

European Colloquium on
Theoretical and
Quantitative
Geography
2021

Manchester
3-5 November 2021

Book of
Extended Abstracts



ECTQG 2021 MANCHESTER
European Colloquium on Theoretical and Quantitative Geography

© Nuno Pinto, Yahya Gamal, Aya Badawy and Harry Odell

Published by Spatial Policy and Analysis Laboratory, Manchester Urban Institute, University of Manchester

Oxford Road, Manchester M13 9PL

ISBN 978-0-9576682-6-3



Scientific Committee

Michael Batty (University College London, United Kingdom)

Denise Pumain (Université Paris I Panthéon-Sorbonne, France)

Mark Birkin (University of Leeds and Turing Institute, United Kingdom)

Roger White (Memorial University Newfoundland, Canada)

Céline Rozenblat (University of Lausanne, Switzerland)

Itzhak Benenson (Tel Aviv University, Israel)

Isabelle Thomas (Université Catholique de Louvaine-la-Neuve, Belgium)

Geoffrey Caruso (University of Luxembourg)

Beniamino Murgante (University of Basilicata, Italy)

Richard Kingston (University of Manchester, United Kingdom)

Chris Brunsdon (Maynooth University, Ireland)

Ann Verhetsel (University of Antwerp, Belgium)

Cecilia Wong (University of Manchester, United Kingdom)

José António Tenedório (Nova University Lisbon, Portugal)

Sarah Lindley (University of Manchester, United Kingdom)

António Néelson Rodrigues da Silva (University of São Paulo, Brazil)

Lei Wang (Nanjing Institute of Geography and Limnology, China)

Helen Zheng (University of Manchester)

António Pais Antunes (University of Coimbra, Portugal)

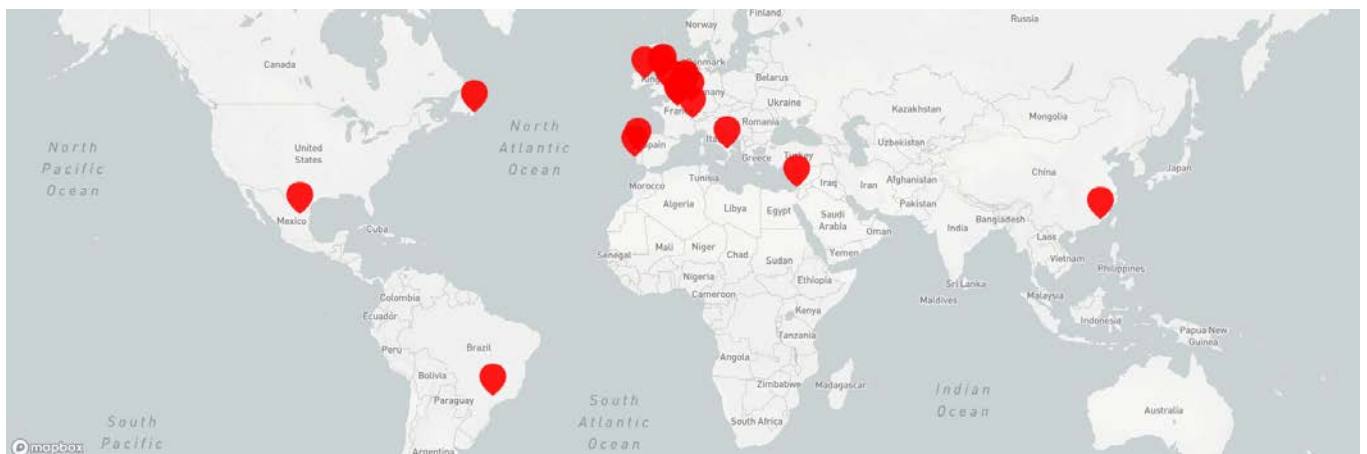
Roberto Ponce-Lopez (TEC de Monterrey, Mexico)

Cyrille Médard de Chardon (Luxembourg Institute for Socio-Economic Research)

Alain L'Hostis (University Eiffel, France)

Federico Amato (University of Lausanne, Switzerland)

Nuno Pinto (University of Manchester, United Kingdom)



ECTQG2021 in Manchester, keeping the ECTQG Community live in times of challenges.

Welcome to the 22nd European Colloquium on Theoretical and Quantitative Geography, ECTQG 2021 Manchester, organised by a group of researchers based at the Spatial Policy and Analysis Lab of the Manchester Urban Institute, at the University of Manchester.

There is a first time for everything, and due to the Covid-19 pandemic, the Colloquium will have in 2021 its first hybrid edition, taking place mainly online.

Worldwide, academic research has been hit by the imposed lockdowns and by the associated economic crisis. Traditional formats of academic meetings were forced to find new ways to ensure that ideas and innovation kept being discussed and shared by academics from around the world online, avoiding the risk of travel.

ECTQG 2021 Manchester raised to that challenge and has been organised with the sole aim of keeping the ECTQG community live during this very hard period of the pandemic. The Manchester Organising Committee in collaboration with the ECTQG Steering Group, faced challenges linked to uncertainties of travelling and available funds to participate in conferences, associated with a very demanding context in the UK Higher Education sector, with a dramatic increase of academic workload.

The Manchester Organising Committee is extremely happy to be able to bring ECTQG 2021 home to the ECTQG community and beyond, bringing together more than 90 accepted submissions by more than 250 authors from all continents in the world. Three great keynote lectures will be delivered by leading academics in theoretical and quantitative geography and urban analytics, along with a workshop on pedagogy of quantitative methods, six special sessions and twelve parallel sessions, all with strong thematic cohesion and interesting research presentations.

We are also very proud of bringing the European and global communities of researchers in theoretical and quantitative geography and urban analytics back to the United Kingdom, even if mainly virtually, in these times of uncertainty and challenges for the British research community.

We are, however, very sad for not welcoming you all in Manchester and at the Department of Planning and Environmental Management as initially planned. The University of Manchester is the home of Alan Turing. The Department is about to start celebrations of its 70th anniversary in 2022, being home of some historical achievements in the field as well as alma matter of some prestigious colleagues in quantitative geography.

Manchester is definitely a wonderful place to organise ECTQG.

We thank the University of Manchester Spatial Policy and Analysis Lab and Institute for Data Science and Artificial Intelligence, as well as the Urban Analytics Programme of The Alan Turing Institute for their support.

We wish to thank the ECTQG community for their resilience in keeping the interest in participating in ECTQG 2021 despite the challenges of the past two years.

We will meet all of you in the virtual spaces of the ECTQG 2021 Manchester.

The ECTQG 2021 Manchester Organising Committee

Nuno Pinto, Yahya Gamal, Aya Badawy, Harry Odell, António Ross, Ron Bar-Ad and Matthew Harrison

Support

ECTQG 2021 Manchester is an event organised with the support of the [Spatial Policy and Analysis Lab](#) of the [Manchester Urban Institute](#), with the support of the Manchester's [Institute for Data Science and Artificial Intelligence](#). ECTQG 2021 has also the valuable support of the [Urban Analytics](#) programme of the [Turing Institute](#).



Manchester Urban Institute
The University of Manchester

Spatial Policy and
Analysis Laboratory



www.mui.manchester.ac.uk/spal
[@ mui@manchester.ac.uk](mailto:mui@manchester.ac.uk)
[@UoM_Spal](#) [@UoMUrban](#)

The Alan Turing Institute



**INSTITUTE FOR DATA SCIENCE
AND ARTIFICIAL INTELLIGENCE**

Gateway to The University of Manchester's Data Science &
Artificial Intelligence activity

Programme

03-Nov

Pre Colloquium Workshop			
10:00	Facilitating quantitative methods pedagogy: Digitalization, innovation, and future needs		
	https://zoom.us/j/99817401229 Session Technical Support: Aya Badawy		
Opening Session			
2:00	https://zoom.us/j/96199348463 Session Technical Support: Ron Bar-Ad		
2:20	Critical Issues in Quantitative Geography	Urban Form	Scaling Laws and Urban Form
	PS_A1	PS_A2	PS_A3
	https://zoom.us/j/91552122661 Session Technical Support: Aya Badawy	https://zoom.us/j/97419958074 Session Technical Support: Yahya Gamal	https://zoom.us/j/93323311053 Session Technical Support: Harry Odell
Keynote Lecture Prof Luís Bettencourt			
5:00	https://zoom.us/j/96622670458 Session Technical Support: Harry Odell		

04-Nov

9:00	Exploration and validation of spatial simulation models	The multitude of spatiotemporal scales in urban systems	Networks	Transport and Urban Form
	SS03_04.1	SS02.1	PS_B3	PS_B4
	https://zoom.us/j/91864397536 Session Technical Support: Aya Badawy	https://zoom.us/j/98057405920?pwd=cUpmSWRGKzRHbHtSVEJFRGdyRHI1dz09C Session Technical Support: António Ross	https://zoom.us/j/93881964088 Session Technical Support: Ron Bar-Ad	https://zoom.us/j/93526715574 Session Technical Support: Yahya Gamal
11:20	Exploration and validation of spatial simulation models	The multitude of spatiotemporal scales in urban systems	Mobility and Distance	Rural to Urban
	SS03_04.2	SS02.2	PS_C3	PS_C4
	https://zoom.us/j/91864397536 Session Technical Support: Aya Badawy	https://zoom.us/j/99254924465?pwd=d1dlQzZV0Npb1lRWVdFY3Z0dmVldz09 Session Technical Support: António Ross	https://zoom.us/j/99983777459 Session Technical Support: Ron Bar-Ad	https://zoom.us/j/96172434092 Session Technical Support: Yahya Gamal
Keynote Lecture Prof Michael Batty				
2:00	https://zoom.us/j/98603869138 Session Technical Support: Yahya Gamal			
4:00	Special Session in honour of Martin Charlton	Energy and Industry	Perception of Urban Space	
	SS10	PS_D2	PS_D3	
	https://zoom.us/j/91830030285 Session Technical Support: Aya Badawy	https://zoom.us/j/96172434092 Session Technical Support: Yahya Gamal	https://zoom.us/j/93379537071 Session Technical Support: Harry Odell	
Celebration Martin Charlton				

05-Nov

9:00	Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance	Climate change mitigation strategies at urban and territorial scales: modelling tools and quantitative impact assessment	Health and Gender	
	SS10.1	SS02	PS_E3	
	https://zoom.us/j/97382220613 Session Technical Support: Harry Odell	https://zoom.us/j/95376416809 Session Technical Support: Ron Bar-Ad	https://zoom.us/j/94120964561?pwd=Vkl6ZEMzTjJlNlU9S2MWhFty9DcWdmLjU09 Session Technical Support: António Ross	
11:20	Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance	Health and Well-being	Diverse topics in quantitative geography	
	SS_10.2	PS_F2	PS_F3	
	https://zoom.us/j/95312241858 Session Technical Support: Harry Odell	https://zoom.us/j/96250677337 Session Technical Support: Ron Bar-Ad	https://zoom.us/j/93695707082?pwd=RmQ3TkZyRVBkeFFBVDdkNXhuZmRaZ09C Session Technical Support: António Ross	
Keynote Lecture Prof Rachel Franklin				
3:00	https://zoom.us/j/96431355959 Session Technical Support: António Ross			
Conclusions and Closing Session				
4:45	https://zoom.us/j/96173707582 Session Technical Support: Ron Bar-Ad			

Detailed Programme

03-Nov			
Pre Colloquium Workshop			
10:00	Facilitating quantitative methods pedagogy: Digitalization, innovation, and future needs		
2:00 Opening Session			
	Critical Issues in Quantitative Geography	Urban Form	Scaling Laws and Urban Form
2:30	1 Sebastien Bourdin and André Torre Geography of contestation: a study on the yellow vests movement and the rise of populism in France	2 Mirjam Schindler Tackling urban air pollution concerns through urban form – Lessons from a GIS-based residential choice model with micro-economic foundation	19 Raphaël Bubloz Convolutional neural networks for building types classification in Switzerland
2:40	4 Anna Dmowska, Tomasz Stepinski and Jakub Nowosad A spatio-racial pattern in the US cities	28 Bowen Zhang, Chen Zhong and Qili Gao A Novel Regionalisation Algorithm for Redrawing Cities' Functional Boundary within Mega-city Region	21 Roger Vazquez and Roberto Ponce-Lopez A data-driven methodology to identify urban form and characterize its functionality
3:10	6 Alicia Bianchi, Giovanni Fusco and Karine Emsellem The real estate ads, a new data source to understand the social representation of urban space.	26 Olivier Bonin and Arthur Benichou Stability of urban forms: modelling the emergence of collective behaviour in residential trajectories	31 Paul Kilgarriff, Remi Lemoy and Geoffrey Caruso A new city definition from the radial homothetic scaling law
3:30	50 Jonny Huck, Duncan Whyatt, John Dixon, Gemma Davies, Brendan Sturgeon, Neil Jarman and Dominic Bryan An Empirical Agent Based Modelling approach to Evaluating Intergroup Contact Theory	14 Michele Tirico, Stefan Balev, Antoine Dutot and Damien Olivier Morphogenesis of street networks. A reaction-diffusion system for self-organized cities	32 Giampiero Lombardini Power law, scaling and the "degrowth geographies". The case of Italy
3:50	70 Pakinam Hassan Perceptions towards improving the participatory urban planning in Alexandria, Egypt	77 Hanae El Gouj, Paul Bouman, Liubov Tupikina and Claire Lagesse Road network characterization to understand city evolution	Discussion
4:40	Discussion	Discussion	
5:00 Keynote Lecture Prof Luis Bettencourt			

04-Nov

Exploration and validation of spatial simulation models		The multitude of spatiotemporal scales in urban systems		Networks		Transport and Urban Form		
9:00	3 Dani Broitman and Yakov Ben-Haim	Forecasting Residential Sprawl Under Uncertainty: An Info-Gap Analysis	54 Kerry Schiel, Geoffrey Caruso and Mirjam Schindler	40 years of sprawl in the USA – Scale and Geographic Effects	8 Aleksey Ogulenko, Itzhak Benenson, Marina Toger, John Östh and Alexey Siretskiy	The fallacy of the closest antenna: Towards an adequate view of device location in the mobile network	15 Romain Loup, Guillaume Guex and François Bavaud	Estimating individual journeys from collective transfers at bus stops
9:20	72 Juste Raimbault and Denise Pumain	How validation through model exploration empowers theories of spatial complexity: example of urban systems	23 Pierre Frankhauser and Olivier Bonin	Multifractal geometry for integrating several challenges of sustainable metropolitan planning	34 Natalia Zdanowska and Robin Morphet	The Northern Cities of the United Kingdom: firm ownership links, productivity and interconnection	24 Efrat Blumenfeld-Lieberthal, Nimrod Serok and Shlomo Havin	Evaluation and prioritization of urban traffic bottlenecks
9:40	52 Jingyan Yu, Alex Hagen-Zanker, Naratip Santittisadeekorn and Susan Hughes	Bayesian calibration of cellular automata urban growth models from urban genesis onwards	30 Thomas Louail and Marc Barthelemy	A dominance tree approach to systems of cities	57 Nikita Sinitsyn	"VKontakte" social network behavior patterns: case study of Vologda region	75 Anna Ibraeva, Gonçalo Correia, Cecília Silva and Antonio Pais Antunes	Post-Metro Mode Choice Changes: from Macro- to Micro-Analysis
10:00	73 Juste Raimbault and Michael Batty	Sensitivity analysis of the MATSim transport model	37 Rémi Lemoy, Geoffrey Caruso and Paul Kilgariff	Radial analysis and scaling of urban land use and sprawl	59 Robin Cura	Leveraging GeolPs as a source of understanding and comparing infra-regional spaces	33 David Burg	Modelling urban population density of employment centres: Israel case study
10:20	90 Jasper van Vliet, Bep Schrammeijer, Yuan Wang and Niels Debonne	What metrics don't reveal: process accuracy and the evaluation of land use models	38 Axel Pecheric, Rémi Lemoy, Marion Le Texier and Geoffrey Caruso	Geovisualization of artificial land use in European cities in 2006, with urban scaling laws	40 Roberto Ponce López, Diego Fabián Lozano-García, Sebastian Fajardo, Alicia Landin Quiros and Pedro Pachecho	The economic cost of urban sprawl approximated by the cost of maintaining road infrastructure in Monterrey	36 Yahya Gamal and Nuno Pinto	Decision Making and Land Market Preferences: A Land Use Agent Based Model
10:40	Discussion		Discussion		Discussion		Discussion	

Exploration and validation of spatial simulation models		The multitude of spatiotemporal scales in urban systems		Mobility and Distance		Rural to Urban		
11:20	67 Aya Badawy, Nuno Pinto and Richard Kingston	Exploring methods for simulating urban negotiations	88 Jan Friedrich, Seraphim Alvanides and Janka Lengyel	On the modeling of the interdependences of spatial scales in urban systems	56 Golan Ben-Dor, Ido Klein, Aleksey Ogulenko, Eran Ben-Elia and Itzhak Benenson	Congestion and Parking Prices in a Shared Automated Vehicle Future: A Case Study of Jerusalem	18 Fanny Sarkissian, Jean-Philippe Antoni and Teddy Loyer	Settling the challenges of a sustainable city through the dynamic representation of an urban agriculture system
11:40	58 Pendula Ferdinand, Bo Pieter Johannes Andree and Eric Koomen	Validating a global urbanisation model	47 Sofia Fellini, Pietro Salizzoni and Luca Ridolfi	Structural fragility of cities to airborne releases	20 Fabian Wenner, Johannes Moser and Alain L'Hostis	Which Regions Benefit from Rail Accessibility? Germany 1990-2030	5 Timur Gusakov	Types of Crimean Rural Settlement
12:00	78 Juste Raimbault	Benchmarking road network growth models	89 Janka Lengyel, Stéphane Roux, François Sémecurbe, Stéphane Jaffard and Patrice Abris	Defining roughness and intermittency within metropolitan regions	7 Satoru Kumagai, Kenmei Tsubota, Toshitaka Gokan, Kazunobu Hayakawa, Ikumo Isono and Souknilanh Keola	Economic and Environmental Impacts of Dedicated Freight Corridors in India	71 Anasua Chakraborty, Jacques Teller and Hichem Omrani	Impact of Static and Dynamic Drivers on Urban Densification
12:20	49 Jingyan Yu, Alex Hagen-Zanker, Naratip Santittisadeekorn and Susan Hughes	Spatial and temporal transferability of urban growth models - A data-driven framework of multi-area, multi-period calibration, growth mode clustering, and scenario development	62 Haocheng Sun, Michael Batty and Chen Zhong	Measuring the Polycentricity based on urban and intercity transportation networks in Greater Bay Area: a cross-scale method in the context of Node-Place model	29 Alain L'Hostis	Introducing "Contextual Metrics" as a mathematical definition for a comprehensive approach of geographical distances	10 Hiroyuki Usui	A bottom-up approach for delineating urban areas: Simultaneous optimisation of criteria for building interval and built cluster size minimising the connection cost of buildings by roads
12:40	Discussion		43 Tomasz Stepinski and Mahmoud Saeedimoghaddam	Multiplicative random cascade models of urban structures	87 Didier Josselin	An extension of the Tobler's law? Influence of the metric to compute the distance between an optimal facility and its demand points.	Discussion	
1:00	Discussion		Discussion		Discussion		Discussion	

2:00 Keynote Lecture Prof Michael Batty

Special Session in honour of Martin Charlton		Energy and Industry		Perception of Urban Space		
4:00	42 Chris Brunson	Whats in a Name? A Spatial Analysis of British Public House Names	79 Fernando Benitez-Paez and Mark Birkin	Studying social inequality in urban energy use in the UK	85 Roel Yosifof and Dafna Fisher-Gewirtzman	Simulating pedestrian's visual perception: Dynamic 3D quantitative analyses
4:20	55 Alexis Comber, Paul Harris and Chris Brunson	A spatial analysis of prescribing data: revisiting the 'Geographical Analysis Machine	86 Marion Nativel	Integrating stated preferences, landscape metrics and attitudes to assess visual impact of wind turbines.	83 Ron Bar-Ad, Geoffrey Caruso, Nuno Pinto and Markel Vigo	Perceived Variables in Virtual Reality: A Methodological Analysis
4:40	91 Steve Carver, Chris Brunson and Alexis Comber	Training Sets and Train Sets: How organs, pubs and railways inter-relate	53 Andrei Korotkov, Artem Kushlevich, Dmitry Rusakov, Roman Menyalschikov, Nikita Sinitsyn and Mikhail Elyutin	Wood trade regions of Russia	25 Osher Ann Lapid and Efrat Blumenfeld Lieberthal	On Tel Aviv's Roofs
5:00	Discussion		82 Denisa Naidin and Sabine Dörny	The Luxembourgish Space Industry – Evolving Between Local and Global Forces	80 Anna Zanchetta and Mark Birkin	The future of high streets: investigating post-pandemic scenarios in the UK
5:20	Discussion		Discussion		Discussion	

7:00 Celebration Martin Charlton

05-Nov

Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance		Climate change mitigation strategies at urban and territorial scales: modelling tools and quantitative impact assessment		Health and Gender		
9:00	41 Junwen Lu and Suhong Zhou	Does high-density living facilitate non-motorized travel for every resident? Examining the moderating effect of individual daily time-space arrangement in Guangzhou, China	60 Giulia Benati and Federico Martellozzo	Decrease of the value of the regulating ecosystem service of CO2 capture by urban trees due to the increase of soil consumption: a quantification of the phenomenon in Rome	81 Stefania Bertazzon, Alison Hanes, Rizwan Shahid and Isabelle Couloigner	The first to be hit: spatio-temporal analysis of first-wave COVID-19 diffusion in two regions of northern Italy
9:20	39 Qili Gao, Chen Zhong and Yang Yue	How activity pattern associates with income status? Evidence from transit smart card data and AI methods	69 Sabrina Lai	Synergies and trade-offs in ecosystem services' provision: identifying spatial bundling in Sardinia.	63 Ido Klein, Golan Ben-Dor, Aleksey Ogulenko, Avner Saadoun and Itzhak Benenson	Spatially-explicit modeling of the SARS-CoV-2 transmission between population groups in Jerusalem
9:40	48 Suhong Zhou, Linsen Wang, Jiangyu Song, Junwen Lu and Zhong Zheng	The Role of Mobility on Daily Spatial Quality Exposure Equality. Empirical evidence from Guangzhou, China	84 Marlène Boura and Geoffrey Caruso	The Urban Carbon Budget (UCB) Model: a high resolution spatio-temporal model of CO2 emissions, dispersion and sequestration in cities	68 Léo Taillandier	A multivariate modelization to catch bicycling spatial heterogeneity within French functional urban areas : the example of Toulouse.
10:00	11 Mingxiao Li, Wei Gao, Wei Tu and Bichen Fang	A congenerous regional simulation model for land use-population-economy changes in the Greater Bay Area	66 Estelle Mennicken, Geoffrey Caruso, Rémi Lemoy, Philippe Gerber and Alex Hagen-Zanker	European urban cores under pressure: quantifying the congestion of trips into and out from city centres as a function of population size	16 Bochu Liu, Michael Widener and Lindsey Smith	Partners' non-food-related activities and arrangement of food-related household labor: a case study of coupled adults in Toronto, Canada
10:20	35 Apeksha Tare, Merten Nefs, Eric Koomen and Erik Verhoef	Spatial preferences of logistics development	17 Yufei Wei, Geoffrey Caruso and Rémi Lemoy	Centrality and city size effects on NO2 ground and tropospheric concentrations within European cities	27 Lindsey Smith, Bochu Liu and Michael Widener	Comparing geographies of care in contrasting neighbourhoods in Toronto, Canada
10:40	Discussion		Discussion		Discussion	

Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance		Health and Well-being		Diverse topics in quantitative geography		
11:20	9 Zifeng Chen and Anthony Yeh	Delineating Functional Urban Areas in Chinese Mega City Regions Using Fine-grained Population Data and Cellphone Location Data: A Case of Pearl River Delta	12 Marie Ramel-Delobel, Thomas Coudon, Lény Grassot, Delphine Praud, Béatrice Fervers and Pietro Salizzoni	Exposure to air pollution and breast cancer risk: taking daily mobility into account in epidemiological studies	45 Yongxin Yang	The Comparative Study of TOD in Metro Station Areas of Guangzhou and Shenzhen Using an Extended Node-Place Model
11:40	62 Haocheng Sun, Michael Batty and Chen Zhong	Measuring the Polycentricity based on urban and intercity transportation networks in Greater Bay Area: a cross-scale method in the context of Node-Place model	13 Madeleine Guyot, Ingrid Peigrims, Raf Aerts, Eva M. De Clercq, Isabelle Thomas and Sophie O. Vanwambeke	Nonresponse in the analysis of the association between mental health and the urban environment in Brussels	61 Tsutomu Tsuboi	Indian Traffic Congestion Model by Shock Wave
12:00	46 Jinzhou Cao, Wei Tu, Chen Zhong, Qili Gao and Qingquan Li	The evaluation of intra-urban developments through urban scaling indicators: empirical evidence from the Greater Bay Area, China	74 Stefania Bertazzon, Caitlin Calder-Bellamy, Rizwan Shahid, Isabelle Couloigner and Yui Tanagi	Preliminary spatial analysis of childhood asthma, air pollution, and socioeconomic status in Calgary, Canada	65 Xiang Ye, Renzhong Guo, Siru Li, Ding Ma and Zhigang Zhao	Preliminary discoveries on the joint impacts of the modifiable areal unit problem (MAUP) and commission error on the classical linear regression model (CLRM)
12:20	76 Zhiyuan Li and Juste Raimbault	High-speed rail development and access equity: the case of Guangdong province, China	64 Lény Grassot, Yohan Fayet, Delphine Praud, Béatrice Fervers and Thomas Coudon	Which local historical information as a proxy for environmental exposure to air pollution in France throughout the 20th century?	44 Angelos Mimis and Kostas Bithas	A geophysical model for estimating CO2 balance in tree crops in Mediterranean region
12:40	51 Daniela Vanessa Rodriguez Lara and António Nelson Rodrigues da Silva	A system dynamics model of urban mobility resilience when exposed to fuel-related threats	22 Guoqiang Wu, Alison Heppenstall and Nik Lomax	Estimating health and social outcomes at small areas in Great Britain by generating and utilising synthetic population data	92 António Ross, Neil Walton and Nuno Pinto	Identifying trip purpose from a dockless bike-sharing system in Manchester
1:00	Discussion		Discussion		Discussion	

3:00	Keynote Lecture Prof Rachel Franklin				
4:45	Conclusions and Closing Session				

Keynote Speaker

Prof Luis Bettencourt

Inaugural Director of the
Mansueto Institute for Urban Innovation and
Professor of Ecology and Evolution at the College
University of Chicago



Luís M. A. Bettencourt is the Inaugural Director the Mansueto Institute for Urban Innovation at the University of Chicago and Professor of Ecology and Evolution at the College. He is also Associate Faculty of the Department of Sociology and External Professor at the Santa Fe Institute. He conducts interdisciplinary research on complex adaptive systems in biology and society and leads research and education programs in Urban Science and Sustainable Development. His research focuses on the identification, modeling and theory of the systemic processes and properties that create and sustain cities. This work uses interdisciplinary concepts together with many different forms of evidence and data to create new theoretical and methodological syntheses that account for the complex properties of urban environments and produce new science-based solutions. This work also involves partnerships and collaborations with international networks of researchers, local governments and NGOs to understand and systematize urban knowledge, and to foster processes of sustainable development. His work is well-known academically and has been influential in developing new theory and new creative approaches to challenges of urbanization worldwide.

3 November 2021, 5.00 PM (GMT)

Keynote Speaker

Prof Michael Batty

Professor of Planning
Chairman of the Centre for Advanced Spatial Analysis (CASA)
University College London



Michael Batty is Bartlett Professor of Planning at University College London. He is Chair of the Centre for Advanced Spatial Analysis (CASA) and a Turing Fellow in the Alan Turing Institute. He was Professor of Town Planning at the University of Wales in Cardiff in the 1980s where he acted as Dean of the Faculty of Environmental Design. From 1990-1995, he was Director of the National Center for Geographic Information and Analysis (NCGIA) at SUNY-Buffalo before he set up CASA at UCL. He has worked on computer models of cities and their visualisation since the 1970s and his recent publications *Cities and Complexity* (2005), *The New Science of Cities* (2013), *Inventing Future Cities* (2018), all published by The MIT Press, and the edited book *Urban Informatics* (Springer 2021) reflect this focus on the applications of digital technologies to urban planning. He is a Fellow of the British Academy (FBA), the Royal Society (FRS), and the Academy of Social Sciences (FAcSS). He was awarded the CBE in the Queen's Birthday Honours List in 2004, the Lauréat Prix International de Géographie Vautrin Lud ('Nobel de Géographie') in 2013, the Gold Medal of the Royal Geographical Society in 2015, and the Gold Medal of the Royal Town Planning Institute in 2016.

4 November 2021, 2.00 PM (GMT)

Keynote Speaker

Prof Rachel Franklin

Professor of Geographical Analysis
Centre for Urban and Regional Development Studies (CURDS)
Newcastle University



Rachel Franklin is Professor of Geographical Analysis in the Centre for Urban and Regional Development Studies (CURDS) at Newcastle University. Prior to Newcastle, she was at Brown University in the U.S., where she was Associate Director of Brown's Spatial Structures in the Social Sciences (S4) initiative. She currently holds visiting appointments in Population Studies at Brown and in Social Sciences at the Gran Sasso Science Institute in Italy. She is a Fellow of the Alan Turing Institute in the UK and is also current editor in chief of Geographical Analysis. Rachel's research interests are in the sources and impacts of demographic change as it occurs at multiple spatial scales, and in novel forms of data and analysis to identify and characterise these changes. She works at the regional and local scales to understand how best to characterize or measure the populations of places; how location and scale are related to demographic change; and how migration, especially internal, affects demographic composition. She is especially interested in how we use data and statistics to understand what sorts of people are located where, how this changes over time, and what this means for our understanding of spatial inequality.

5 November 2021, 3.00 PM (GMT)

Pre-Colloquium Workshop

3 November 10:00-11:30 am (GMT)

Facilitating quantitative methods pedagogy: Digitalization, innovation, and future needs

Proponents

Jane Bunting, Oliver Gronz and Cyrille Médard de Chardon

COVID-19 has only accelerated the digitalization of teaching. Beyond the known draw backs, there are advantages and opportunities available in blended learning as well.

This workshop will moderate discussions on existing challenges to quantitative geography teaching, recent innovations, and the the future and emerging skill requirements for students and teaching.

The meeting will use a variety of digital platforms to facilitate discussion. We welcome participation by researchers of all levels.

This special session is part of the DigiLEGO project (<https://project.digilego.eu>), funded through the EU Erasmus+ Strategic Partnership, which aims to support the geography HE community by enhancing the capacity of HE teachers through collaborative creation of high-quality, highly shareable open educational resources (OERs) for discipline-specific methods training, training in digital and blended methods delivery, and training resources on content creation and digital pedagogies.

Parallel Session PSA1

3 November 2:30-5:00 pm (GMT)

Critical Issues in Quantitative Geography

Geography of contestation: a study on the yellow vests movement and the rise of populism in France

Sebastien BOURDIN; André TORRE²

¹EM Normandie Business School, Métis Lab

^{2**}University Paris-Saclay, INRAE, AgroParistech

Keywords: Contestation, discontent, vote, electoral geography, left-behind places, Yellow-vests

Introduction

There is a new trend in the literature to analyze the extent to which the places and the types of territories where people live influence electoral behaviour. In particular, researchers have been interested in the geography of discontent (McCann, 2016 and 2020; Dijkstra et al., 2020) and the voting of populations located in places which do not matter (Rodríguez-Pose, 2018). The study of recent events such as Brexit (Los et al., 2017; Abreu & Öner, 2020), the American elections (Rodríguez-Pose et al., 2020), and the European elections (Di Matteo & Mariotti, 2020) highlights a rise in extreme voting in places which do not matter, namely rural territories, peripheral areas, urban districts in difficulty, etc.

Therefore, it is hypothesized that the characteristics of these marginalized territories affect the behaviour and attitudes of the people who live there (Gordon, 2018). Simultaneously, the reasons linked to their economic decline are far from explaining why populist electoral behaviours are observed in several places. In political science, it is highlighted that these anti-establishment votes were a form of contestation of the political power which had been in place for several years, alternating between the traditional left and right parties. As a result, a form of disenchantment regarding the 'system' and globalization appeared (Rodrik, 2019), which the high number of protest votes in the ballot boxes confirmed.

In France, the rise of populism and the failure of the traditional parties during the 2017 presidential election call the relevance of the vote's traditional determinants into question. The progressive polarisation towards extreme voting is a striking feature of the political dynamics which has been at work since the early 2000s, accompanied by a regular decreasing influence of the traditional moderate right (Republicans) and left (Socialists) parties. At the same time, the social media's influence is also getting more and more preponderant in the orientation of the voters' political choices and political mobilization (Jost et al., 2018). In particular, the Yellow Vests' social movement was born on Facebook.

In 2018, the French society was shaken by this large-scale protest movement. At first, the movement brought together angry motorists protesting against the rising fuel prices and the government's decision to reduce authorized speed limits on secondary roads. But it soon turned into a general protest against the government's policy. Its participants were thus invited to block traffic as close as possible to their homes with, as early as the first Saturday of the demonstration, many new roadblocks, especially at roundabouts, but also big protestations in the major French cities. If the increase in the tax on petroleum products was the movement's triggering factor, it does not seem to be the only explanation. Several protest movements took place at the beginning of Emmanuel Macron's five-year term of office, although they didn't succeeded in getting more unified. In addition to the Government's political choices, fundamental questions have been raised about public policies for several decades, in a context of falling public spending and growing inequality, with significant territorial repercussions. Thus, the mobilization took place in this context of a deep questioning of liberalism and social democracy in many countries (Pappas, 2019).

Based on the CEVIPOF (French Center for Political Science) Barometer of Confidence survey, Algan et al. (2019) investigated the socio-demographic characteristics of the Yellow Vests' supporters. The study showed that the majority of the voters are extreme right's partisans, extreme left's ones, or abstainers. It should be noted that the survey focuses on the political position of those supporting the

Yellow Vests movement, without dealing directly with the demonstrators themselves. Moreover, this analysis, carried out by researchers in political science and electoral sociology, does not include issues regarding the places' characteristics. Another paper wrote by Grossman (2019) points out that most studies remain qualitative and do not provide any synoptic view of the Yellow Vests' real identity and the places they live in.

By analyzing the Yellow Vests movement and comparing it to the characteristics of the French populist votes, our article contributes to the literature on the question of the geography of discontent, disenchantment, resentment, and contestation. This movement can be seen as a behaviour of disenchantment; in other words, voters no longer bother to go to the polls because they no longer trust politicians or are no longer part of the political offer (Kostelka, 2017). They express their opinions (Hirschman, 1970) in a different way, through conflictual behaviours instead of expressing a kind of loyalty or trust by voting. Voting for the extreme right party, which is a very widespread practice in French politics, can be rather considered as a different behaviour: a discontent expressed by a vote (Davezies et al., 2013; Rodríguez-Pose, 2018; Guilluy, 2019). Even if both behaviours express protest, we wonder if it is the same socio-economic and territorial factors which underlie them, and if these different types of behaviours are driven by the same types of people.

We tested two hypotheses: that the Yellow Vests are the expression of the votes for populist parties, and that the characteristics of a territorial decline or stagnation have a parallel influence on these two movements. When we compare the determinants of the Yellow Vests and those of the populist votes, we cannot confirm our hypotheses. In other words, populist supporters cannot be confused with activists of the Yellow Vests movement, and their sensitivity to territorial dimensions differs as well. We highlighted that spatial variables better explain the geography of the Yellow Vests, whereas the French popular votes are explained by a mix of spatial variables and variables traditionally used in electoral sociology (socio-demographic and economic).

Mainly, the French geography of contestation sends a signal which raises serious concerns about the territories' dynamics. The first one is that of the driving power of the metropolises over the whole country. Like the rest of the world, France is subject to the logic of metropolization, tending to concentrate values in large urban areas. Economic geography has largely highlighted the close links between globalization and metropolization, and the effects of diffusion and redistribution of wealth in large cities. As a result, people living outside of these metropolises do not feel that they benefit from globalization, or even suffer from its consequences. With housing prices soaring in these large cities, some people have to live in the periphery or rural areas, increasing their distance from public services and their transportation costs to get to their jobs.

A large part of the contestation is also related to the development of a movement, the new public management (Hammerschmid et al., 2019 ; Chouraqui, 2020), presented as a miracle recipe for copying the public sector organisation on the imperatives of profitability. In France, and generally speaking in developed countries, the search for public policy performance has generated a decrease in public services in the territories which were already in difficulty. Therefore, it is not surprising to see that parts of the population that no longer benefits from these services – or whose quality has deteriorated – no longer believe in the government and assert their claims by protesting.

This brings us back to the third question, that of territorial governance. The recent French territorial reforms focus on the institutional architecture of large regions and cities at the expense of the construction of local policies. The Yellow Vests' revolt has highlighted the need for more participative democracy: rethinking the territorial governance of economic development projects is a key issue. As Torre and Traversac (2011) pointed out, the main resource of a territory is the intensity of the links between all the stakeholders, public but also private ones, including beyond their immediate environment. The contemporary economy is relational. This is as true for the functions which animate the heart of the world's major metropolises as it is for the initiatives born in rural territories (Boschma and Iammarino, 2009). Consequently, the question of territorial governance is

becoming more and more central. To be effective, it must include all the stakeholders in the territory, including citizens, who are too often forgotten and left aside (Stead, 2014). It is precisely this feeling of being ignored and the lack of consideration for the citizens' expectations which generates resentment and encourages people to vote for anti-system parties or to sep up a revolt like the Yellow Vests movement.

A spatio-racial pattern in the US cities

Anna DMOWSKA¹; Tomasz F. STEPINSKI²; Jakub NOWOSAD¹

¹Institute of Geoecology and Geoinformation, Adam Mickiewicz University, Poznan, Poland

B. Krygowskiego 10, Poznan, 61-680, Poland

+48618296344, dmowska.amu.edu.pl (correspondent author)

²Space Informatics Lab, University of Cincinnati

215 Braunstein Hall, Cincinnati OH, 45221, USA

Keywords: residential segregation, racial diversity, pattern analysis

1. Introduction

Traditionally, racial distribution is studied using two separate concepts: racial diversity and residential segregation. Racial diversity characterizes the relative ethnoracial heterogeneity of the population (White, 1986) and is measured by entropy of racial composition. Residential segregation is defined as the degree to which two or more groups live separately from one another (Massey and Denton, 1988). In the multiracial context, it is assessed by the information theory index H (Reardon and Firebaugh, 2002). H measures the difference between the diversity (entropy) of the city and the weighted average of diversities of individual city' subdivisions, expressed as a fraction of the total diversity of the city (Massey and Denton, 1988; White, 1986; Reardon and Firebaugh, 2002). The segregation measure is dependent on the choice of aggregation level of city' subdivisions (the finer the subdivisions, the larger segregation) and their racial composition. The spatial configuration of city' subdivisions is not considered. Consequently, residential segregation, a spatial concept, is routinely analyzed and presented in a non-spatial manner. Such an approach is dictated by the form of population data that are commonly available in tabular form, listing the racial composition for each city's subdivision.

We introduce a Racial Landscape method (RL) – a completely re-imagined, pattern-based approach for analyzing and visualizing residential segregation and racial diversity (Dmowska et al., 2020). Our method addresses many limitations of the current approaches. Here we give an overview of our method and present the result of applying the RL method to 26 urban areas in the United States to analyze and visualize the racial pattern in 2010 at three different spatial scales (1.8 km, 3.6 km, 7.2 km).

2. Racial landscape method

In the RL method, the racial composition is represented by a high-resolution raster grid. Each cell in such a grid contains only inhabitants of a single race and is described by two attributes: race category and population density. Such a grid consists of many large and small patches (racial enclaves) formed by adjacent raster grid cells having the same race category. Thus, the distribution of racial enclaves creates a specific spatial pattern (Fig. 1C).

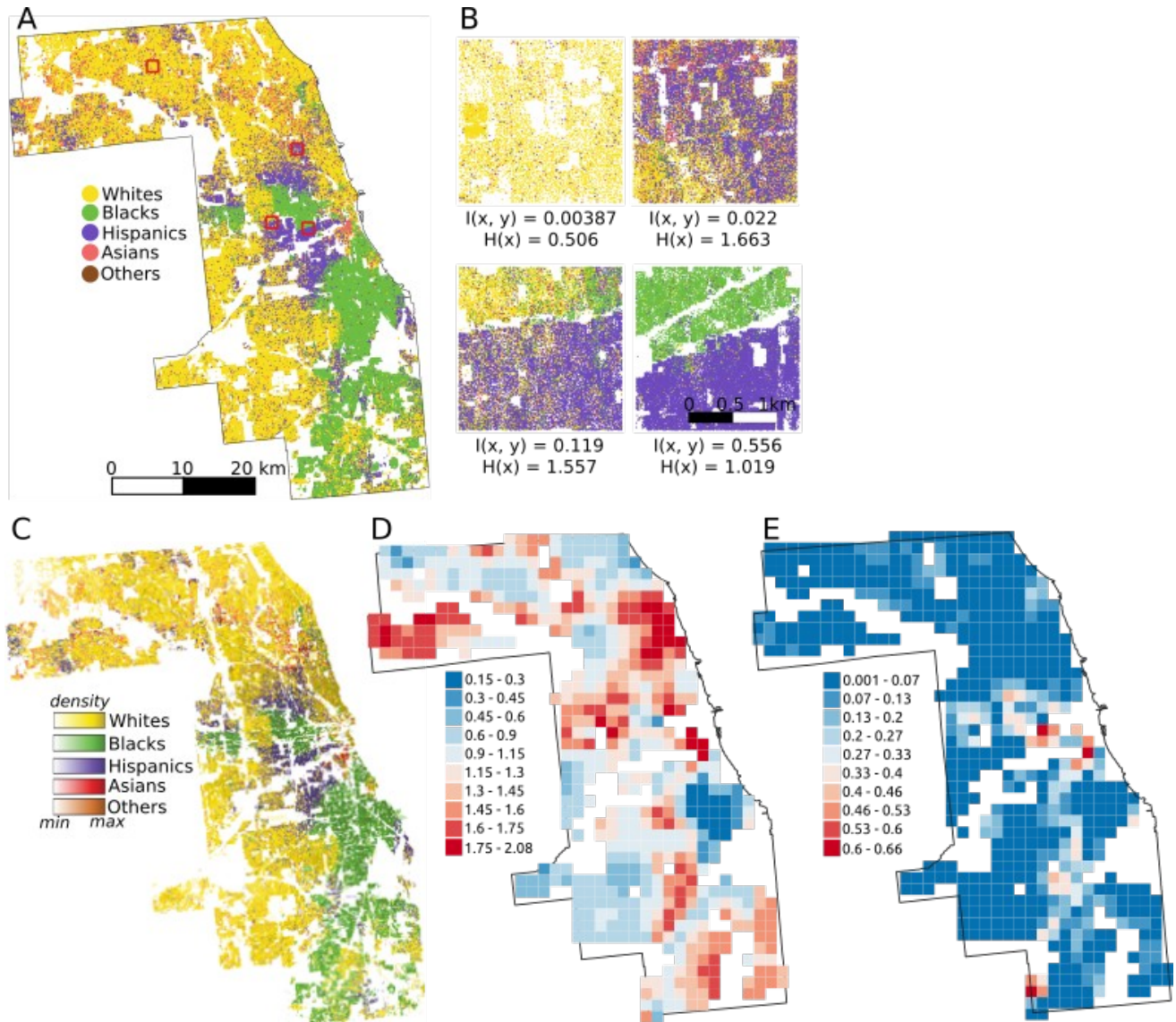


Figure 1: (A) A spatio-racial pattern in Cook County, IL, in 2010 under the assumption of constant population density. (B) Pattern magnified to the extents of four small regions selected from different locations in Cook County. Local values of $H(x)$ (diversity) and $I(x; y)$ (segregation) are given. (C) A racial landscape map illustrating a spatio-racial pattern in Cook County, IL in 2010. Each cell has two attributes: racial category and population density. (D) Racial diversity map measured on spatial scales of 3.6 km. (E) Racial segregation map measured on spatial scales of 3.6 km.

Figure 1 shows an example of a spatio-racial pattern in 2010 in Cook County, IL that includes the city of Chicago. In the field of landscape analysis, the spatial pattern can be quantified using a co-occurrence matrix. The co-occurrence matrix is calculated by moving through each cell, looking at

its category, and counting how many neighbours of each category our focus cell has. The result is a two-way table of the size $K \times K$ (K – number of categories) summarizing cells' adjacencies. We modify a co-occurrence matrix, so each pair of the cells contribute the average population densities instead of 1. The inclusion of population density is necessary to assess residential segregation correctly. We called a modified co-occurrence matrix an exposure matrix. The normalized exposure matrix can be further summarized using the Information Theory metrics (marginal entropy, conditional entropy, and mutual information). Nowosad and Stepinski (2019) proved that two metrics (marginal entropy and mutual information) are sufficient to quantify landscape pattern. Marginal entropy is a measure of racial diversity (the larger marginal entropy, the larger diversity). Mutual information is a measure of residential segregation (the larger mutual information, the larger same-category areas).

The Racial Landscape method (RL) does not require any subdivision for assessing the segregation and diversity of the entire city/region. The calculation can also be performed for any specific spatial scale. In such a case, the city is divided into local patterns (squares with a specific spatial scale, i.e., 1.8 km, see Fig. 1B), and for each local pattern, marginal entropy and mutual information are calculated. The results can be presented in the form of racial diversity (Fig. 1D) and segregation (Fig. 1E) maps for any specific spatial scales. The results can also be summarized as a table listing the metric (entropy, mutual information) for the entire city and selected spatial scales. The high-resolution grid (the input data to the RL method) can be visualized as easy to understand racial map that conveys the racial character of a city (Fig. 1C).

The RL method is implemented in the R package – *raceland* (<https://cran.r-project.org/package=raceland>).

3. Results

Figure 2 shows the diversity-segregation plot with the values of entropy and mutual information for each of the 26 urban settings at the scale of 1.8 km, 3.6 km, 7.2 km and the whole area. The values for the specific scale are an average value of entropy/mutual information calculated from all local patterns. The diversity-segregation plot illustrates that the value of entropy and mutual information increase as the scale increase. The values of the mutual information calculated for the urban settings (without dividing it into the local patterns) are much higher than at the local scale. Selected urban settings significantly differ in respect to the diversity and segregation level.

Based on the diversity/segregation values, the urban settings can be divided into few groups:

1. mixed pattern characterized by the low level of segregation (A) dominated by one racial group with the small share of the other group (Knox County (22), El Paso County (26)); (B)

- constituted of two racial groups (Phoenix AZ (16), Bexar County TX (20)) and (C) with a significant share of three or more racial groups (Sacramento CA (21), Las Vegas NV (23), San Francisco CA (18), San Jose CA (19))
2. pattern characterized by a medium level of segregation with a significant share of (A) two groups (St. Louis County MO (7), Shelby County TN (8), District of Columbia (6), Philadelphia County PA (3)), (B) three or more racial groups (Atlanta GA (9), Harris County TX (15), Los Angeles CA (13), New York NY (5));
 3. clumpy pattern constituted of two or more racial groups inhabiting different parts of the area indicates urban areas with a higher level of segregation and diversity (Cook County IL (1), Wayne County (2)).

Examples of urban settings representing different kinds of patterns are shown in Figure 3.

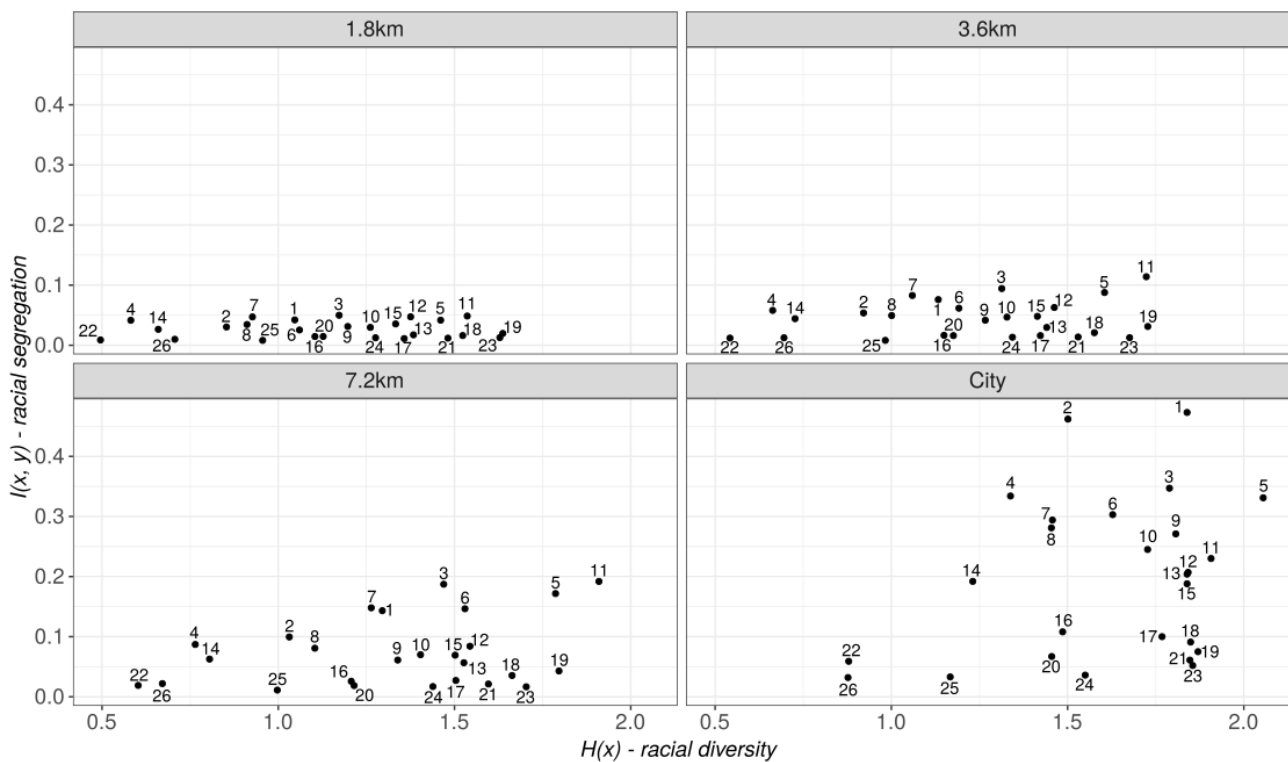


Figure 2: Diversity-segregation plot. The numbers in the plot indicates analysed urban settings: 1 – Cook County, IL; 2 – Wayne County; 3 – Philadelphia County, PA; 4 – Jefferson County, AL; 5 – New York, NY; 6 – District of Columbia, DC; 7 – St. Louis, MO; 8 – Shelby County, TN; 9 – Atlanta, GA; 10 – Miami, FL (UA); 11 – Suffolk County, MA; 12 – Dallas County, TX; 13 – Los Angeles, CA (UA); 14 – Hamilton County, OH; 15 – Harris County, TX; 16 – Phoenix, AZ (UA); 17 – San Diego, CA (UA); 18 – San Jose, CA (UA); 19 – San Francisco County, CA; 20 – Bexar County, TX; 21 – Sacramento, CA (UA); 22 – Knox County, TN; 23 – Las Vegas, NV (UA); 24 – Seattle, WA (UA); 25 – Salt Lake City, UT (UA); 26 – El Paso, TX (UA).

4. Conclusion

We introduce a Racial Landscape method – a new, pattern-based approach to assess the level of residential segregation and racial diversity that addresses the limitation introduced by the current approaches. Here, racial segregation is a measure of clumping (larger same-color enclaves larger segregation), calculated based on high-resolution grids without using city’s subdivisions. The results can be presented for the analysed area or for any specific racial scale in the form of the map or quantitative summaries listing the measure of segregation and diversity. The examples of applying this method are illustrated using 26 urban settings in the US.

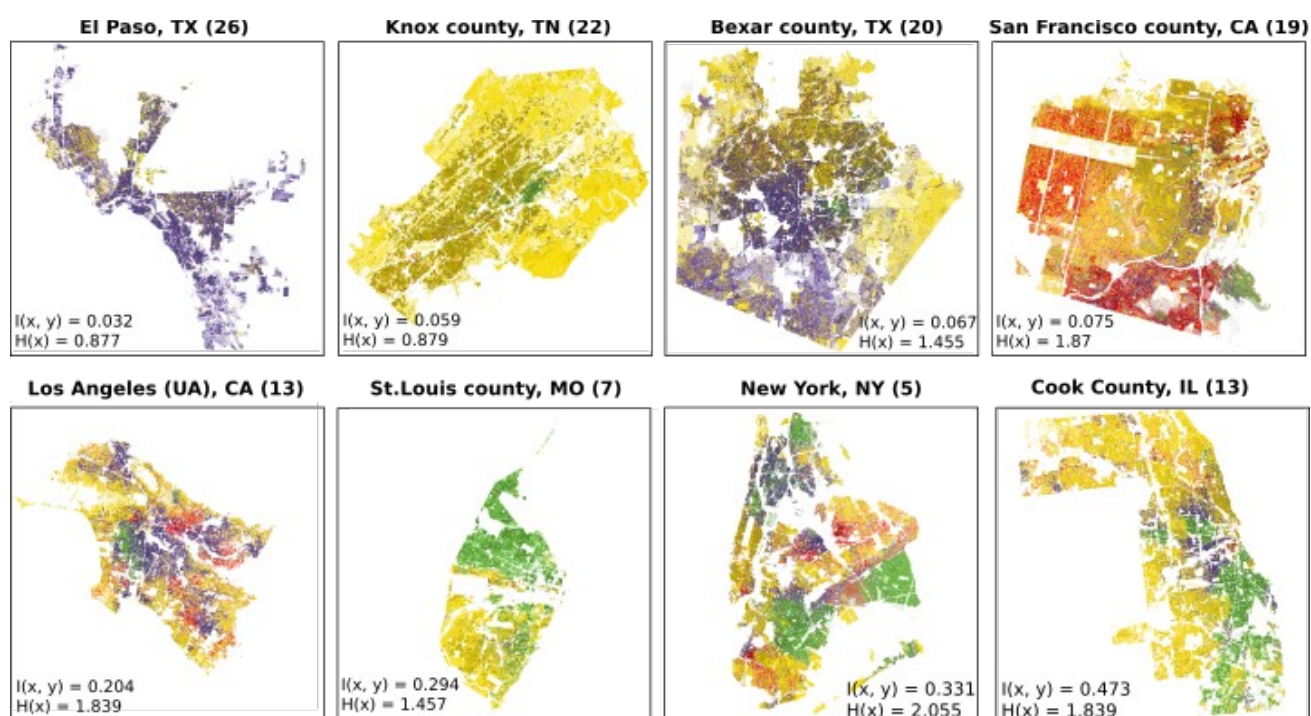


Figure 3: Examples of urban settings representing different kinds of patterns. The number in the brackets are the same as in the diversity/segregation plot.

References

- Dmowska, A., Stepinski, T. F., & Nowosad, J. (2020). Racial Landscapes—a pattern-based, zoneless method for analysis and visualization of racial topography. *Applied Geography*, 122, 102239.
- Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, 67(2), 281–315.
- Nowosad, J., & Stepinski, T. F. (2019). Information theory as a consistent framework for quantification and classification of landscape patterns. *Landscape Ecology*, 34(9), 2091–2101.
- Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological methodology*, 32(1), 33–67.

White, M. J. (1986). Segregation and diversity measures in population distribution. *Population index*, 198– 221.

The real estate ads, a new data source to understand the social representation of urban space.

A.BLANCHI ; G.FUSCO ; K.EMSELLEM

Université Côte d'Azur, CNRS, ESPACE

98 Bd Herriot, 06200, Nice, France

alicia.blanchi@etu.univ-cotedazur.fr, +33638055656

Keywords: Real estate ads, urban space, socio-spatial representation, residential context, textual data.

Real estate ads are a rich source of information for urban geographers. They relate properties inscribed in urban space and are intended for a target population. They reflect the reality of a property and the projection of potential buyers in it, as well as in urban space. They speak indirectly of individuals, social groups, places, urban forms, objects in space as they are represented by its authors and readers (Figure 1).

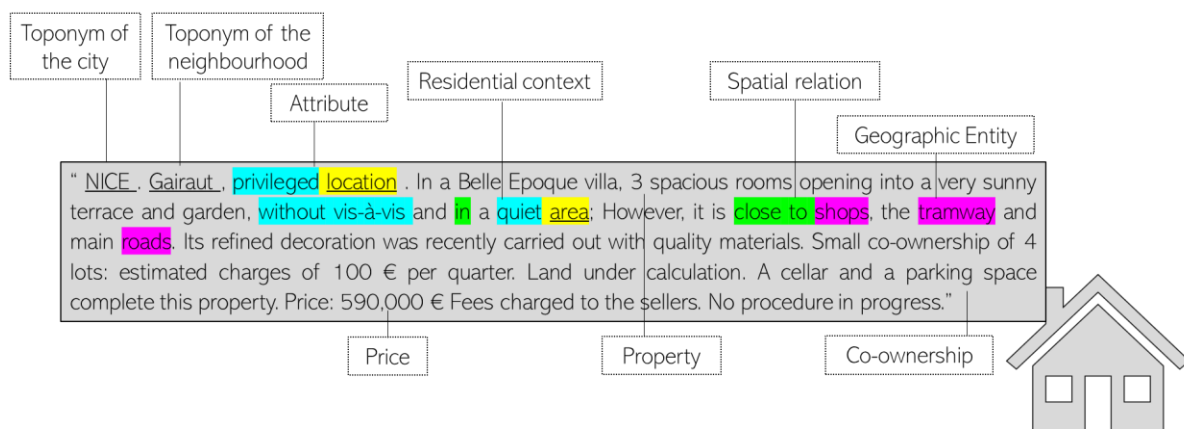


Figure 1: Real estate ads from a web about a property located in Nice (France) written in English

Previous research on real estate ads usually focuses on the analysis of the market value of properties in urban space (Ribardi re and Valette, 2017). Descriptors of the piece of real estate and of its asking price are of central importance in these analyses. However, real estate ads also contain descriptive elements of the spatial context of the proposed property and, indirectly, offer the possibility to investigate the social representation of the different subspaces within a wider urban real estate market.

In this contribution we propose to expose how French real estate ads can constitute as a new corpus to study the social representations of urban space in French metropolitan areas.

We define social representations of urban space as a set of specific meanings in relation to the urban space shared by a social group (Jodelet, 2015).

Usually, the methods used to collect representations of urban space are mainly based on surveys or interviews with individuals or groups of individuals in the form of speeches, quizzes, or mind maps according to the stakes of the research (Alba 2004, Dernas, 2018). However, it is a time-consuming, and tedious method which takes a lot of time to set up and to realize. It is necessary to have a significant number of individuals surveyed even. In recent years, new data sources have been proposed to study the social representations of urban spaces (Hu et al. 2019).

We think that real estate ads can be added among these new data sources, and this for several reasons.

Firstly, real estate ads are punctual data located in the space. Although still in the minority, some ads are geolocated to the property (address or XY coordinates). They allow to conduct studies at a very fine scale.

Secondly, real estate ads always have a text that corresponds to the description of the property and its environment. Different strategies to highlight the property are chosen depending on the target population (to use this or that word, enhancement of this or that characteristic, choice of platform). Real estate ads are the witness of the urban society which produced them because their text testifies at a given historical moment the information relating to the urban space mentioned and differently valorised. They mention geographical entities of several kinds (human and administrative boundaries, natural features, roads and transit facilities, amenities, architectural features of the neighbourhood) and toponyms (Figure 1). The territorialized resources to which housing gives access (Sigaud, 2015) are important in residential choice. The location of the property can be valorised in the text across different spatial scales: Either by the location “Located in...” or by a distance “Nearby...”, whether real or perceived (Figure 1). The geographical elements or atmosphere of the city are meaningful for the author and the reader. We can safely assume that the reported features are already the results of a cultural filtering for the purpose of the ad, mentioning those considered important for the target population and omitting others.

If ads relating to properties in a given subspace consistently highlight the same contextual features, addressing the same target population, they indirectly characterize that subspace as being the place of the young, of the established or wannabe upper-class, of suburbanite families with children, etc. The presence/absence of toponyms is also particularly informative of the capacity of each subspace to become a recognized place in the social representation of city dwellers.

Third, we believe that real estate advertisements are traces of urban realities and representations. Thereby, in our analyses, we make an important assumption: the social representation of a given neighbourhood or residential context relates less to the ads' authors and more to their target population. Real estate agents must reflect the prevailing social representation of urban space of their potential clients if they want to successfully propose their service on the real estate market.

The authors speak of the individuals either directly by mentioning the target population ("Ideal for families", ...) or indirectly by describing the residential context. For example, the specific attractions mentioned, the vocabulary used, the attributes of the residential context, can all echo the residential trajectories of the target populations who are based at least on lifestyle, lifecycle and social status (Shearmur and Charron, 2004, Fusco and Scarella, 2017).

However, analysing real estate ads present many challenges to overcome:

To study real estate ads in geography, you have to pay attention to the location of the ads. Even if the real estate ads are punctual objects in space, their location are not perfect (erroneous location or not located) but also can have different scales of location (Department, city, neighbourhood, residence, address). Therefore, a fine-grained spatial analysis is only possible when the location is identified at the address.

The other challenge is relating to the specific character of the text. The latter can be written in telegraphic style, sometimes not real sentences with only sequences of words, sometimes are neglected with spelling mistakes. It is therefore necessary to manage textual biases to be able to capture important information and their contexts.

Real estate ads also need careful attention to the biases of real estate marketing. They are marketing objects who respond to business strategies introducing cognitive biases. Their speeches are very oriented: Only positive aspects are mentioned, and the other aspects are implied or unspoken. These specificities must be addressed by any analytical procedure of real estate ads.

Moreover, to qualify the social representation of urban space we need to specifically address the words related to the spatial context (Toponyms, geographical entities, spatial relations among them and the piece of real estate). The very fact that features of the spatial context represent a small fraction of real estate ads makes a lexicometric approach relatively difficult. A grammatical analysis of the ad could thus be necessary to understand which attributes are specifically associated to geographical entities. Therefore, the greatest challenge would be to extract the desired information by separating information relating to space and to the property.

Finally, real estate ads are massive data which are regularly updated and available throughout the national territory thanks to the development of online platforms. The downside is the notion of temporality: It's difficult to have the old data because the ad disappears from the platform after the sale. Contrary to surveys, the description of social representation of space in real estate ads is poorer but their massive nature allows a relatively thorough coverage of whole metropolitan areas, as well as the automation and reproducibility of the methodology at different scales.

An analytical protocol tackling these challenges is presently under development and will be tested on the corpus of real estate ads from French metropolitan areas.

Acknowledgement: This research was carried out thanks to a research grant by KCityLabs, KINAXIA Group (CIFRE Agreement with UMR ESPACE).

References

- Alba, M.D., 2004. El Método ALCESTE y su Aplicación al Estudio de las Representaciones Sociales del Espacio Urbano: El Caso de la Ciudad de México. *Papers on Social Representations*. 13, pp. 1.1-1.20. Available at: http://www.psych.lse.ac.uk/Psr/PSR2004/13_01Alb.pdf.
- Dernat, S., Bronner A-C., Depeau S., Dias P., Lardon S., et al., 2018. Représentations sociocognitives de l'espace géographique. *Cartotête. 2e Journée d'études*, Apr 2017, Strasbourg, France. 978 - 2 - 7380 -1422 - 1. hal-01934636
- Fusco G., Scarella F., 2014. Bayesian Clustering of Housing Preferences and Lifestyles: a Comparative Analysis of Home Ownership in South-Eastern France. halshs-01505436
- Hu, Y., Deng, C., Zhou, Z., 2019. A Semantic and Sentiment Analysis on Online Neighborhood Reviews for Understanding the Perceptions of People toward Their Living Environments. *Annals of the American Association of Geographers*. doi:10.1080/24694452.2018.1535886
- Jodelet, D., 2015. Les représentations socio-spatiales de la ville. In D. Jodelet, *Représentations sociales et mondes de vie* (pp. 145-177). Paris, Editions des Archives contemporaines. https://www.researchgate.net/publication/328175517_Les_representations_socio-spatiales_de_la_ville
- Ribardièrre, A., Valette, J.-F., 2017. Géographie des prix immobiliers à Mexico : variabilité et hétérogénéité des valeurs enregistrées dans les annonces en ligne. *Cybergeo*. doi:10.4000/cybergeo.28040
- Shearmur, R., Charron, M., 2004. From Chicago to LA. and back again: A Chicago-inspired quantitative analysis of income distribution in Montreal. *Professional Geographer* 56, 109–126. doi:10.1111/j.0033-0124.2004.05601016.x
- Sigaud, T., 2015. Accompagner les mobilités résidentielles des salariés : L'épreuve de l'entrée en territoire. In « *Espaces et sociétés* » 2015/3 n°162. 129-145. <https://www.cairn.info/revue-espaces-et-societes-2015-3-page-129.htm>

An Empirical Agent Based Modelling approach to Evaluating Intergroup Contact Theory

Huck J.J.¹, Whyatt, J.D.², Dixon, J.³, Davies, G.⁴, Sturgeon, B.⁵, Hocking B., Jarman, N.⁶, Bryan, D.⁷

¹ MCGIS, Department of Geography, The University of Manchester
Oxford Road, Manchester, M13 9PL, UK

jonathan.huck@manchester.ac.uk (correspondent author)

^{2,4} Lancaster Environment Centre, Lancaster University
Library Avenue, Lancaster, LA1 4YQ, UK

d.whyatt@lancaster.ac.uk, gemma.davies@lancaster.ac.uk

³ Department of Psychology, The Open University
Walton Hall, Milton Keynes, MK7 6AA, UK

john.dixon@open.ac.uk

^{5,7} Department of Anthropology, Queens University
University Rd, Belfast, BT7 1NN, UK

b.sturgeon@qub.ac.uk, d.bryan@qub.ac.uk

⁶ Peace Direct

Second Floor, 72-74 Mare St, London E8 4RT, UK

n.jarman@conflictresearch.org.uk

Keywords: Contact Hypothesis, Bayesian, Segregation, Activity Space, Belfast

Introduction

Intergroup contact theory (ICT) states that if contact occurs between members of diverse groups under reasonably favourable conditions then it will reduce intergroup prejudice (Allport, 1954). ICT remains one of the most strongly supported theories in social psychology, and a fundamental component in the study of segregation (Pettigrew, 2021). However, the implications of this effect at the population level have yet to be examined in detail. There is a long history of the study of segregation using ABM approaches, dating back to Schelling (1969; 1971), but these studies generally remain focused only on the production of residential segregation, in spite of an increasing recognition that segregation affects individuals across a range of socio-geographical spaces (Wong and Shaw, 2011; Farber et al., 2012; Huck et al., 2019). This paper briefly describes the use of an empirical ABM in order to evaluate the progression of activity space segregation and implications of ICT using a case study location in North Belfast, which is a community with a long history of segregation and sectarian violence.

Method

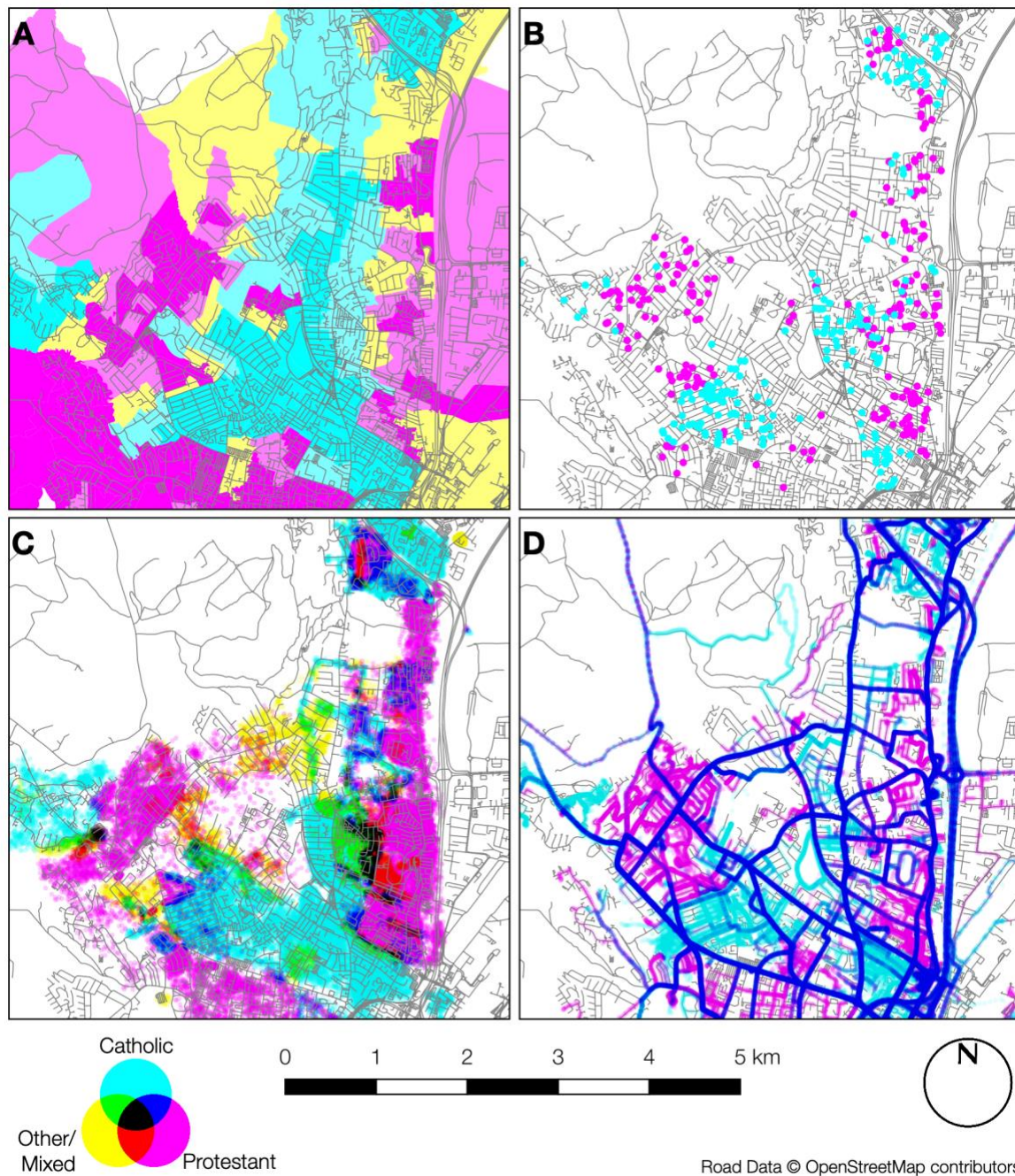


Figure 1: Input Datasets: (A) Census; (B) Survey; (C) PGIS; (D) GPS Traces; reproduced from Huck et al. (2021).

This study relies upon four key datasets:

1. Northern Ireland Small Area (SA) data of the 2011 census was obtained from NISRA¹, including the percentage of Catholic, Protestant and ‘Other’ residents in each area (Figure 1A).

¹ <https://www.nisra.gov.uk/support/geography/northern-ireland-small-areas>

2. A paper-based questionnaire was administered to 488 participants (242 Catholic, 246 Protestant) in a door-to-door survey, exploring various aspects of identity, perceptions and experiences of segregation in Belfast (Dixon et al., 2020a; 2020b) (Figure 1B).
3. A Participatory GIS (PGIS) dataset was collected using the freely available Map-Me platform² (Huck et al., 2014; 2019). 33 participants (14 Catholic, 17 Protestant, 2 ‘Other’) used the system to define places that they consider to be *Catholic*, *Protestant*, and *Mixed* (Figure 1C).
4. GPS traces were collected using a bespoke Android mobile application³ from 197 participants (93 Catholic, 92 Protestant, and 12 ‘Other’) for a period of up to 14 days (Dixon et al., 2020a) and cleaned as per Davies et al. (2017) (Figure 1D).

Modelling Community Belonging

The experience of segregation is incompatible with the official administrative boundaries with which communities are frequently represented by authorities and researchers (Huck et al., 2019). In order to better reflect community understandings of territory, we build upon the ideas posited by Huck et al. (2018) by using Bayesian Inference to produce evidence-based surfaces of *community belonging* to define territorial patterns across the study area. Formally, these surfaces represent the probability that the next individual observed at a given location will from a given group; but here are used as a proxy for the extent to which a location might be considered to be part of Catholic or Protestant group territories. A separate model was used for each cell in a continuous raster surface with 20m resolution. ‘Prior’ belief values for each model taken from the census data, after which evidence is iteratively added from the Survey, PGIS and GPS data using an IDW² approach. Each set of ‘evidence’ was treated as a multinomial distribution of observations, with the parameter vector drawn from a Dirichlet distribution. Finally, we used a Markov Chain Monte Carlo (No U-Turn Sampler) approach to estimate the posterior distribution from 1000 sample draws. We took the mean and credible interval (CI; 95% of the values around the mean) from the resulting distribution, giving both estimated values and uncertainties for each cell in the output surface. This analysis was implemented in Python⁴ using a High-Performance Computing (HPC) facility. The resulting surfaces are illustrated in Figure 2.

² <http://map-me.org>

³ <https://github.com/jonnyhuck/bmp-pathways-app>

⁴ <https://github.com/jonnyhuck/bmp-community-belonging>

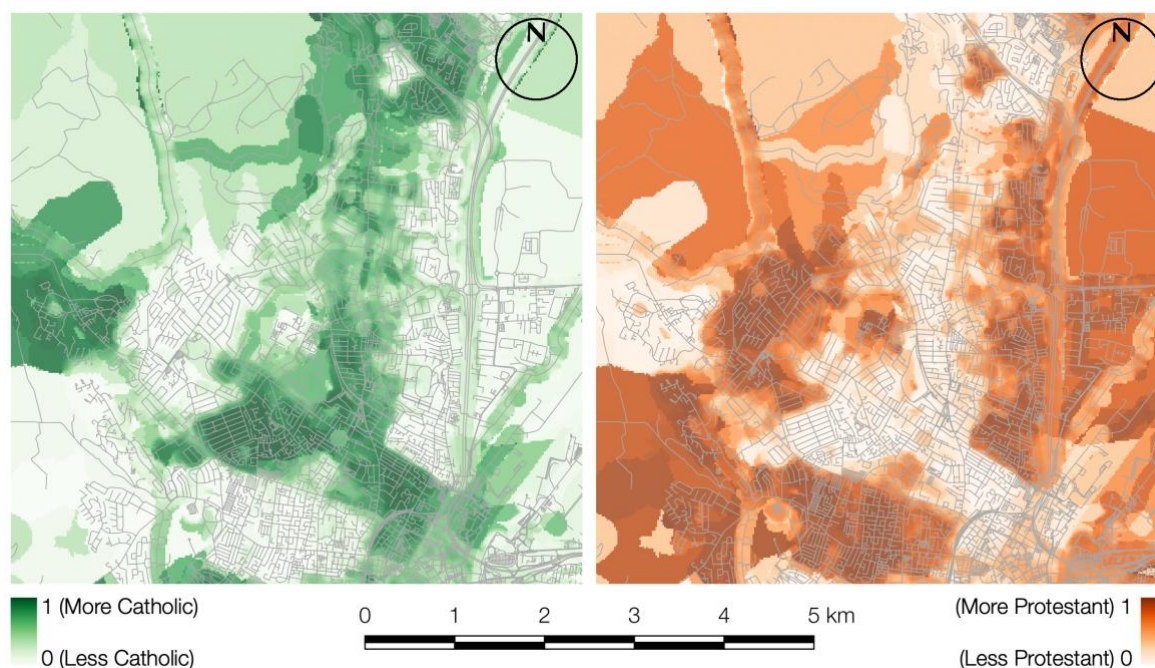


Figure 2: Probabilistic community belonging surfaces describing the degree to which a point in space ‘belongs’ to Catholic / Protestant groups; reproduced from Huck et al. (2021).

Agent Based Model

These *community belonging* surfaces (Figure 2) were used alongside the GPS data to create an empirical ABM of individual behaviour in the study area. Each individual agent’s behaviour is determined by a set of values drawn from probability distributions for a range of variables including group; gender; tolerances residence location, movement and destination location (based on the *community belonging* surface for the ‘other’ group); number, type, timing and mode of daily journeys; car ownership; walking speed etc. Distributions were created through analysis of the GPS, Survey and *community belonging* datasets. Routing for agent journeys is performed using modified implementation of the A* algorithm (Hart et al., 1968), which identifies the shortest route within the agents’ specified tolerances. The model runs in ‘steps’ of 1 minute (in ‘model time’), with segregation evaluated for each agent at each step using the *interaction* index proposed by Palmer (2013) as a generalised index for activity-space segregation (Equation 1).

$$interaction = \sum_{i=1}^n \frac{y_i}{b_i n}$$

Equation 1.

Where: y_i is the number of out-group person-hours that fall within individual i 's personal interaction space; b_i is the number of person-hours of both groups that fall within individual i 's personal interaction space; and n is the total in-group members. In order to maximise efficiency, the model runs on an HPC facility, with agent movements at each step undertaken in parallel, whilst contact encounters and interaction metrics are calculated in series after all movements are complete. Positive and negative contact encounters affect agent tolerances based on effect sizes reported in the literature, and therefore influence their future behaviour. The mean of all agents' *interaction* scores is recorded each step as a measure of 'overall' segregation for the whole area. Prior to the introduction of any interventions, model performance will be validated through comparison of the interaction index profile of the agents against that generated using the 'reference' dataset of GPS traces. This will permit us to confirm that the simulation 'baseline' is reflective of the input datasets prior to the introduction of any interventions and examination of their impact on the temporal profile of segregation values.

Conclusion

This research has the potential to reveal novel insights into the implications of ICT at the population level, thus advancing our understanding of how ICT might be deployed in order to influence patterns of sharing and segregation through policy intervention.

Acknowledgments

This research was conducted as part of the Belfast Mobility Project (<https://belfastmobilityproject.org>), funded by the ESRC (ES/L016583/1). The authors thank the residents of North Belfast who contributed openly and generously to this study.

References

- Allport GW (1954) *The Nature of Prejudice*. Boston: Addison-Wesley.
- Davies G, Huck J, Whyatt D, et al. (2017) Belfast mobility: Extracting route information from GPS tracks.
- Dixon J, Tredoux C, Davies G, et al. (2020a) Parallel lives: Intergroup contact, threat, and the segregation of everyday activity spaces. *Journal of personality and social psychology* 118(3): 457-480.
- Dixon J, Tredoux C, Sturgeon B, et al. (2020b) 'When the walls come tumbling down': The role of intergroup proximity, threat, and contact in shaping attitudes towards the removal of Northern Ireland's peace walls. *British Journal of Social Psychology* 59(4): 922-944.
- Farber S, Páez A and Morency C (2012) Activity spaces and the measurement of clustering and exposure: A case study of linguistic groups in Montreal. *Environment and Planning A* 44(2): 315-332.
- Hart PE, Nilsson NJ and Raphael B (1968) A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics* 4(2): 100-107.

- Huck J, Whyatt J and Coulton P (2014) Spraycan: A PPGIS for capturing imprecise notions of place. *Applied Geography* 55: 229-237.
- Huck J, Whyatt J, Dixon J, et al. (2021) Empirical Agent-Based Modelling for exploring Intergroup Contact in a Segregated Society. *Proceedings of the 29th Geographical Information Science Research UK Conference*.
- Huck JJ, Whyatt JD, Dixon J, et al. (2019) Exploring segregation and sharing in Belfast: A PGIS approach. *Annals of the American Association of Geographers* 109(1): 223-241.
- Huck JJ, Whyatt JD, Hocking B, et al. (2018) A Bayesian Approach to 'Community Belonging' and Segregation in a Divided City. In: *The 26th Annual GIS Research UK Conference*, Leicester, UK.
- Palmer JR (2013) *Activity-space segregation: Understanding social divisions in space and time*. Princeton University.
- Pettigrew TF (2021) Advancing intergroup contact theory: Comments on the issue's articles. *Journal of Social Issues*.
- Schelling TC (1969) Models of segregation. *The American Economic Review* 59(2): 488-493.
- Schelling TC (1971) Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143-186.
- Wong DW and Shaw S-L (2011) Measuring segregation: An activity space approach. *Journal of Geographical Systems* 13(2): 127-145.

Perceptions towards improving the participatory urban planning in Alexandria, Egypt.

Pakinam Hassan

Ph.D. candidate in Human Geography at School/Department School of Geography, Earth and Environmental Sciences,
The University of Birmingham, Birmingham, United Kingdom.

And

Teaching Assistant at the Architectural Engineering Department, Faculty of Engineering, Alexandria University.
61 Sheepmoor Close, B17 8TD, Birmingham, United Kingdom

Email address: pah945@student.bham.ac.uk

Telephone: (UK) 0044 7587264638

Keywords: public participation, state of power, urban development, mixed method research, new strategy.

1. Abstract

Public participation in planning became essential in securing well-functioning and efficient planning processes. In the last ten years, Egypt witnessed rapid development in the field of construction and planning. However, this fast development was without adequate involvement from the residents, which led to a feeling of exclusion, which contracts with the 2011 revolution concepts driven by public voices and participation. Consequently, this study brings public participation into focus. Furthermore, it aims at introducing a new strategy of the urban planning decision-making process. This is based on the participatory approach using a mixed method composed of qualitative (interviews with experts in the planning and architectural field) and quantitative (surveys with Alexandria's residents) methods via analyzing the case study of Alexandria, Egypt.

In this respect, the main research question is: How could people be involved in the planning process's decision-making, and in what stages would their engagement be beneficial? Through an in-depth analysis of the interviews' discussions and the surveys' outcomes, the author highlights the main obstacles to implementing public participation in Egypt: time, weak local governance, authenticity, trust, centralization of decision making, and others. Furthermore, this paper identifies the different

perceptions and ideas from planners, NGOs, and residents and the existing opportunities to implement the participatory approach by engaging people in some initial steps and contributions to small-scale projects. Finally, the author presents a new proposal of the planning strategy in Egypt.

2. Introduction:

Participation in community development is considered a vital human need and essential to communities' psychological health. It is influenced by many factors, such as the community context in which participation happens, and the level of diversity and priorities. Furthermore, it is crucial to understand the link and relationship between the change in the built environment, the participation level and human wellbeing (Butterworth, 2000).

Egypt is one of the developed countries that were a part of the Arab spring. In this event, people practiced public participation in political affairs. Directly after 2011, Egypt went through many changes at the political and developmental levels. Previous attempts to apply this strategy in the urban planning decision-making process through the General Organization of Physical Planning (GOPP) introduce a comprehensive strategy for urban planning problems in 2005. They tried to decentralize the decision-making process in planning incorporation with the UN-Habitat.

Previous trials to implement and study public participation in planning have highlighted difficulties with implementation on the ground and questioned the nature and extent of participation beyond consultation (Mahmoud 2011). Further, Ali (2011) argued that without democracy stakeholder alienation would continue.

Consequently, the debate in this study is framed around the possible way of implementing the public participation concept in urban planning decision-making. Furthermore, the research problem is focused on three keywords: the gap in the literature, state of power, and the current situation. Therefore, this study brings public participation into focus. It aims at introducing a new strategy of the urban planning decision-making process.

3. Materials and Methods

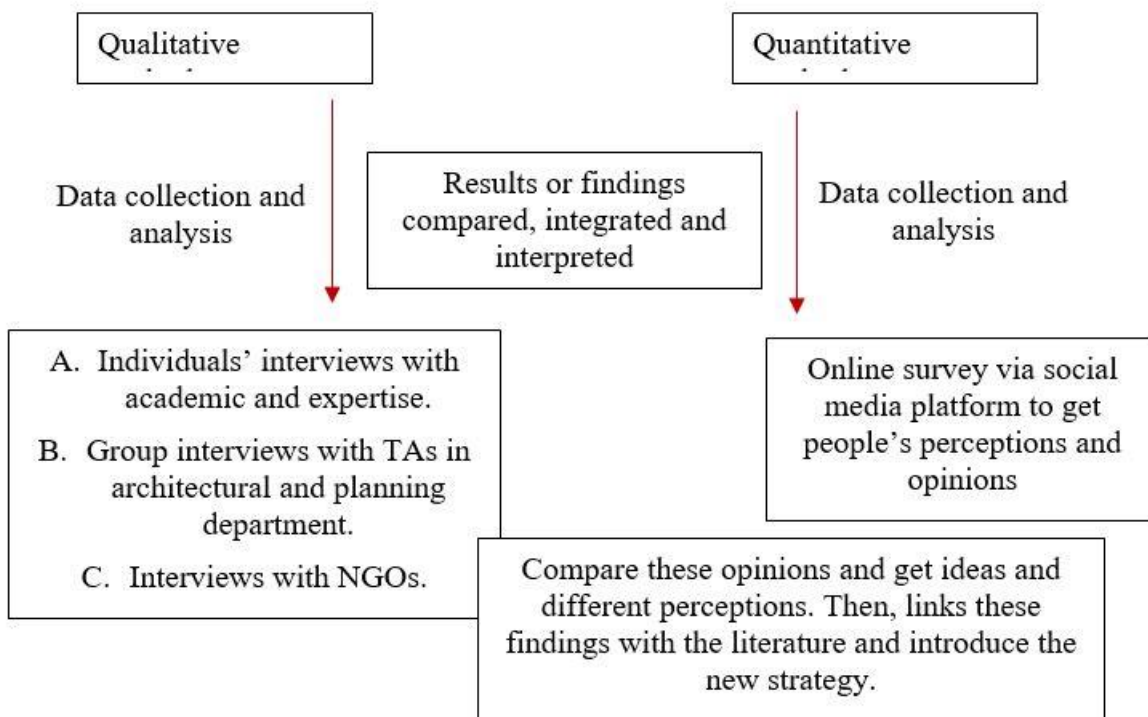
The public participation in the planning decision-making process in Alexandria, Egypt, was presented as a case study in this research with a mixed-method approach. The reason behind choosing Alexandria than it witnessed a massive change without the involvement of the public.

This study is inductive; it is mainly focusing on observing the context of public participation in Alexandria. Then, collecting and analyzing people's different perceptions and perceptions, and finally, the research aims to develop recommendations and improve the existing planning decision-making process or strategy. The empirical analysis in this study concentrated on data collection of secondary resources (maps, previous reports of the development projects, and previous studies) and quantitative and qualitative methods.

On the one hand, the quantitative method (Duminy & Organizing) helped the researcher represent the current situation with some percentages and numbers and understand the real world (objectivity). On the other hand, the qualitative method (interviews) introduced people's perceptions, experiences, and the social and cultural system in which they live (subjectivity) (Nicholls, 2009). It is necessary to know the different perceptions and ideas from planners, NGOs, and residents. The used mixed method is convergent parallel mixed-method research (Figure 1).

The qualitative method was 36 one-to-one online interviews with some academic staff in architectural planning departments in Alexandria, planners, architects, and NGOs who focus their efforts on the development problems in the city and the protection of its heritage. The quantitative method was in the form of an online survey distributed via Facebook, social media platform. The researcher collected 490 completed surveys at the end of the data collection stage. Then, it was followed by the analysis of the interviews and surveys. After that, there was a discussion and comparison between the interviews and the survey findings. Finally, these findings and comparisons led the research to the recommendations and the new strategy.

Figure 1 diagram presents the convergent parallel mixed method used in the study (by the author)



4.Results and discussion

The findings of this study could be summarized on some essential points that influence the implementation of the participatory approach in the Egyptian planning system. Through the deep discussion with the interviewees and the online survey findings, there were many points that both residents and experts agreed on the same opinion. However, not all the experts had the same opinions on some of the discussed points, such as the desire to participate. On the one hand, some experts see that we should give people a chance to participate because their opinions are valuable and know more about their needs. On the other hand, other experts see that people in Egypt do not have time to participate, and they are busy with their economic problems. While in the results of this question in the survey, the majority agreed that they want to be more involved (table 1).

Table 1 the residents' responses to the question : Would you like to be more involved?

Much more involved	34.69%
--------------------	--------

More involved	53.88%
Neither more nor less involved	5.31%
Less involved	2.04%
Much less involved	0.41%
Do not know	3.67%

Furthermore, one of the discussed points was the issue of trust. A common comment was that we have a two-way trust issue. For example, an architect and a member in an NGO said:

“ We have a two-way trust issue. The government does not trust that the public opinions are valuable, and the public does not believe that the government (Butterworth, 2000) will hear their voices, and also they do not trust that all this development will fulfill their needs or that these projects are beneficial.”

Also, the interviewees did not just highlight the problem but give several solutions for each problem. So, in the trust issue, an architect suggests how can the government gains the trust of the society and said:

"For me, I will begin to trust the officials if I see small steps or initiatives that fulfil my needs.... For example, if they fix infrastructure issues that affect my neighborhood or doing a green public space for free. I want to feel that the citizens' wellbeing, needs, and comfort are their priorities.“

Additionally, many other points were discussed, such as time, centralization (Top-down approach), lack of information about projects and the ways and channels of objections, the weakness of local governance, lack of compensation for owners of heritage buildings, the use of media (conferences, advertisement in TV and social media), Corruption' money talks!' Activating laws and violations and essential training for the employees in the local administrations.

5.CONCLUSION

To conclude, this research focuses on the significance of updating the planning decision-making process in Egypt based on the participatory approach by investigating the case of Alexandria city. The

research findings highlight the significance of treating people like adults; give them respect, recognition, and responsibility. Furthermore, it clarifies the main opportunities, strengths, weaknesses, and threats in Egypt's current public participation process and gives some solutions to implement a more successful process.

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. Patricia Noxolo and Dr. Natasha Cornea and my funders, the Egyptian Cultural Affairs and Missions Sector.

References

Aly, S., Amer, M. J. W. T. o. E., & Environment, t. (2011). Public involvement in sustainable development: a public participation process in the Sidi Gaber railway station development project, Alexandria, Egypt. 150, 537-552.

Butterworth, I. J. M. V. H. P. F. (2000). The relationship between the built environment and wellbeing: a literature review.

Duminy, J. J. B. I. s. c. f. t. U. P. P. o. t. p.-r. n. W. i. i. E. G., & Organizing. (2011). Literature survey: Informality and planning.

MAHMOUD, H., & ARIMA, T. J. (2011). Evaluating The Participation Policy in Strategic Planning of Egyptian Villages. 76(662), 809-818.

Nicholls, D. (2009). Qualitative Research: Part One – Philosophy. *International Journal of Therapy and Rehabilitation*, 16, 526-534. doi:10.12968/ijtr.2009.16.10.44562

Parallel Session PSA2

3 November 2:30-5:00 pm (GMT)

Urban Form

Tackling urban air pollution concerns through urban form – Lessons from a GIS-based residential choice model with micro-economic foundation

Mirjam SCHINDLER¹

¹School of Geography, Environment and Earth Sciences, Victoria University of Wellington

Kelburn Parade, Wellington 6021, New Zealand

+64 4 463 5645, mirjam.schindler@vuw.ac.nz

Keywords: residential location choice model, urban simulation modelling, GIS, urban form, air pollution

Air pollution causes environmental and health effects in urban areas and residents are increasingly concerned particularly about being exposed to harmful levels of pollution concentrations. Worldwide, over 4 million pre-mature death are attributed to air pollution annually and urban planners and policy makers are debating potential remedies to reduce environmental and especially health impacts due to air pollution. Traffic is a main source of pollution under debate in urban planning, also because it is the cause of high variability in pollution concentrations within urban areas. This study adds to recent research which suggests that urban form, and most notably local urban characteristics, can play a pivotal role in mitigating population exposure.

This paper presents a GIS-based residential location choice model grounded in urban economic theory to study impacts of i) road network design and ii) cadastre design, and iii) the level of residents' health concern on intra-urban patterns of population density, pollution concentration and population exposure to those concentrations. A systematic scenario comparison informs on how characteristics of urban form and awareness among the population can steer urban and health outcomes. It helps planners to better understand these complex urban processes to inform the design of future interventions.

By proposing a spatially explicit residential location choice model implemented on a hybrid empirical-synthetic space in GIS, it advances methodological approaches in urban simulation modelling. Previous studies have explored these feedback effects between either residential and transport choices and population exposure using spatially explicit models or both. Yet, the models were either based on 1D or 2D (regular) cell-based representations of space. Such representation allows for the provision of interesting insights into underlying processes but offers limited flexibility to explore more advanced, spatially explicit urban and transport design components. Cell-based urban models such as cellular

automata or agent-based models in this context offer valuable but limited possibilities to represent the relationships between spatial objects such as between a house and the road network. Resultant urban systems are still far away from real cities in that the shape and complexity of emerging urban structures are simplistic representations of real cities. However, specifically for urban phenomena where the relationship between objects in space is at the heart of the debate, a spatial representation beyond cell-based structures is desirable.

The urban modelling community has explored alternative pathways where the geometry of objects can be considered explicitly, such as irregular cellular automata, vector-based, graph theory-based or GIS-based modelling in different application contexts. This project looks into these spatially advanced modelling methods, in particular vector-based modelling, to investigate how spatially explicit urban and transport design policies can mitigate population exposure to traffic-induced air pollution, taking into account residents' residential preferences. Thus, the methodological contribution of this study is a residential location choice model in the context of localized air pollution in a GIS-based modelling setting and with solid microeconomic foundations. As such, it bridges urban economics, quantitative geography and urban simulation models.

Relevant previous work in this context has focussed on theoretical modelling space. We turn to a simulation approach which uses initial empirical urban transport settings to generate synthetic urban environments. These resulting urban areas are thus not spatially constrained by cell environments but are spatially more flexible. This is relevant for local urban and transport policy makers and aligns the modelling work more closely with real-world contexts. This modelling setup, combining empirical and synthetic space, allows for a systematic testing of the impacts of urban and transport design. This provides a realistic starting point for exploration of alternative urban developments based on residential preferences and/or urban and transport planning interventions. The cadastre design (i.e. the container space) is initially exogenous in the model. This study therefore compromises to some degree on the endogeneity of space in relation to previous work in this context, but instead advances the complexity of the spatial representation. This also allows a comparison of modelled urban forms as to relevant network and landscape metrics and their performance with respect to pollution concentration and population exposure.

A Novel Regionalisation Algorithm for Redrawing Cities' Functional Boundary within Mega-city Region

Bowen Zhang¹; Chen Zhong²; Qili Gao^{3,4}

¹Department of Geography, King's College London
London, WC2R 2LS, UK

+44 (0) 207 848 8720, bowen.zhang@kcl.ac.uk (correspondent author)

²Centre for Advanced Spatial Analysis, University College London
London WC1E 6BT, UK

+44 (0)20 3987 2708, c.zhong@ucl.ac.uk

^{3,4}Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University
Hong Kong, 999077, Hong Kong SAR

Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University
Shenzhen, 518060, China

+852 95763906, qi-li.gao@polyu.edu.hk

Keywords: urban movements, cell phone data, spatial interaction model, community detection, mega city-region

1. Introduction

The majority of commonly used transport and urban models assumes that flows of people follow a spatial distribution with a distance decay from urban centres to the periphery (Anderson, 2011, Batty and Milton, 2021). Moreover, spatial heterogeneity exists across cities in this distribution due to zoning system, administrative boundaries, transport linkage, jobs and housing balance, and other complicated urban contexts. Because of the limitation of the data and computation, most of the previous studies on flow and urban structures investigated one city only (Zhong et al., 2014). Since the conception of the mega-city regions and the related research topics is getting more important in recent years, some research looked into megacity regions, adding insights into the relationships between cities predominantly using network analysis and emerging mobility data (Zhang et al., 2020). However, the traditional network analysis and most community detection algorithms usually only consider the absolute value of flow volume for dividing the partitions regardless of the spatial factors like travel

distance/cost. When detecting the functional boundary between cities, the current methods focus on the strength of local linkages between particular zones but overlook the reasonability for distributing trips in a global sight. Considering the inter-city flow is usually tiny (10% or less) in overall flows, which may fail to support planners and policymakers appropriately. This paper proposed a novel regionalisation algorithm that re-assigned zones' belonging of cities by searching for the best partition with the best goodness of fitting in the modular spatial interaction model (MSIM). This algorithm is not only more sensitive for cross-boundary trips, and this algorithm can reflect the dynamic change of cities' functional boundaries by fitting cross-boundary trips into global trip distribution, providing support for governments and planners to understand the spatial structure in mega city-region.

2. Data and methods

This research takes the cell phone data in a typical metropolitan region (Shenzhen-Dongguan-Huizhou) in China provided by China Unicom. The data contains Origin sub-district ID, Destination sub-district ID, the volume of travel flow and travel time. It detected 8921 flows with 13,588,846 commuters in the SDH area, including both intra-city and inter-city, figure 1 below shows the Distribution of inter-city flows within the SDH area.

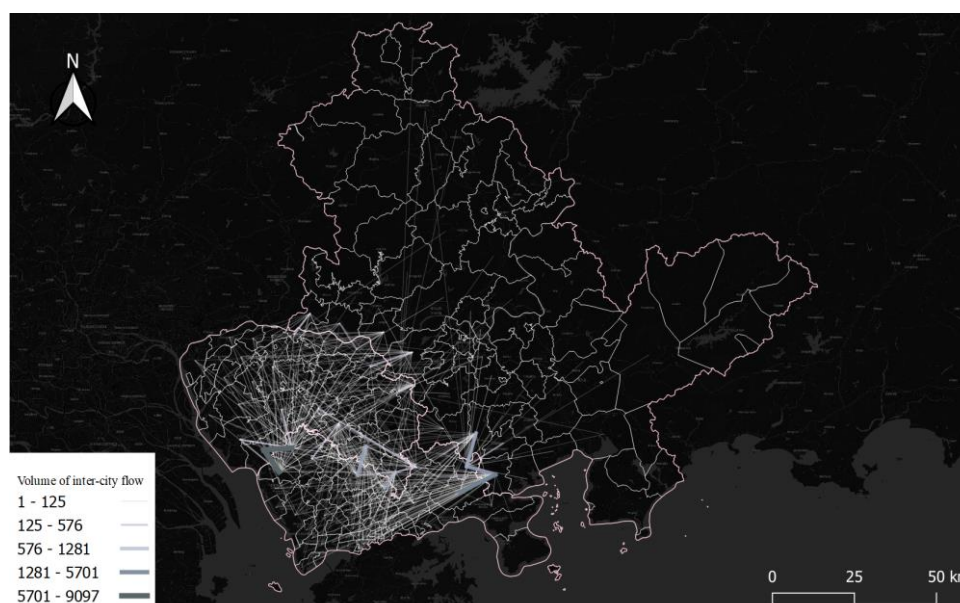


Figure 1 Distribution of inter-city flows within the SDH area

A set of spatial interaction models has been established based on 172 administrative sub-district zones based on the framework of the QUANT model (Batty and Milton, 2021), which is a single constrained gravity model assuming the distribution of trips roughly follows the format of the negative-power function. We found that constrained the administrative boundaries, modifying the spatial interaction model (SIM) by dividing one regional-size 'global model' into several city-level sub-SIM models can significantly improve the goodness of fitting of travel flows by summing of the sub-SIMs. This method is called modular spatial interaction model (MISM). This finding suggests that the travel behaviours of people who belong to the same functional city may yield better performance in fitting the specific distribution of trips. In other words, the goodness of fitting by MISM would be an indicator for assessing the reasonability of functional boundaries of cities. Based on this idea, in contrast to existing administrative city divisions, we draw the boundaries of functional cities by fitting the specific spatial distributions with real human movement flows. Specifically, this algorithm attempts to re-assign the 10% of zones with the highest ratio of inter-city flow and predict the flows by MISM, selecting one re-assigned zone by comparing the fitting results of flows. Then this algorithm adopted the framework of the greedy algorithm, updating boundary with selected zones and repeating the iterating processing above until finding the global best partition with the best goodness of fitting.

We set two scenarios for the case study area: the first scenario is the inter-city flows for each zone is static according to the current administrative boundary during the iteration. This scenario is set based on the situation that current cities' core functional zones can only spill over to zones close to the administrative boundary due to the local authority's current land-use planning and management scope. Moreover, the second scenario is the inter-city flows for each zone is dynamically updating according to the current boundary in iteration processing. It means the changed zones will further influence other unchanged zones as original inner-city flow could become inter-city flow. In this scenario, the cities' core functional zones can spill over freely without the restriction by the current administrative boundary, forecasting the potential functional boundary in the long term.

3. Results and discussion

Figure 2 shows the result of scenario 1 (statistic inter-city flow), indicating the current functional boundary within the SDH area. The functional core of Dongguan city is in the west of its administrative boundary because of its good transport connection with Shenzhen and Guangzhou. Therefore, the only zone in Dongguan that should be re-assigned to Shenzhen is Fenggang, as it has been known as the 'sleep city' for workers in Shenzhen. Moreover, few zones near the Dongguan-Huizhou boundary will be re-allocated to Huizhou from Dongguan because these zones are away from the city centre.

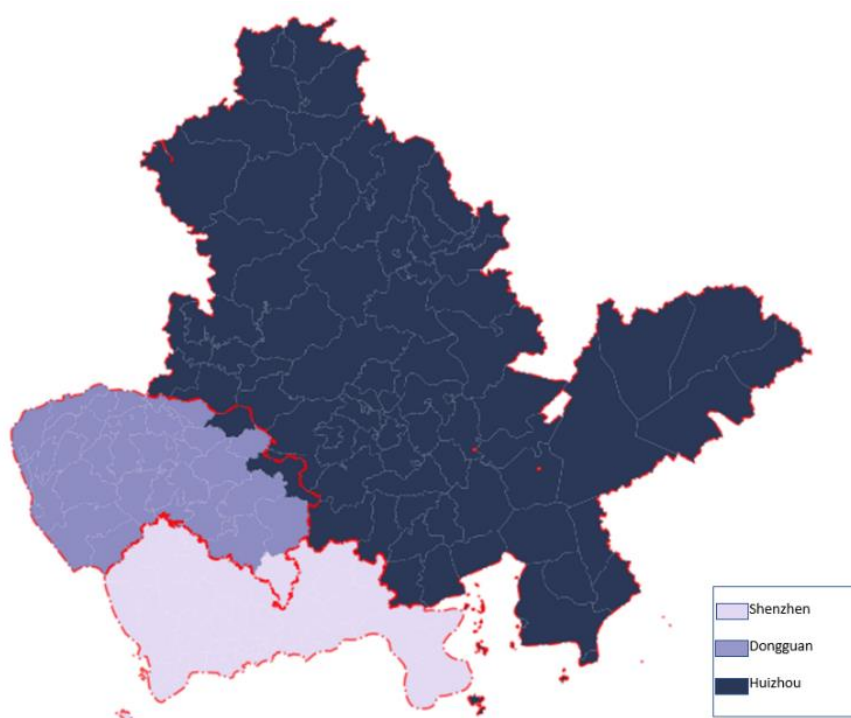


Figure 2 Current functional boundary within SDH area

As for the dynamic inter-city flow scenario, the result (shown in figure 3 below) suggests that the two cities' functional areas will 'invade' the original administrative boundaries starting from the administrative boundaries of Huizhou and Dongguan, respectively.

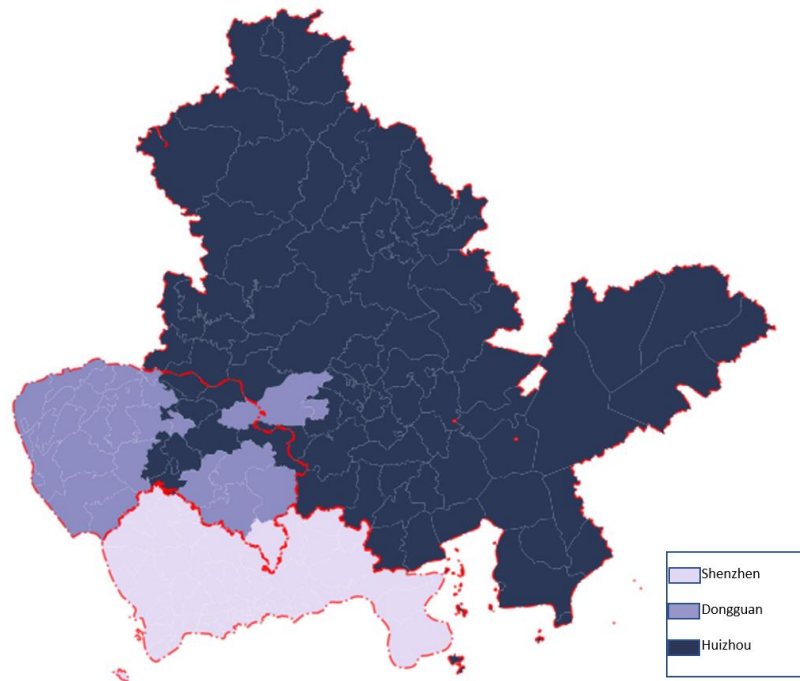


Figure 3 Predicted functional boundary within SDH area in long-term

The re-assigned zones in Dongguan are mainly from the ‘East industrial park’. Historically, these zones are the cluster of manufactory industries but lack commuting connection with the city centre. Therefore, when few zones near Dongguan-Huizhou boundaries are re-assigned to Huizhou, these zones will change their belonging following the zones near the boundary. As a result, more functional areas are potentially staggered between Dongguan and Huizhou, which also means there will be more interaction between Dongguan and Huizhou city in the future.

4. Conclusion

Figure 2 shows the result Previous research on spatial interaction models about zoning systems and the modifiable areal unit problem mainly focuses on how spatial resolution affects the SIM. This research answers how different partitions of zones in the same spatial scale could significantly influence the SIM and its meaning in urban studies for redrawing the functional boundary. These

empirical-based results can help the governments and planners understand the spatial structure in mega city-region and support their urban integration policy.

Acknowledgments

This project has received funding from the ESRC under JPI Urban Europe/NSFC (No. ES/T000287/1), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 949670), and the China Scholarship Council (CSC) from the Ministry of Education of P.R. China.

References

- ANDERSON, J. E. 2011. The Gravity Model. *Annual Review of Economics*, 3, 133-160.
- BATTY, M. & MILTON, R. 2021. A new framework for very large-scale urban modelling. *Urban Studies*, 0042098020982252.
- ZHANG, W., FANG, C., ZHOU, L. & ZHU, J. 2020. Measuring megaregional structure in the Pearl River Delta by mobile phone signaling data: A complex network approach. *Cities*, 104, 102809.
- ZHONG, C., ARISONA, S. M., HUANG, X., BATTY, M. & SCHMITT, G. 2014. Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, 28, 2178-2199.

Stability of urban forms: modelling the emergence of collective behaviour in residential trajectories

Arthur BENICHOU¹; Olivier BONIN²

¹LVMT, Université Gustave Eiffel, ENPC

Champs-sur-Marne, 77420, France

arthur.benichou@eleves.enpc.fr

²LVMT, Université Gustave Eiffel, ENPC

Champs-sur-Marne, 77420, France

+33.1.81.66.88.68, olivier.bonin@univ-eiffel.fr (correspondent author)

Keywords: urban morphogenesis, catastrophe theory, diffusion, residential trajectories, complex systems

1. Introduction and position of the problem

It is difficult to define what a city is, as it is an object that escapes definition by criteria of population, spatial extent, political regime, functions or activities. However, cities are prominent, present in all continents and in almost all cultures, and they are remarkably stable. This stability also partly characterises the internal organisation of cities. Most large cities still have very old traces of their historical spatial organisation, often in the form of districts with fuzzy boundaries but with a clear identity. One can think of the 'Latin Quarter' in Paris, a student district around the historic heart of the Sorbonne, whose identity persists despite the increasing relocation of universities and Grandes Ecoles outside Paris.

Understanding the evolution of cities, but also the obstacles to the transformations driven by public policies, is particularly important in the context where the geographical forms of cities, which are too sprawled out, consume too much energy and are not resilient enough, are no longer adapted to the challenges of climate change. However, the task is immense: how can one analyse the evolution of cities and the mechanisms involved when these objects are not even clearly defined?

One possible approach to overcome these difficulties is to adopt a holistic rather than functionalist approach, i.e. to approach a city as a whole, rather than breaking down its elements: buildings, networks, infrastructures, populations, services, flows, etc. One can then focus on the question of stability of form, emergence and transitions, within the classical paradigm of complexity.

2. Modelling cities as complex systems

Most approaches to the city are either purely individual, in the tradition of methodological individualism (e.g. urban economy models), or integrate a collective dimension which results directly from individual choices (e.g. multi-agent models). These approaches are not fully in line with the complexity paradigm, which posits that groups are not reducible to the sum of the individuals that make them up, and studies the issues of emergence and bifurcation of systems. This problem is particularly clear for residential choices: how can one explain that cities and neighbourhoods are so stable, insofar as individuals behave differently and have no explicit collective strategy), but also insofar as choices are constrained by public power, lack of information and, finally, by the history of successive settlements in the various places?

In his early work, Gilles Ritchot (1992) proposed a simple yet powerful “theory of urban form”. He postulated that different residential trajectories over time shape an abstract urban form, and that this urban form constrains new trajectories. He identified two axiological dimensions to what he calls the “political control of mobility”, both of which have, to varying degrees, an individual and a collective dimension. A household may or may not have control over his mobility (endo- or exo-regulation), and his residential trajectory may or may not be concomitant with other trajectories (focus or diffusion).

Gilles Ritchot's work, although theoretical in nature, has not been mathematically formalised (except in a semiotic dimension which is not the problem addressed here). However, it is possible to propose a model of catastrophe theory which, by crossing the two previous axiological dimensions, makes it possible to explain, at first sight, the emergence of more or less stable urban forms (Bonin, 2021). The main conceptual weakness of this catastrophe theory model is that it is based on the dimensions of focus/diffusion and endo/exo-regulation, without explaining the individual and collective part in these behaviours.

The model proposed in this article therefore aims to study the stability of urban forms by explicitly modelling the emergence of collective behaviours from the residential trajectories of households.

3. Methodology

3.1 Proposal of an axiological model of household residential choice

The present model incorporates two slightly different dimensions from Ritchot's. The endo-/exo-regulation axis of residential trajectories becomes an axis of size of the universe of choice for residential settlement. The axis of focus/diffusion becomes an axis of motivation (conviction, adhesion) of residential settlement.

The first axis remains relatively similar as the main instruments of endo-/exo-regulation are regulation and legislation. The second axis is different, since the question of focus (or diffusion) results rather from the collective: it is proposed here to focus on potential mechanisms leading to the focus or to the diffusion of residential trajectories, rather than to a system of values. Then, the position of a mean point, representing a population, will be interpreted as endo-/exo-regulation and focus/diffusion, which results from the individual axiological axes.

3.2 Catastrophe machine: emergence of collective behaviour

In this model, each household is identified in the complex plane by its coordinates in these two new axiological dimensions. In order to describe the transition from the individual positioning in this axiological system to the collective behaviour inducing transitions between different states of the "city" system, we take inspiration from the catastrophe machine introduced in the Thom-Pomian model (Petitot, 1978). This model describes transition cycles between different political regimes which have been observed in many European countries: monarchy-revolution-dictatorship-restoration. A regime in place obviously depends on the individuals governed by that regime and their opinions; these individuals are similarly located in the complex plane by means of axiological axes defined by Thom and Pomian. The position of the barycentre of individuals (representing society as a whole) results in more or less stable forms of power and can induce transitions.

The barycentre mechanism is relevant in a system where all individuals have the same weight (otherwise it can be generalised by a centre of gravity), and where there is no diffusion of

opinions. Residential choices induce local settlement systems which will then, from one person to another, spread spatially or in an abstract manner (for example, the spread of the suburban lifestyle). Weights are therefore replaced by temperatures, which makes it possible to calculate diffusions in axiological space. The system also keeps track of its previous states. Hotspots can represent social groups (precursors or pioneers, etc.), local planning policies, infrastructures, or even social innovations (bank accounts, housing loans).

3.3 Implementation through simulation

The model adopts a resolution of the heat equation, which is the simplest model for diffusions in the axiological space.

The emergence of a hot spot on the right of the x-axis corresponds to many residential choices, which arises for instance with the democratisation of the private car or with the construction of new buildings. On the left of the x-axis, it may correspond to the creation of large social housing complexes. The emergence of a hot spot on top of the y-axis might correspond to the effect of a change in lifestyles (for example, development of shared accommodation). In negative values, this would be a repellent effect; it is therefore unlikely that a household would be present there.

4. First Results

Figure 1 represents the set of positions of households in the axiological system for two consecutive time stamps (left and right). The red dot represents the barycentre of households and the green dot represents the most stable urban form that is selected by the system.

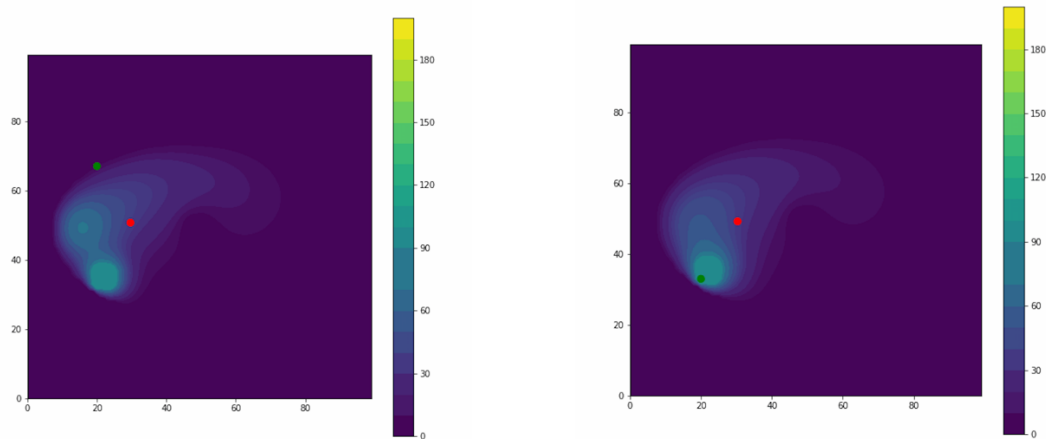


Figure 1: Representation of the heat map summarizing the position of households in the axiological system: size of the universe of choice on the x-axis, and willingness to settle on the y-axis; left and right showing two consecutive times of an evolving system.

Here on Figure 1 (left) households have very little options for settlements (position on the left on the x-axis), and are less and less inclined to favour their residential positions (thus have decreasing values on the y-axis). The barycentre shifts continuously towards the negative values on the y-axis and reaches the point where a stable state is selected: on the left (initial state of the system), the green point is on the upper left corner, indicated a situation of exo-regulation and focalisation, corresponding to an area of concentration such as working-class districts or dense suburbs, and on the right (next state of the system) the green point is on the lower left corner, corresponding to areas of dispersion such as the countryside or to the outer suburbs: the transition can be interpreted thus as the one from the paternalist industrial city to the sprawled city, where the working class households can occupy small houses far from the city centres.

References

- Bonin, O., 2021. Que peut apporter la théorie des catastrophes à l'étude des villes ? Deux exemples et une question. To be published in the proceedings of the ARTHOM seminar.
- Ritchot, G., 1992. La valorisation économique de l'espace géographique. Cahiers de géographie du Québec 36, 175–214.
- Petitot, J., 1978. Sur le modèle historique de Thom-Pomian. Mathématiques et sciences humaines, tome 64, p. 43-70.

Morphogenesis of street networks. A reaction-diffusion system for self-organized cities

Michele TIRICO¹, Stefan BALEV¹, Antoine DUTOT¹, and Damien OLIVIER¹

¹*Normandy Univ, UNIHAVRE, LITIS, 76600 Le Havre, France, [name].[surname]@univ-lehavre.fr*

Keywords Reaction-diffusion system, street network model, urban morphogenesis, complex systems, spatial networks

1 Introduction

Morphogenesis is the natural or artificial process that determine physical forms, shape and patterns of a self-organized system and specialize its subpart (Bourgine et al. 2011; D’arcy 1917). Within urban systems, morphogenesis concern the formation of its physical elements (built-up areas, street networks, public spaces) and the specialization of its suburbs (residential, productive, leisure). Street networks represent a major organizational component of the urban systems (Marshall 2005). Streets are the backbone of the city, the structural support of human activities, the physical witness of the evolution of the urban area. Understanding street network evolution is revealing important information about the growth of cities and their functioning. Modelling the urban system, simulate realistic scenarios and quantitatively measure properties of results is one of well adapted approaches to study the street network morphogenesis (Pumain and Reuillon 2017).

Complexity theory (Morin 2005) provides a comprehensive framework to study urban morphogenesis and more generally urban dynamics (Portugali 2006). A general agreement can be find into the principle that urban systems are characterized by decentralized interactions between their constitutive elements (Batty 2013; Pumain, Sanders, et al. 1989). This context is a source of self-organization and it surrounds the emergence of unanticipated and new properties. Spatial patterns can be observed, as like an aggregation of elements, a regularity of properties or a specialization of a part. The system is surrounded by an environment, which exchange with it and play the role of a morphogenetic actor. Forms appear as an

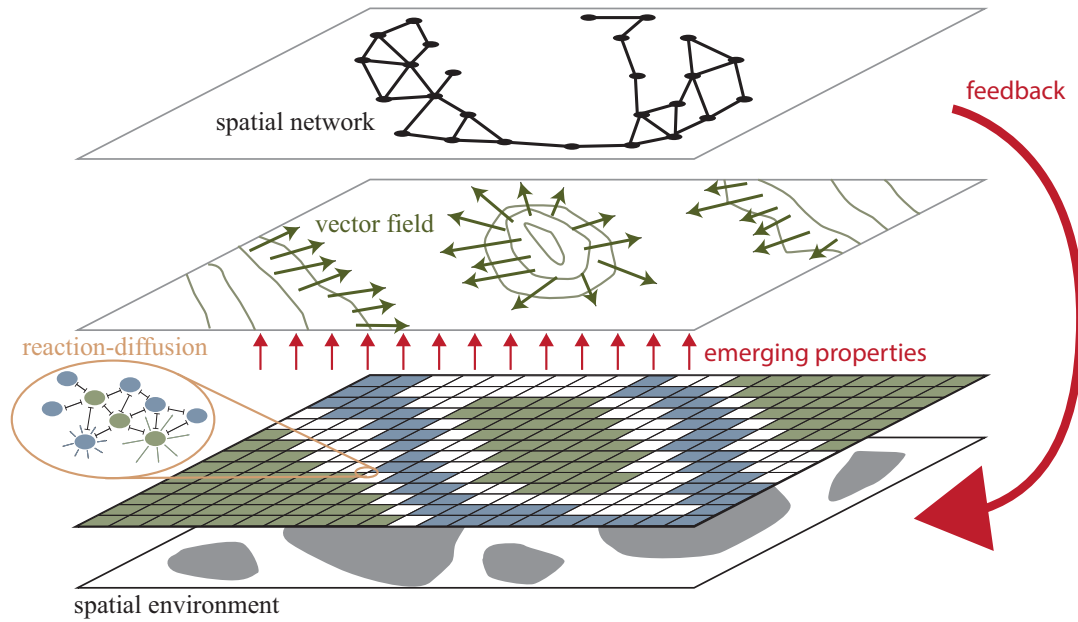


Figure 1: The concept. Morphogens govern the formation of a network which in turn retro-acts to their behaviour.

overlapping of different processes, becoming the result of opposite actions which favour or inhibit the growth. To understand those phenomena we cannot sum it up with an overall design intent and we must consider them in their wholeness.

In the following, we present a street network model generator, we proposed an application in an real case (Fécamp, Normandy, France), we measured different scenarios and and we discuss them.

2 Modelling spatial network morphogenesis

Among different element that composed an urban system, we can suppose that some of them have a dominant role into the morphogenesis of the street network. This latter is constrained by the spatial arrangement of such a form-producer elements and by exogenous factors (like the geomorphology or political decisions). Turing (1952) in his influential paper on chemical morphogenesis call them morphogens. The emergence of the network feeds back to the morphogens, affecting their capacity to organize themselves. We formalize those dynamics with a model composed by three layers surrounded by an environment (fig. 1). The first layer is cellular automata (Wolfram 2002) constituted which simulates the spontaneous emergence of morphogens. The concentration of morphogens updates in accordance with the Gray et al. (1983) model, a particular model of the RD theory. The dynamic is expressed as a RD system (Turing 1952), where two kinds of morphogens (A and B) react and diffuse. At a microscopic scale, A catalyses its

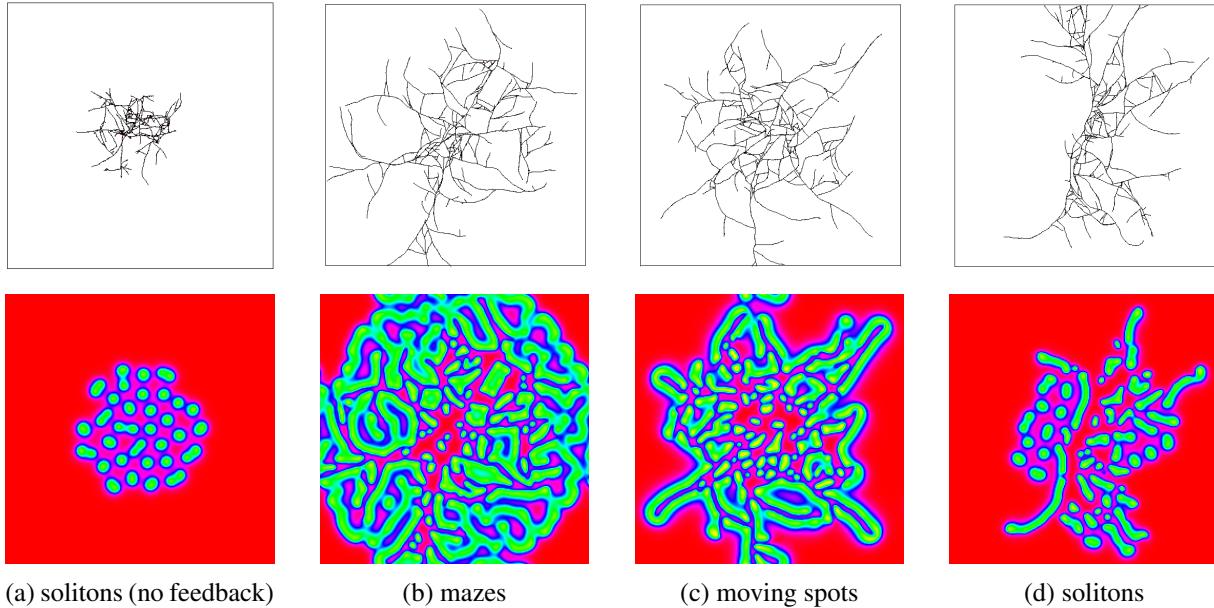


Figure 2: Four experiments (columns). The first row indicates four final spatial networks and the second row the corresponding reaction-diffusion (RD) layer.

own production and also the production of B . At the same time, B inhibits the production of A . B diffuses faster than A . At a macroscopic scale, the RD system can yield evolving patterns of concentration (Pearson 1993). The intermediate layer is a dynamic vector field that reflects the constraints generated by the RD and by the environment. Its role is to model the impact of those emerging patterns to the spatial network (the latest layer). The network formation feeds back to its morphogenetic elements. The mechanism consists to reduce the possibility of morphogens to cluster modifying parameters of the RD layer. The model is completed with a spatial environment, which aim to represent exogenous urban constraints for the street network (as like the geomorphology of the study case, socio-economical forces or policy decisions). More details of the model and a complete overview about experiments can be founded in (Tirico 2020; Tirico et al. 2019).

3 Results

In fig. 2 we present 4 simulations obtained with three pattern formation parameters. The first simulation does not consider the feedback process. Morphogens seem to bypass the graph. Like a physical limit, a barrier defines a changing of behaviour of morphogens: on one side no organizations can form, and on the other side, morphogens create patterns according to the unperturbed pattern formation process. This behaviour is unexpected, it diverges to classical simulation, making a wealthy of new and mixed patterns.

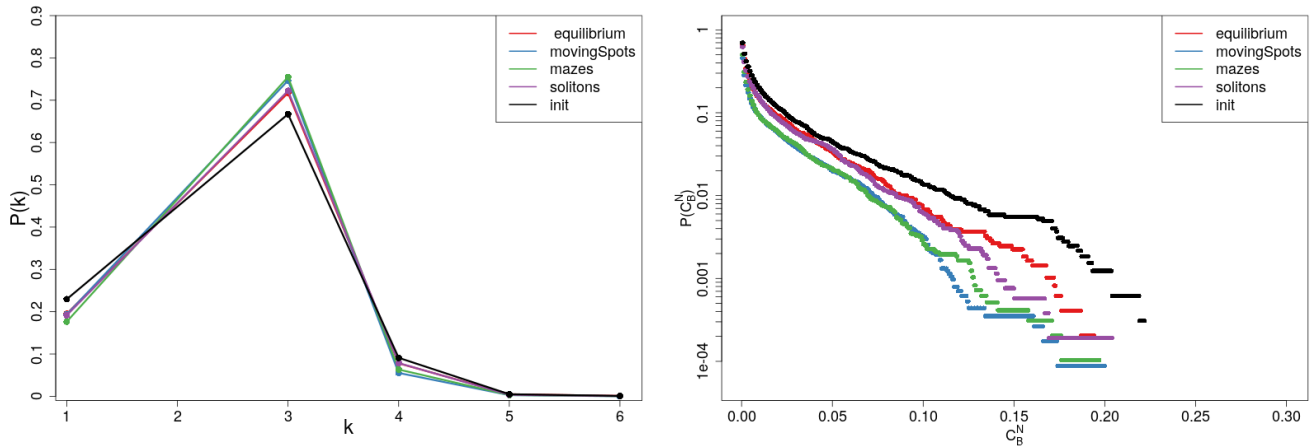


Figure 3: Vertex degree (left) and betweenness centrality (right) statistical distributions of Fécamp. Each panel compare the initial configuration and the configuration at the end of four simulation.

The second experiment concerns the application of the model in a real context: we study the street network morphogenesis of Fécamp town. The environment integrate several morphogenetic aspects that had been considered during the simulation: the geomorphology (a second vector field was computed from an orographic map and combined with the main one), the build-up and the green area. In green and built-up areas we reduced the probability of the network to develop. In this way we integrate in the process, in a stylized manner, those political decisions that can impact the network evolution in a defined area.

We measured the degree distribution (fig. 3, left) and the betweenness centrality distribution (fig. 3, right) (Porta et al. 2006) of the street network of nowadays and the networks after four simulations. Simulated networks are obtained with four different patterns. They conserve the main characteristics of the starting network with variations. Due to our decentralized approach, we observe that the growth of Fécamp town move toward an organic form, rich of tree-like networks and bifurcations. This process is captured by both distributions, where we observe an increment of vertices with degree 3 and a more hierchized distribution of the betweenness centrality.

4 Discussion and conclusion

This work proposes a street network generator model. Inspired by complexity theories, dynamics are completely decentralized and driven by feedback mechanisms between layers. An external control can be eventually integrated by imposing over the environment some probabilities of growth. The application of the model to study the evolution of the Fécamp town suggests that the growth of a street network

is the result of an unpredictable combination of rate of growth, exogenous factors and feedback. The main properties of generated graphs are similar to real street networks. In this context, morphogens can represent real activators and inhibitors of urban growth, e.g. population, economical factors and political actors. We think the model may be helpful both to investigate urban growth and to support decisions of urban planners.

The fundamental brick of our stylized and general approach is that the morphogenesis of the street network is driven by morphogens. They are spatially embedded, interact each others, are involved in competition/cooperation processes, move and may arranged with regularity. We can find a correlation with those dynamics and population or economical actors of cities. We can suppose that it exists a relation between the concentration of people in a region of space and the existence of buildings and streets. In this field, population could be considered for the street network as a morphogen because it affects the network growth. The spatial relation between the density of population and the street network is not in all situations ensured; it is still unpredictable, rarely synchronic, and often affected by a large amount of socio-economical and natural factors. Those preliminary observations open to many perspectives. We start a process of validation. More precisely, we plan to test our model in other urban contexts and validate our approach with a comparison between different real case studies.

References

- Batty, M. (2013). *The New Science of Cities*. Publisher: The MIT Press. Cambridge: The MIT Press.
- Bourgine, P. and A. Lesne, eds. (2011). *Morphogenesis: Origins of Patterns and Shapes*. Berlin Heidelberg: Springer-Verlag.
- D'arcy, T. (1917). *On Growth and Form*. Cambridge: Cambridge University Press.
- Gray, P. and S. K. Scott (1983). "Autocatalytic reactions in the isothermal, continuous stirred tank reactor: Isolates and other forms of multistability". In: *Chemical Engineering Science* 38.1, pp. 29–43.
- Marshall, S. (2005). *Streets and Patterns*. 1 edition. Abingdon and New York: Sporn Press. 336 pp.
- Morin, E. (2005). *Introduction à la pensée complexe*. Paris: Points. 160 pp.
- Pearson, J. E. (1993). "Complex Patterns in a Simple System". In: *Science* 261.5118, pp. 189–192.
- Porta, S., P. Crucitti, and V. Latora (2006). "The Network Analysis of Urban Streets: A Primal Approach". In: *Environment and Planning B: Planning and Design* 33.5, pp. 705–725.

- Portugali, J. (2006). “Complexity Theory as a Link between Space and Place”. In: *Environment and Planning A* 38.4, pp. 647–664.
- Pumain, D. and R. Reuillon (2017). *Urban Dynamics and Simulation Models*. Lecture Notes in Morphogenesis. Springer International Publishing.
- Pumain, D., L. Sanders, and T. Saint-Julien (1989). *Villes et auto-organisation*. Paris: Economica. 191 pp.
- Tirico, M. (2020). “Morphogenesis of complex networks. An application in urban growth”. PhD thesis. Le Havre: Le Havre Normandy University. 203 pp.
- Tirico, M., S. Balev, A. Dutot, and D. Olivier (2019). “Morphogenesis of Complex Networks: A Reaction Diffusion Framework for Spatial Graphs”. In: *Complex Networks and Their Applications VII*. Springer International Publishing, pp. 769–781.
- Turing, A. M. (1952). “The Chemical Basis of Morphogenesis”. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237.641, pp. 37–72.
- Wolfram, S. (2002). *A New Kind of Science*. 1 edition. Champaign: Wolfram Media. 1192 pp.

Road network characterization to understand city evolution.

Hanae EL GOUJ¹, Paul BOUMAN², Liubov TUPIKINA^{3,4}, Claire LAGESSE¹

¹Université Bourgogne Franche-Comté

32 rue Mégevand 25030 Besançon, France

hanae.el_gouj@univ-fcomte.fr, claire.lagesse@univ-fcomte.fr (correspondent author)

²Econometric Institute, Erasmus University Rotterdam

Burgemeester Oudlaan 50, 3062PA, Rotterdam, The Netherlands

bouman@ese.eur.nl

³Nokia Bell labs

Paris, France

⁴Centre de Recherche Interdisciplinaire

8bis rue Charles V, 75004 Paris, France

liubov.tupikina@cri-paris.org

Keywords: City morphogenesis, road networks, spatial analysis, graph theory

1 Introduction

Cities combine different kind of processes that drive their evolution. It is of a great interest to understand the effects of different kinds of spatial patterns on those processes, to be able to minimize space consumption and optimize mobility and transportation.

There are various approaches to model and analyze the evolution of cities. Hillier (1987) proposed the space syntax as a method for analysing spatial patterns; by adopting the tools and methods of graph theory, modelling the city by axial maps.

Based on this pioneering work, Porta et al. (2006) and Jiang and Claramunt (2004) developed the morphological approach where the spatial structure of the city is studied via a graph representing the road network, where arcs are road segments and nodes are intersections. Lagesse et al. (2015) introduces a geographical object called the *way*, which is built by iterative combination of road segments. By using ways, network indicators are less time-consuming to compute, simpler to analyze, and reveal many features of the road patterns without being sensitive to border effects. Barthélemy and Flammini (2008) proposed a computational simulation of road network growth via the creation of new segments connecting to the previous ones from points distributed in space according to a density law. Following this work, Courtat et al. (2011) introduced a simulation model with simple morphological rules. New settlements and road segments are iteratively added based on a potential function. Pousse (2020) proposed a theoretical network growth model of road networks that includes multi-scale aspects.

In this research, we aim to contribute to the understanding of city evolution by a focus on the road network of the North-Eastern French city of Dijon's. Firstly, we select three network indicators that can be used to characterize the city structure as it grows. We apply these to Dijon's historical road networks from 1650 until 2019. Finally, we discuss how we can use computational simulations of city growth with these indicators to explore the parameters leading different evolution processes. Prospectively, we aim to quantify the difference between the trajectories of these indicators, for different cities, through time.

2 Methodology and results

2.1 Indicators based on ways

We apply the method proposed by Lagesse et al. (2015): road networks are modeled as hyper-graphs composed of nodes and *ways*. Ways are constructed from the road network by combining consecutive road segments according to their minimum deviation angle (below a sixty degree threshold) at each intersections. Local indicators (depending on the way itself and on its direct surroundings) are then computed on ways.

Indeed, Strano et al. (2012) and Barthélemy et al. (2013) show that local indicators can be particularly useful to assess evolutionary mechanisms related to local densification and urban expansion. Inspired by this observation, we focus on three indicators that characterize city

structure and their evolution over time: *degree centrality*, *orthogonality*, and *spacing*.

The *degree centrality* indicator corresponds to the number of other ways a certain way intersects.

This indicator characterizes main structural properties of the network.

Orthogonality characterizes a way according to its angle of connection with its direct neighborhood. It highlights how each way is included in the network, between two extreme cases: highly meshed structures such as the perfectly regular Manhattan grids; and flexible connections such as tree-like slums Lagesse (2015).

Spacing defines the average distance between two connections by dividing the way length (in meters) by its connectivity. This characterization can be used to identify parts where the road network is concentrated in dense clusters, and those with wider spacing such as the ring roads Lagesse (2015).

2.2 *Analysis of Dijon*

Applying these indicators to the road network of Dijon, we observe in Figure 1 that ways with high degree centrality and low orthogonality correlate with tree-like structures, whereas ways with high degree centrality and high orthogonality correlate with gridded structures.

Next we consider how different indicators vary as the city grows. We visualize this process by plotting a trajectory of the coefficient of variation of two chosen indicators against each other through time. Figure 2 applies this approach on Dijon's road network at eight points in time, based on two combinations of indicators: spacing versus orthogonality and degree centrality versus orthogonality. In both cases, we observe a clear linear trend between these indicators, with some interesting deviations.

In Figure 3 we observe that ways with high spacing and low orthogonality indicators correlate with urban expansion, whereas ways with high orthogonality and low spacing correlate with densification.

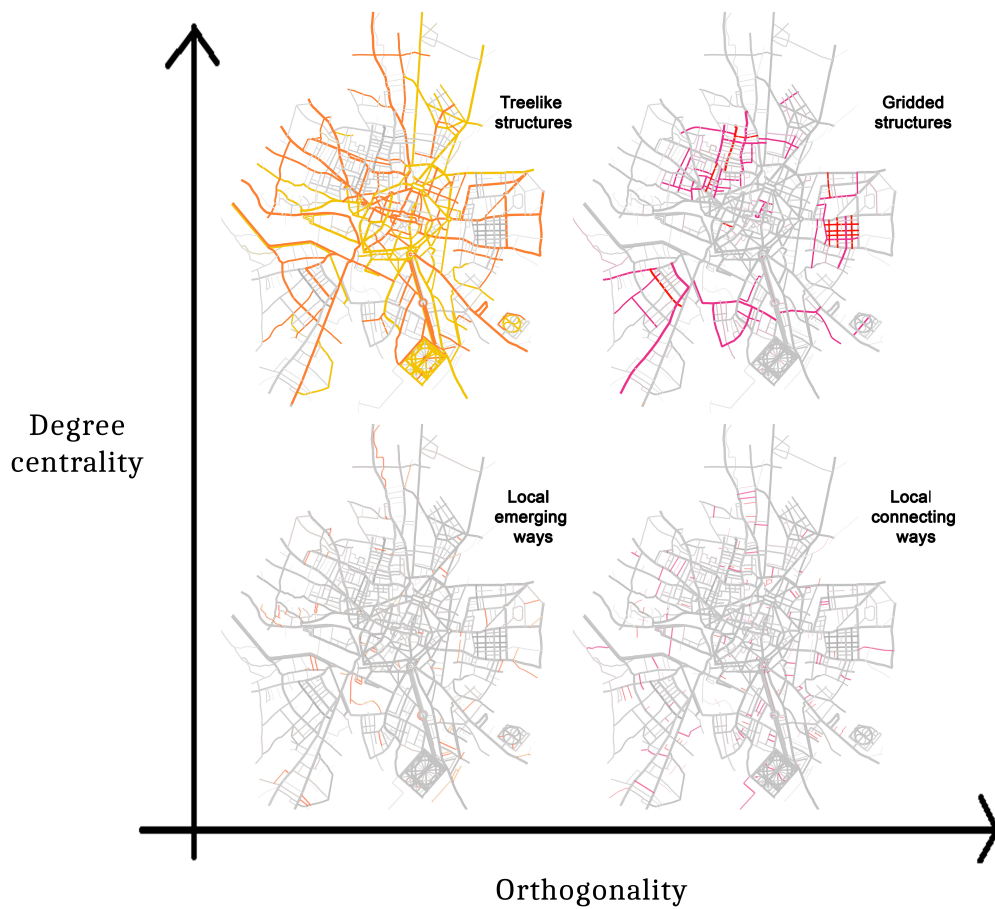


Figure 1: Highlight of the treelike and the gridded structures based on crossing degree centrality and orthogonality indicators on the 1944's network of Dijon.

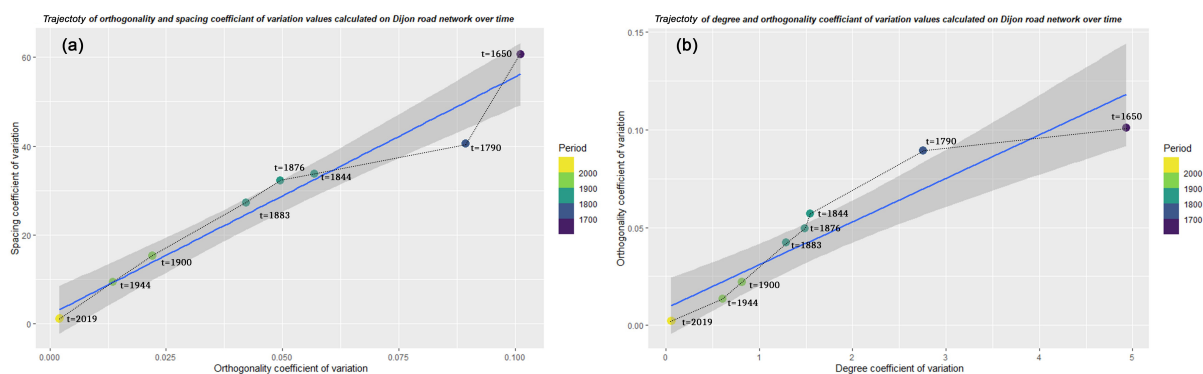


Figure 2: Trajectory of Dijon road network evolution from the 17th to the 21st century based on the coefficient of variation: (a) spacing and orthogonality, (b) orthogonality and degree centrality. The coefficient of variation is measured by dividing the standard deviation of each indicator by the number of ways.

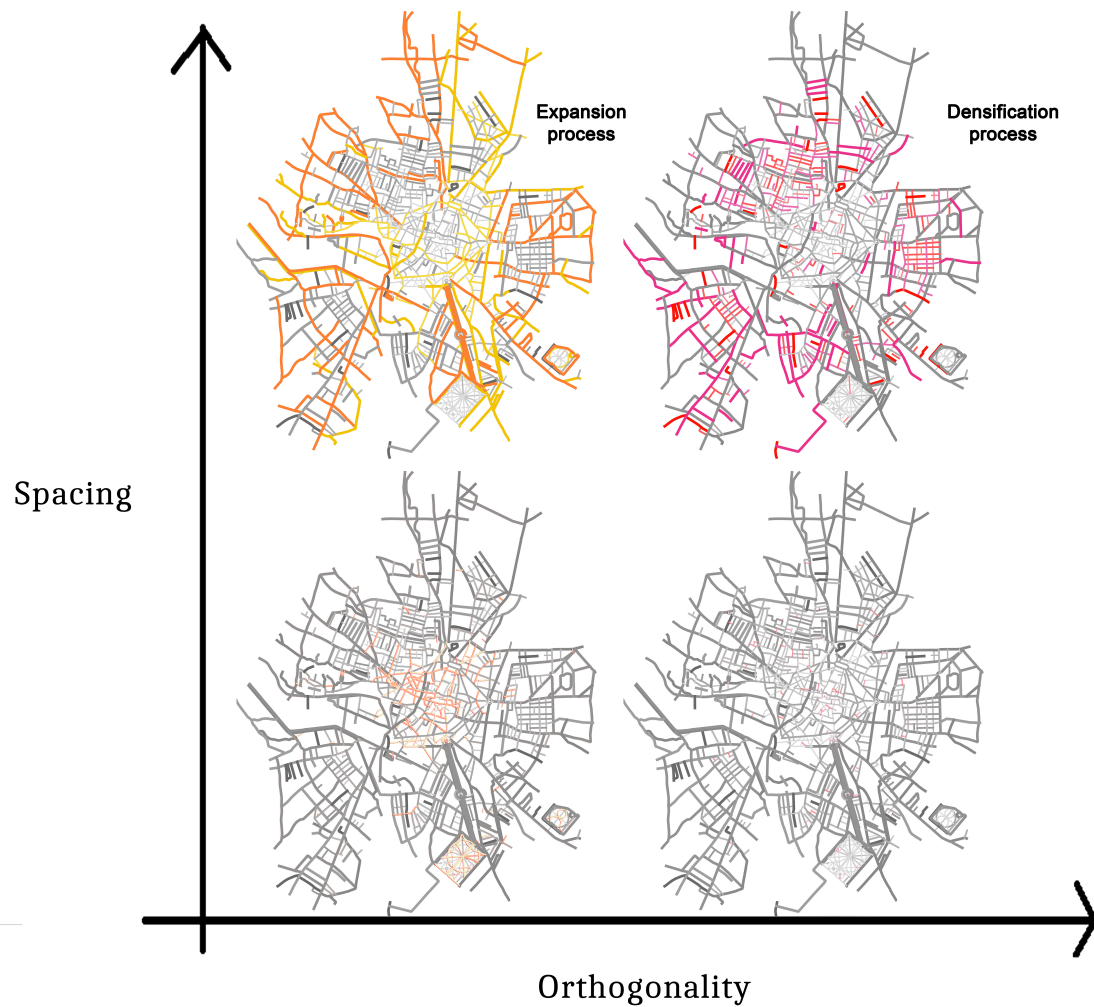


Figure 3: Highlight of the exploration and densification processes based on crossing spacing and orthogonality indicators calculated on the historical network of Dijon in 1944.

2.3 City growth simulation

Since our exploratory analysis only considers a single city and the collection of historic data of actual road networks is a time consuming process, we intend to combine the analysis of actual data with simulated city growth models using the model from Courtat et al. (2011) as implemented by Graftiaux (2021). With this model we can vary four simulation parameters: the *planning type* controls whether the city grows organic or centralized structures, the *construction factor* controls the number of roads added in each step, the *organization factor* determines the effort spent on finding good locations for expansion and the *sprawling factor* controls whether expansion focuses on inner development or expansion outside current boundaries.

3 Discussion and conclusion

Cities can evolve in different manners, e.g. spontaneously acquiring its irregular shape by following the topographical features of the landscape or via conscious planning where a regular, symmetrical form is imposed. As they are growing, a complex combination of planned gridded structures and organic structures emerge (Buhl et al., 2006).

In this research we analyze the evolution of the road network of Dijon through indicators based on the *ways* of the city. Using three indicators, different kinds of structures are highlighted in the road network. Our approach can be generalized to different cities as long as historical road data is available, as well as to simulated networks. It is then possible to compare the trajectories through time of different cities, according to the evolution of their road networks characterization. This type of analysis can also be performed with other indicators, such as closeness centrality and betweenness centrality.

For future research we aim to apply these indicators on more road networks, both from actual cities and simulations, and search for patterns that provide more insight in the different processes driving cities evolution.

References

- Barthélemy, M., Bordin, P., Berestycki, H. and Gribaudi, M. (2013), ‘Self-organization versus top-down planning in the evolution of a city’, *Scientific reports* **3**(1), 1–8.
- Barthélemy, M. and Flammini, A. (2008), ‘Modeling urban street patterns’, *Physical review letters* **100**(13), 138702.
- Buhl, J., Gautrais, J., Reeves, N., Solé, R. V., Valverde, S., Kuntz, P. and Theraulaz, G. (2006), ‘Topological patterns in street networks of self-organized urban settlements’, *The European Physical Journal B-Condensed Matter and Complex Systems* **49**(4), 513–522.
- Courtat, T., Gloaguen, C. and Douady, S. (2011), ‘Mathematics and morphogenesis of cities: A geometrical approach’, *Physical Review E* **83**(3), 036106.
- Graftiaux, M. (2021), The impact of structure and growth parameters on street networks’ efficiency, Master’s thesis, Erasmus University Rotterdam.
- Hillier, B. (1987), ‘La morphologie de l’espace urbain: l’évolution de l’approche syntaxique’, *Architecture et Comportement/Architecture and Behaviour* **3**(3), 205–216.
- Jiang, B. and Claramunt, C. (2004), ‘Topological analysis of urban street networks’, *Environment and Planning B: Planning and design* **31**(1), 151–162.
- Lagesse, C. (2015), Lire les lignes de la ville: méthodologie de caractérisation des graphes spatiaux, PhD thesis, Sorbonne Paris Cité.
- Lagesse, C., Bordin, P. and Douady, S. (2015), ‘A spatial multi-scale object to analyze road networks’, *Network Science* **3**(1), 156–181.
- Porta, S., Crucitti, P. and Latora, V. (2006), ‘The network analysis of urban streets: a primal approach’, *Environment and Planning B: planning and design* **33**(5), 705–725.
- Pousse, R. (2020), Caractérisation et modélisation du réseau viaire, PhD thesis, Université de Paris.
- Strano, E., Nicosia, V., Latora, V., Porta, S. and Barthélemy, M. (2012), ‘Elementary processes governing the evolution of road networks’, *Scientific reports* **2**(1), 1–8.

Parallel Session PSA3

3 November 2:30-5:00 pm (GMT)

Scaling Laws and Urban Form

Convolutional neural networks for building types classification in Switzerland

Raphaël BUBLOZ¹

¹Institute of Geography and Sustainability, University of Lausanne, Switzerland

UNIL – IGD, Géopolis, CH-1015 Lausanne, Switzerland

+41 21 692 43 76, raphael.bubloz@unil.ch

Keywords: Deep learning, Built environment, Building typologies, Open data, Switzerland

The collection and provision of three-dimensional vector data sets of buildings at the scale of entire cities or countries is becoming increasingly common. These building models, whose advancing accuracy allows the inclusion of the shape of roofs and sometimes even balconies and chimneys, support the development of opportunities in various fields such as the estimation of photo-voltaic energy production, the study of urban micro-climates or the modelling of the impacts of natural events. However, these 3D data rarely have attribute information on the function or type of the buildings in question, although this would help to better target the modelling results.

Several studies have sought to classify urban building types based on topographic data such as (raster) maps or 2 or 3D vector databases (e.g. Henn *et al.*, 2012; Hecht, Meinel and Buchroithner, 2015). The authors usually use methods that require the inclusion of additional geographical information such as land use category or proximity to specific points of interest (university, kindergarten, hospital). They also point out the multiplicity of definitions of building 'types', the need for a study area with sufficient labelled data and the possibility for further research with more detailed 3D models.

At the same time, the development of computing power and the ever-growing availability of images (aerials or street views) have led to an increase in computer vision and Convolutional Neural Networks (CNNs) use in urban studies (Ibrahim, Haworth and Cheng, 2020). CNNs can notably be trained to classify images and have been applied to land-use classification or cartographic building generalisation.

With the upcoming of three-dimensional building data now incorporating roof details, it is possible to retain a height information while transforming 3D vector data into a 2D image (matrix). This conversion allows to consider using CNNs to perform classification of building type and let the models focus primarily on the shape of the building and the morphology of its immediate surroundings without needing to consider extra POIs. This research will set out to implement building type classification using CNNs on open data consisting of 3D building models and road networks.

1. Materials and methods

The chosen study area for this analysis is Switzerland as 3D buildings models and a comprehensive federal register of buildings are available for the whole country. Switzerland's territory offers a variety of regions ranging from highly urbanised plains to rural areas with less dense constructions and even more isolated structures in some mountainous zones.

1.1 Datasets

Three building datasets were used in this study. The first dataset is derived from the Register of Buildings and Dwellings maintained by the Swiss Federal Statistical Office (FSO). It contains information on buildings, dwellings, projected works, entrances and streets. For buildings, notable data are the construction year, category and class, and projected coordinates of the building's centre.

The second dataset is the swissBUILDINGS3D 2.0 set of the Swiss Federal Office of Topography (swisstopo). It includes a 3D model of buildings, with roofs and height modelised, corresponding to a Level of Detail 2 (LoD2) in CityGML. Buildings inferior to a 3m x 8m surface are not included except for some cartographic (e.g. tower, silo) or circumstantial exceptions (e.g. tiny building in a small village).

The final dataset was extracted from OpenStreetMap, a collaborative geographic database freely editable and accessible. Among the numerous elements available, footprints of buildings and roads located in Switzerland were downloaded. Figure 1 offers an illustrated view of the geographic datasets with the 3D buildings in grey/white and the roads in black.

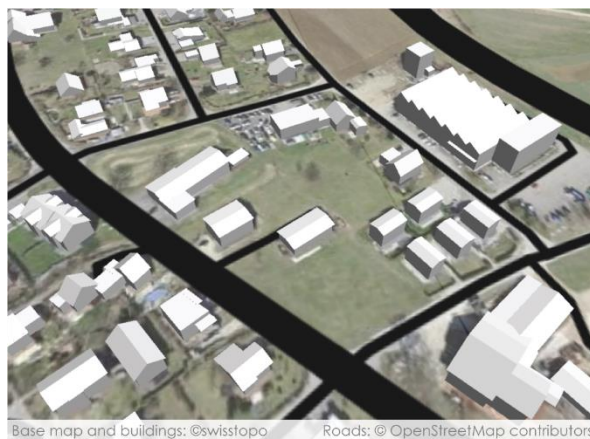


Figure 1: Overview of the geographic datasets

1.2 Creating an image dataset

At the time, there is no one-to-one official relation between a building as defined by the FSO and a building as modelised by swisstopo. It is therefore necessary to spatially combine the statistical information using the registered x,y coordinates. If there is a unique match on the building footprint,

the class and category are assigned to the target geometry. If there are several data points falling on one geometry, class and category are assigned only if they are, respectively, the same for all points. Geometries with several class or categories or no matching record are discarded.

Remaining geometries are then transformed in squared 128 x 128 matrices with a spatial resolution of 1 meter, centred on the building's centroid. The difference of height between the building's base and the roof is assigned to each pixel. The height are then normalized between 0 and 1, where 1 corresponds to the greatest height of all buildings' pixels. Surrounding buildings' coverage is saved in a matrix of the same dimension, with a resolution of 4 meters. This also applies to roads with a resolution of 3m. The values of the coverage are also in the range 0 to 1, where a value of 0.75 indicates that roads or buildings make up three quarter of the pixel's surface. Finally, each image constitutes of a 128x128x3 array and a unique class and/or category label. Figure 2 depicts all three matrices for an industrial building (visible in the top right of Figure 1) with a characteristic saw-tooth roof.

Three channels of a final image

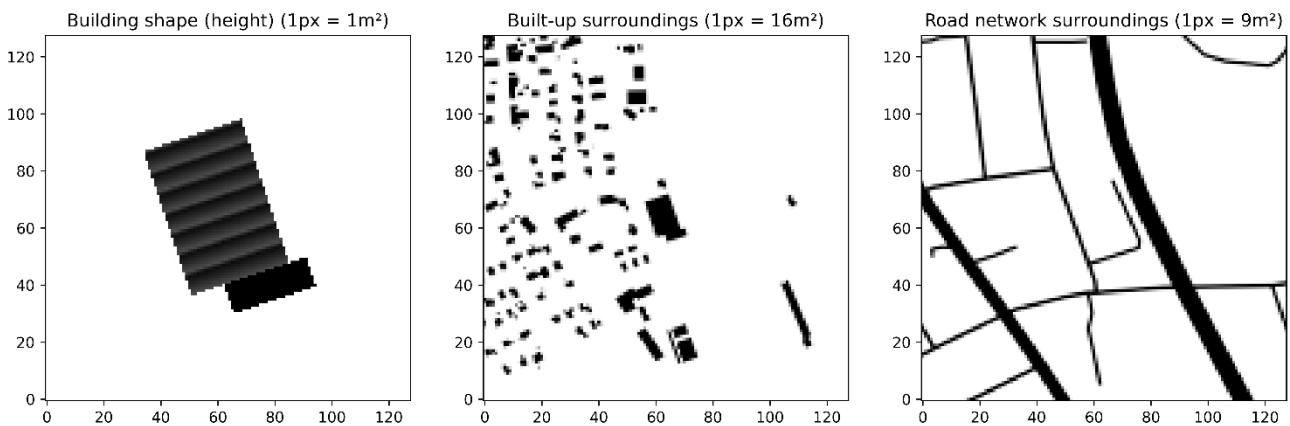


Figure 2: Decomposition of an image into its three channels

2. Challenges and expected results

In this ongoing research, the exact outlines of the CNN architecture are yet to be defined. More than 2 million images were obtained following the procedure above and were split in independent training (80%), validation (10%) and testing (10%) sets. Due to the nature of the thematic, the categories show a heavy class imbalance as, without surprise, building exclusively for residential use have far higher occurrences as other categories. This issue may influence the way to perform sets creation or the metrics used to evaluate the model performance.

The accuracy of the model may also suffer from the fact that differences between categories are tenuous. An apartment complex with or without shops at ground floor will not be in the same category and may yet have similar shape and surroundings. Moreover, buildings in the same class, such as religious buildings, can have quite disparate shape.

A successful application of deep learning focusing on the shape of building and morphology of surroundings could open the way for further predictive and generative research in the architectural and urban research fields.

References

- Hecht, R., Meinel, G. and Buchroithner, M. (2015) ‘Automatic identification of building types based on topographic databases – a comparison of different data sources’, *International Journal of Cartography*, 1(1), pp. 18–31. doi:10.1080/23729333.2015.1055644.
- Henn, A., Römer, C., Gröger, G. and Plümer, L. (2012) ‘Automatic classification of building types in 3D city models: Using SVMs for semantic enrichment of low resolution building data’, *GeoInformatica*, 16(2), pp. 281–306. doi:10.1007/s10707-011-0131-x.
- Ibrahim, M.R., Haworth, J. and Cheng, T. (2020) ‘Understanding cities with machine eyes: A review of deep computer vision in urban analytics’, *Cities*, 96, p. 102481. doi:10.1016/j.cities.2019.102481.

A data-driven methodology to identify urban form and characterize its functionality

Roger Vazquez¹; Roberto Ponce-Lopez^{1,2}

¹Tecnologico de Monterrey

Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849

rvztz@tec.mx

²Institute for the Future of Education

Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849

rpl@tec.mx (correspondent author)

Keywords: Functional zones, urban form, points of interest, machine learning, natural language processing

1. Introduction

Urban form is critical for the analysis of spatial structures and for planning land use and transportation for a city. A proper characterization of the morphology of the city requires the identification of the main subcenters of activity, and second, their characterization based on the profile of jobs and visitors who arrive either for work or leisure. Research on urban form has proposed data-driven approximations to analyze urban form. Natural language processing methods typically used for text mining in combination with the widespread availability of geocoded point-of-interest (POIs) data have led to the development of several functional zone inferring techniques. Our paper builds on such methods to propose a data driven method to study urban form using the density and topology of POIs.

2. Literature Review

Yuan et al. (2012) split a city into different areas based on major roads, each region containing a particular set of points-of-interest with text-based metadata describing the category of the POI. Then, each region is considered as a “document” --similar to text mining- and the distribution of functions is regarded as a distribution of topics to be discovered by a text-based topic-inference model. Gao et al. (2017) and Hu et al. (2019) take a similar approach by randomly selecting POIs within a given area and building up the dataset by selecting points near the sampled ones; the label data of each point is used to discover a distribution of topics in the data. The resulting topics are clustered into groups, which are regarded as the typology of urban function in the study area.

There are two problems with such quantitative approximation found in the literature. First, most of the proposed frameworks require thorough examination of the area of study or previous knowledge of the area to parametrize the NLP models that produce an output as a characterization of urban form and functionality. Second, the literature that characterizes urban form oversimplifies the delimitation of the zones of activity by assuming pre-existing administrative units or defining *ad hoc* units made by the convergence of major roads. Our work contributes to solve those two restrictions.

This research introduces a combined framework to learn functional zones of urban areas based on big and geocoded data. First, our framework uses a spatial clustering algorithm with built-in stability measures to define the unit and scale of analysis. Second, it applies the NLP models of topic inference to such identified units in the previous step to categorize their functionality. The pipeline takes a bottom-up approach: it constructs the geospatial units of analysis through a density-based algorithm on the points-of-interest. Then, it infers the topic distribution of the labels that describe the type of business of the same points-of-interest within the unit of analysis defined in the first step. The main contribution of this paper is a systematic and replicable method to study form and functionality based on POI data.

2. Methodology

Our methodology consists of the combination of two already established techniques in machine learning: a point density algorithm identifies the subcenters of activity; a NLP algorithm characterizes the content of the identified subcenters using the mixture, density and diversity of the POIs.

2.1 Identifying subcenters of activity

The algorithm to identify subcenters of activity is Approximate-DBSCAN (Arribas-Bel et al. 2019). It is a variation of the density-based clustering algorithm DBSCAN which consists of an ensemble of traditional DBSCAN replications calculated for a random sample of the total dataset where only the labels that show in the majority of the solutions are kept. This process ensures that the density-based identified subcenters have sufficient evidence to be a representation of underlying spatial structure. The output of this model is a geometry of points that enclose the sites with the highest density of POIs.

2.2 Characterizing subcenters of activity

Our method uses the well-documented Latent Dirichlet Allocation (LDA) algorithm on the POI, combined with the introduction of topic coherence metrics to find the optimal number of categories or topics. LDA constructs a number of topics using the labels of type of establishment from the POI within each cluster that resulted from the ADBSCAN.

LDA is an NLP processing technique (Blei et al. 2003), widely used in text documents to classify them by topic. The technique fits a model based on the occurrence and combinations of words within the documents and calculates a probability for each document to belong to each topic category. The critical part is to parametrize the model with the number of topics that we want beforehand. The CV topic coherence metric is used to find out the optimal number of topics. In our case of application, the ADBSCAN identified zones are the equivalent to text documents; and the labels of the POIs within the ADBSCAN clusters are the equivalent to the words within documents. The outcome of this technique is a classification of the ADBSCAN clusters defined by the text patterns from the POIs' description labels for type of business.

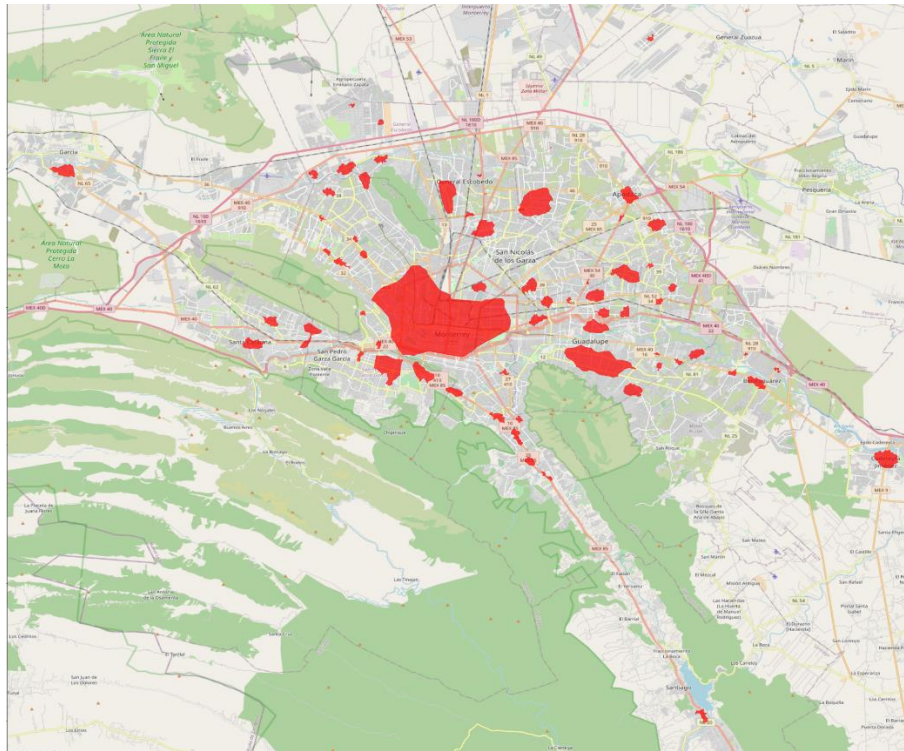
2.3 Case Study and Data

We show the effectiveness of our work by applying the pipeline to a set of 169,020 POIs to characterize urban form in the Metropolitan Area of Monterrey, Mexico. Monterrey is a monocentric city transitioning to polycentricity. It is the second largest metropolitan area in Mexico with 5.5 million inhabitants, and it is the industrial hub of the country. There are 18 municipalities that compose the metropolitan area. The POIs are establishments with latitude and longitude and type of business. The coding and labels for type of business follow the North American Industry Classification System (NAICS). The National Census Bureau of Mexico collects and releases these data every six months.

3. Results

3.1 Identifying subcenters of activity

The pipeline used for our case study resulted in 66 interpretable subcenters of activity that concentrate 37% of the establishments in metropolitan area (see Map 1).



Map 1: Resulting 66 subcenters of activity that concentrate 37% of the jobs in the metro area

3.2 Characterizing subcenters of activity

The LDA model was trained with a learning decay of 0.5. Figure 1 shows the highest coherence scores ranged from 5 to 7. After comparing the results with those parameters, we trained LDA with 6 topics.

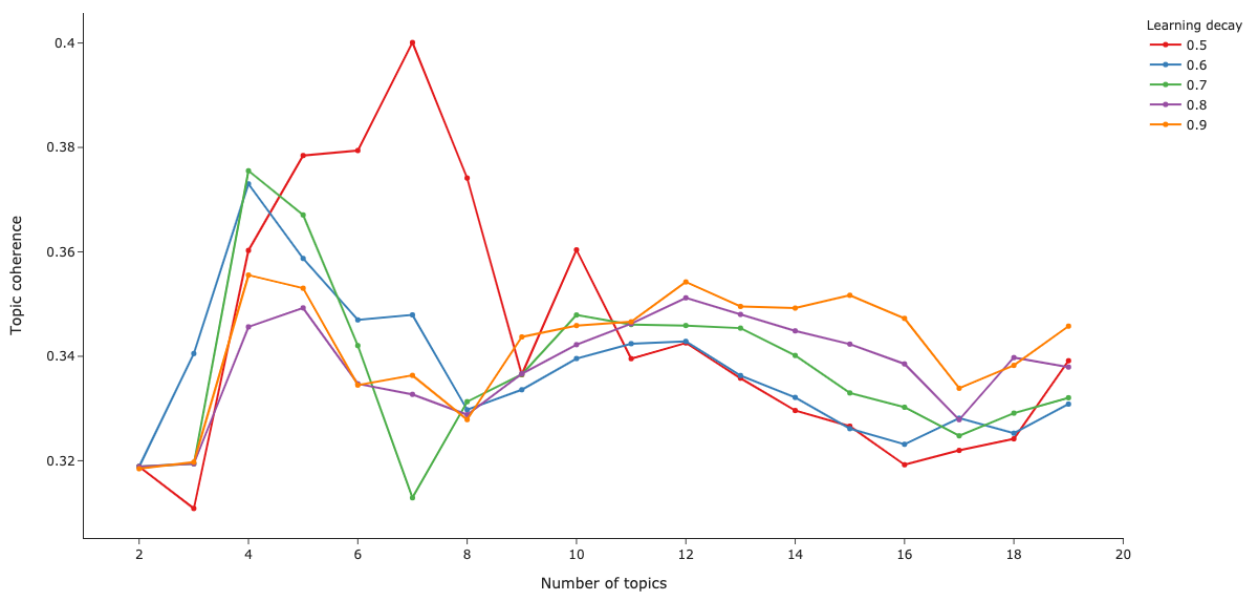


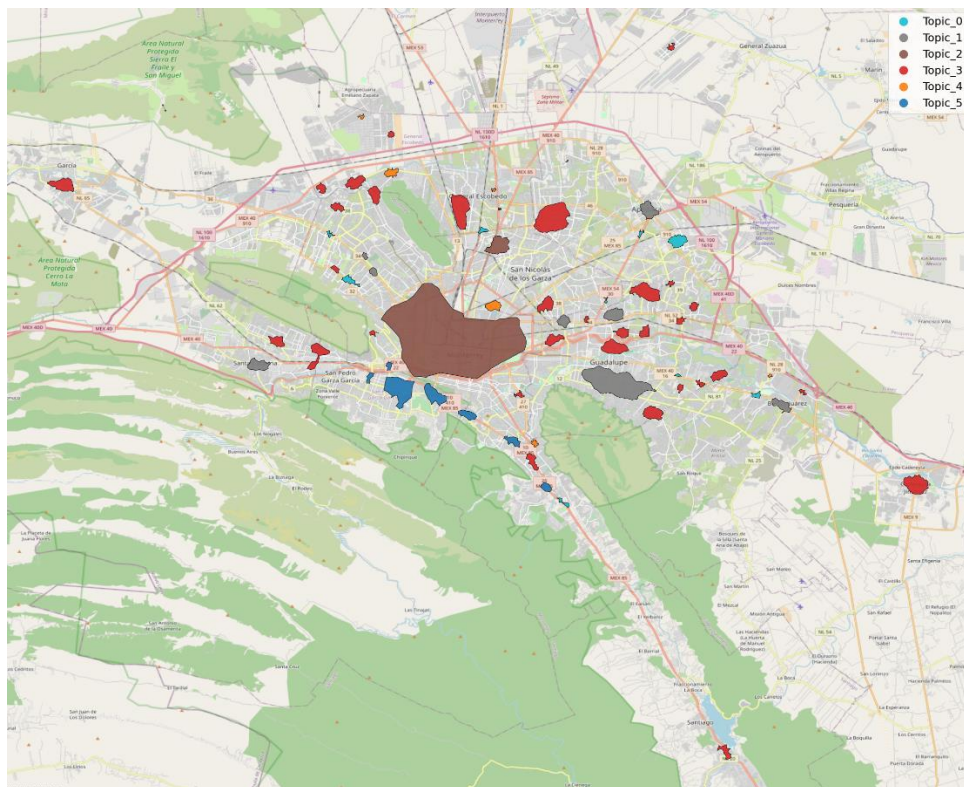
Figure 1: Coherence scores for a given number of topics.

Table 1 shows the most relevant words for each of the topics, by ranking order. The rank and mixture of businesses reflect the degree of spatialization by each category or topic.

Rank	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	trade	services	services	services	drinks	services
2	retail	trade	trade	retail	food	items
3	items	retail	retail	trade	trade	trade
4	services	food	items	drinks	tobacco	retail
5	textiles	drinks	drinks	food	ice	technical
6	products	preparation	food	items	groceries	scientific
7	jewelry	items	preparation	preparation	retail	professional
8	footwear	maintenance	repair	groceries	services	medical
9	clothing	repair	practice	personal	mayor	external
10	accessories	products	maintenance	ice	industry	use
11	drinks	personal	external	tobacco	preparation	practice
12	food	external	ice	repair	maintenance	recreation
13	preparation	medical	tobacco	maintenance	repair	stationery
14	stationery	tobacco	products	use	items	personal
15	personal	practice	groceries	personal	materials	products
16	use	ice	medical	recreation	agricultural	drinks
17	recreation	groceries	mayor	products	waste	footwear
18	brokering	financial	use	stationery	raw	textiles
19	credit	stationery	industry	clothing	forestry	accessories
20	institutions	educational	clothing	trading	supplies	clothing

Table 1: Most relevant words per topic.

Finally, Map 2 presents the subcenters of activity by each topic. The topic 5 is that with the largest degree of spatialization and corresponds to the wealthiest area of the city, that concentrates services, leisure and medical. The smallest patches where food and retail predominate locate mostly in the periphery and less wealthier areas of the city.



Map 2: Resultant 66 subcenters with their respective category.

4. Discussion

A benefit of our methods is that it is replicable to any city as it only needs POIs latitude and longitude, which are commonly available to any city. Our work is different from other research in that we constructed the units of analysis using the density of the same POI data rather than using predefined census data or other administrative units. We use a bottom-up approach to define the scale and size of the subcenters based on the same density of the points. Then, the POIs within the resulting units of analysis were mined to infer the functionality of the area, taking a systematic approach to define the number of categories for the typology of subcenters.

Acknowledgments

This research was funded by Fundación FEMSA.

References

Arribas-Bel, D., Garcia-López, M.À. and Viladecans-Marsal, E., 2019. Building (s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics*, p.103217.

Gao, S., Janowicz, K. and Couclelis, H., 2017. Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS*, 21(3), pp.446-467.

Hu, S., He, Z., Wu, L., Yin, L., Xu, Y. and Cui, H., 2020. A framework for extracting urban functional regions based on multiprototype word embeddings using points-of-interest data. *Computers, Environment and Urban Systems*, 80, p.101442.

Yuan, J., Zheng, Y. and Xie, X., 2012, August. Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 186-194).

Blei, D., Ng, A. and Jordan, M. 2003. Latent dirichlet allocation. In *Journal of Machine Learning Research* (pp. 993-1022).

A new city definition from the radial homothetic scaling law

Paul Kilgarriff¹; Remi Lemoy²; Geoffrey Caruso^{1,3}

¹Urban development and Mobility, Luxembourg Institute of Socio-Economic Research

11 Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg

p.kilgarriff1@gmail.com (correspondent author)

²University of Rouen

IDEES Laboratory UMR 6266 CNRS, France

³Department of Geography and Spatial Planning, University of Luxembourg

11 Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg

Keywords: city definition, delineation, scaling, spatial analysis, radial analysis

1. Introduction

How to define a city is a long-standing issue in geography and spatial analysis. An appropriate definition depends upon the research question and study area. Maunier (1910) wrote that a city definition should be (1) universal and common to all cities and (2) should be uniform everywhere.

One potential issue with using administrative local units to define cities across multiple regions is the introduction of bias either through the way in which the local units were created or using non-comparable local units that vary in size and scale (Dijkstra et al., 2020). The modifiable areal unit problem (MAUP) is a long standing issue for analysts (Gehlke and Biehl, 1934; Openshaw, 1984). Using a functional definition as opposed to an administrative definition, is one method of improving the overall fit (Veneri, 2016).

In Europe, the most widely used definitions of a city are provided by the EU-OECD definition of a functional urban area (Dijkstra et al., 2019; Dijkstra and Poelman, 2012) and the Global Human Settlement Layer (GHSL) definitions (Florczyk et al., 2019; Schiavina et al., 2019). Both provide definitions of the urban centre/city core and the wider Functional Urban Area/city region.

One criticism of functional approaches is the requirement of origin-destination commuting data. Such data however is not available in many countries of the developing world. The EU-OECD definition is based on an urban centre and a commuting threshold of 15% to that centre (Dijkstra et al., 2019; Dijkstra and Poelman, 2012). The GHSL-OECD Functional Urban Area definition was developed to provide comparable FUAs for world cities without the need for local unit or commuting information. Called eFUA, they are derived using GHSL population data, travel times (cell to cell) from the global friction matrix and Urban Centre GDP (Schiavina et al., 2019). The GHSL definition differs slightly from the EU-OECD urban centre as it also uses degree of urbanisation. In addition to functional approaches there are other alternatives. Rozenfeld et al. (2011) used the City Clustering Algorithm (CCA) to build cities from the ground up.

Our method differs from previous attempts at delineating cities in that we use a convergence procedure and scaling methodology. The convergence procedure arrives on the ‘best’ population and the scaling methodology ensures we control for initial city size, this makes comparisons of city metrics (density, land use) comparable across Europe.

2. Methodology

We utilize the EU-OECD FUAs as the baseline to obtain a starting population for a city. This is also the same definition used in the Urban Atlas. The population of the FUA is estimated by downscaling the GEOSTAT data at a 1km² resolution using Urban Atlas land use data, available at a finer 5 metre resolution. Land use categories are then weighted according to their use, this attributes the majority of a 1km² cell’s population into predominately residential land use classes. Urban Atlas and Corine Land Cover (CLC) data are also combined so there is coverage beyond the Urban Atlas area. Theoretically a city’s delineation might be larger in area than the current EU-OECD FUA definition on which the Urban Atlas area is calculated.

This analysis uses the homothetic scaling law devised by Lemoy and Caruso (2018). The population density of a city was found to scale with city size measured by its total population in a homothetic manner. population density is proportional to the cube root of total population. This is the standard relationship between the volume and the side length of a solid in three dimensions (cube or sphere for instance). This homothetic or isometric scaling uses a fixed factor for the entire considered system. In contrast allometry uses different rates of growth (Thompson; 1917; Huxley; 1932) for different parts of the system.

Lemoy and Caruso (2018) found that the radial population density profiles $p(r)$ of different cities are quite similar if the distance r to the city center is rescaled to a distance r' given by:

$$r' = r \times \sqrt[3]{\frac{N_{Paris}}{N}} = r \times k$$

where N is the population of the city being analysed and N_{Paris} is the population of the largest city in the dataset, used as a reference. k is the rescaling factor. For London and Paris, $k \approx 1$.

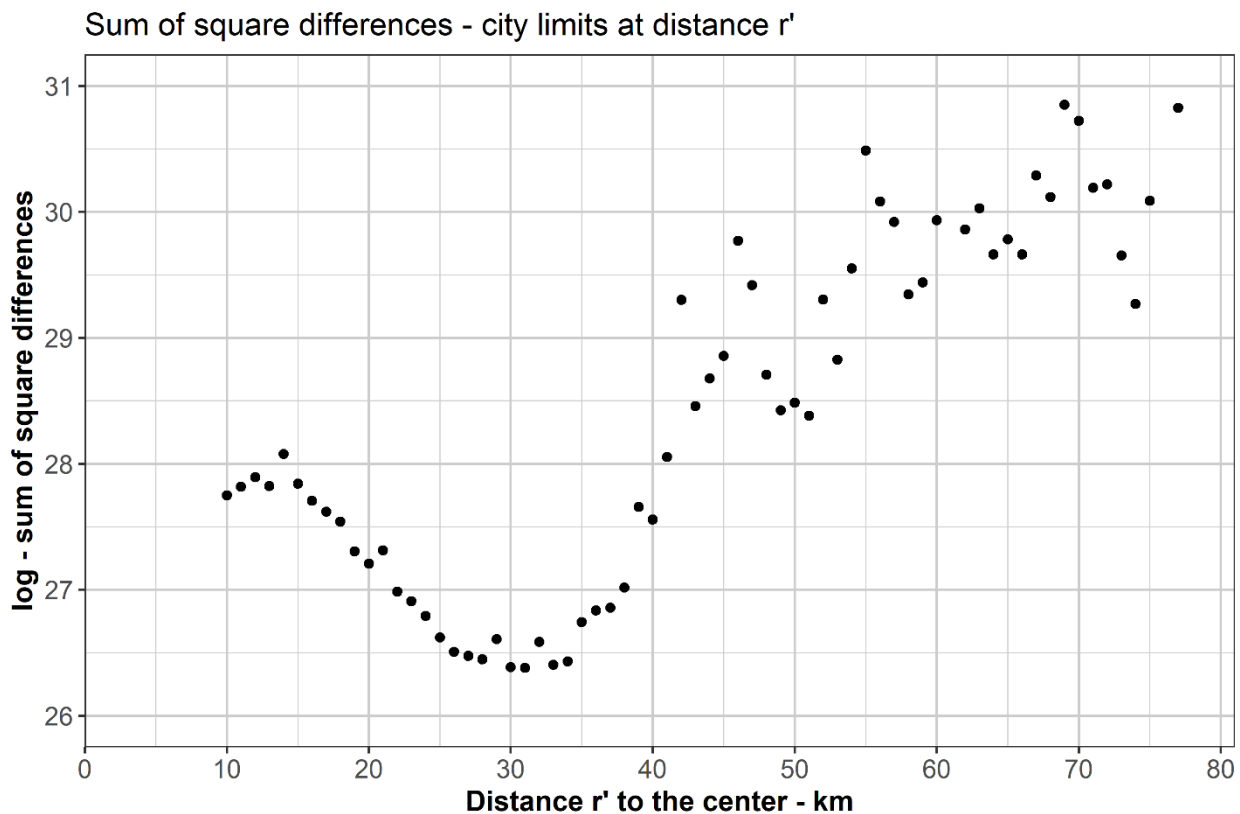
Using the population of the FUA as the starting point k is calculated for each city. Every distance 1-100km is examined. The following is example converging on $r'=40$ km:

1. Using the population of the FUA we calculate k for the city.
2. We rescale each distance from 0 – 100km using the scaling factor k .
3. We cut the city at $r'=40$ km and record the cumulative population at this rescaled distance.
4. Taking the population from step 3, we calculate k^{i+1}
5. We repeat steps 2, 3 and 4. The population typically converges on a population after 4-5 iterations. 15 iterations are performed to ensure all cities converge on a population.

The following procedure is repeated using every distance r' 0km to 100km.

3. Comparing the New City Definition – $r'=31$ km

Figure 1: Sum of square differences using different rescaled distances to define a city



We assess the convergence process examining the sum of square differences at distance $r'=x$ and $r'=x+1$. Figure 1 shows that the sum of square errors converge at $r'=31$ km. Taking this new distance $r'=31$ km as the delineation of a city, we compare with existing administrative and functional definitions and also examine the optimal city fringe distance.

A sensitivity analysis is performed between our city definition and those from the EU-OECD and GHSL. The first test examines the population within $r'=31$ km and the four other urban definitions. From figure 2 we can see that $r'=31$ km for cities is on average larger than both the EU-OECD urban centre and the GHSL urban centre. $R'=31$ km is also smaller than both the EU-OECD and GHSL FUA definitions for cities with the closest similarity to the GHSL FUA definition. There are however outliers in the case of all definitions.

Figure 2: Boxplots showing GHSL and UA populations as a share of $r'=31\text{km}$ population



Focusing on the comparison between $r'=31\text{km}$ and UA FUA (also called the EU-OECD definition), we examine the difference in the ranks for the 540 European cities.

Figure 3: Scatter plot city population rank UA FUA and $r'=31\text{km}$

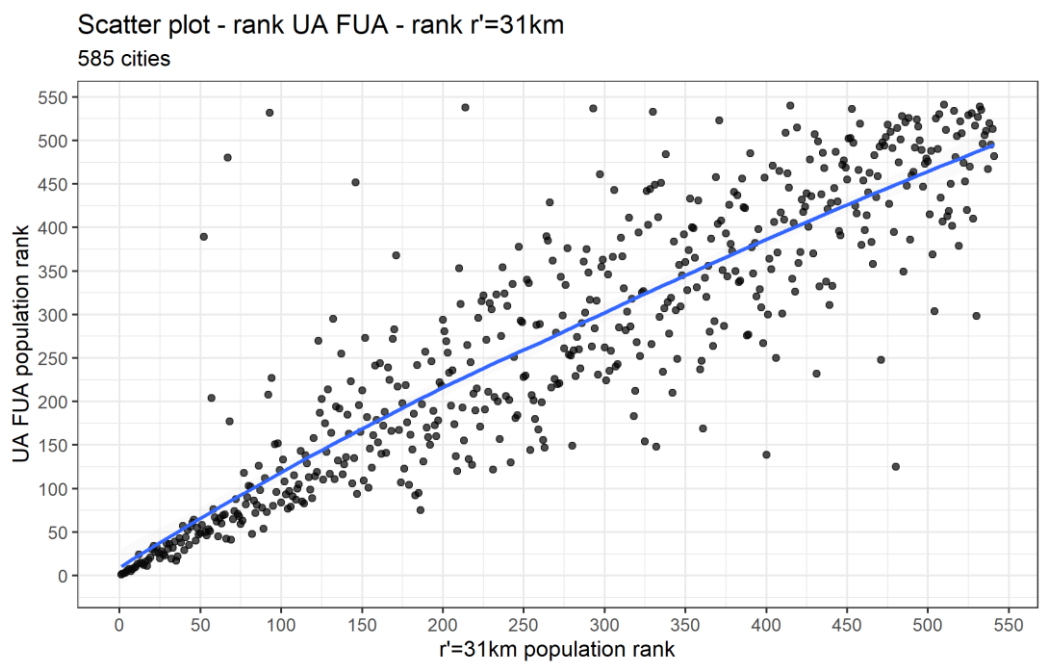
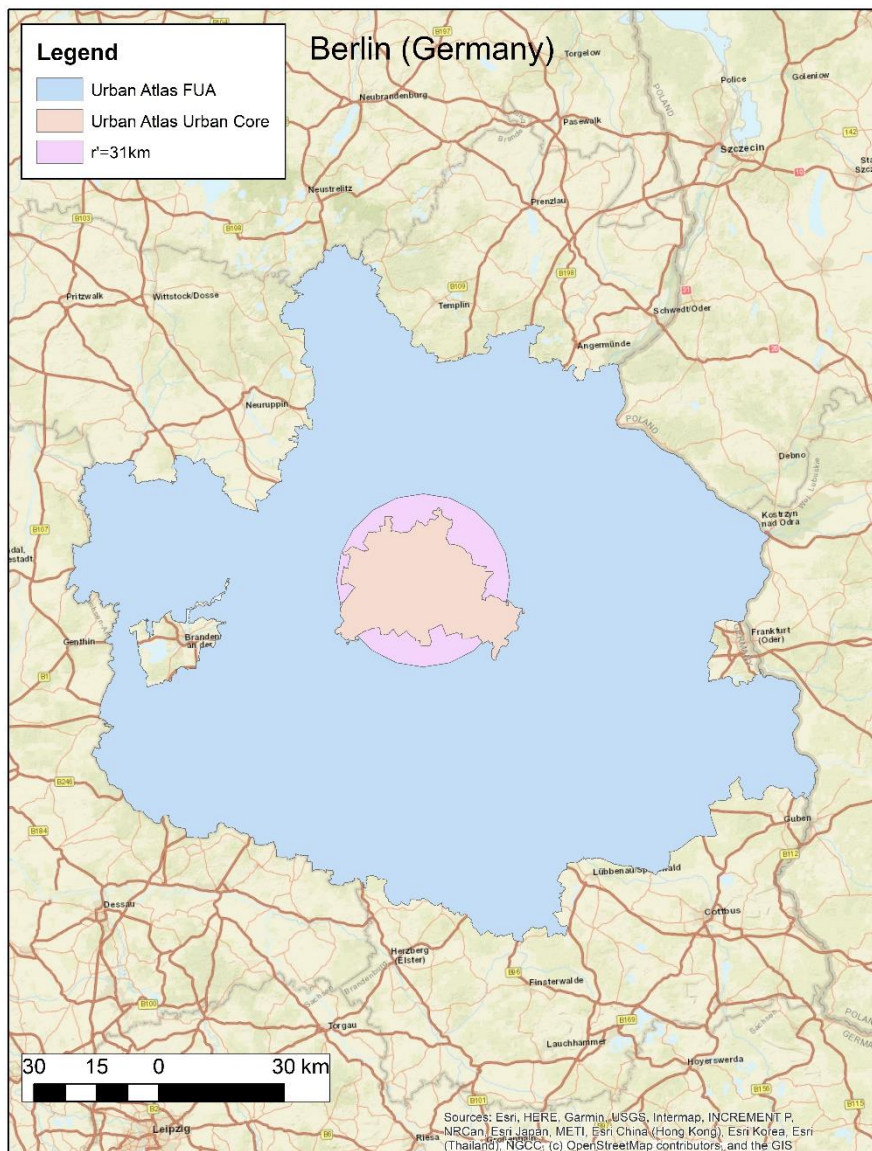
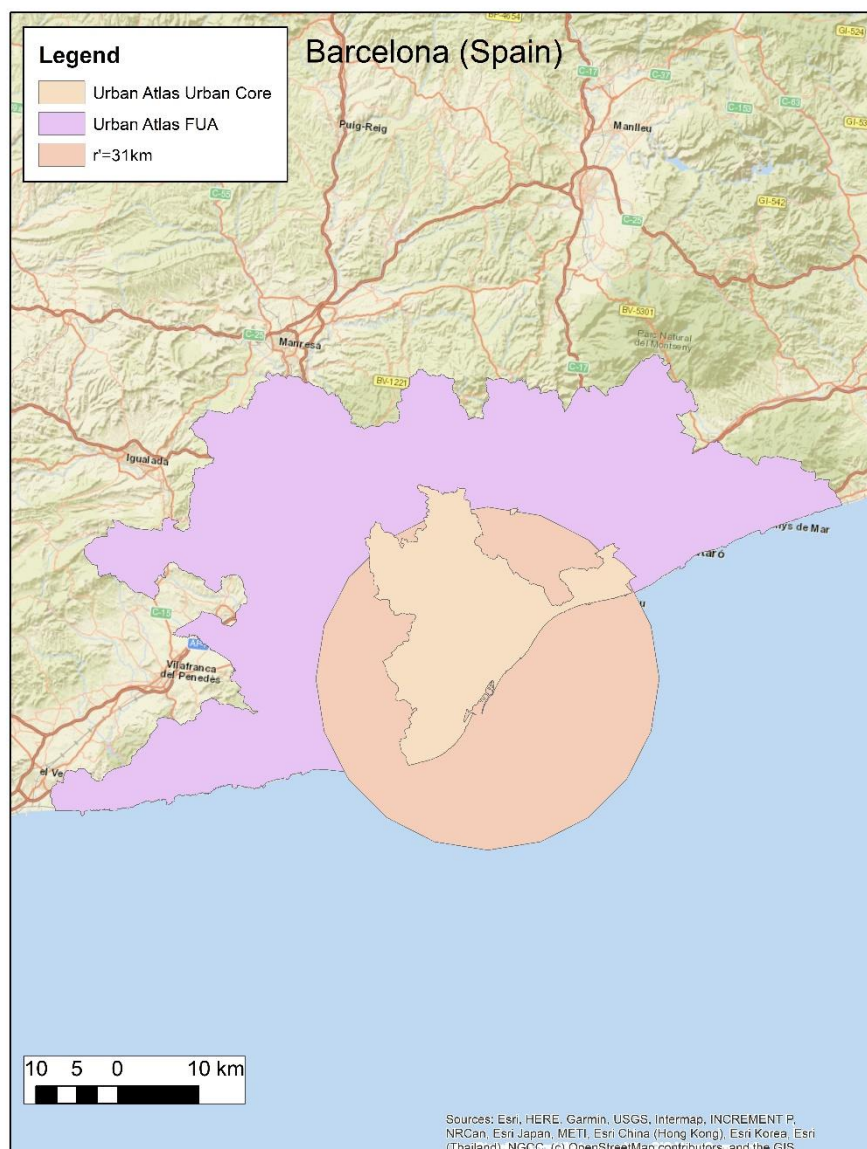


Figure 3 shows that there are several cities which have a higher population using the $r'=31\text{km}$ definition in comparison to the GHSL FUA definition. Using the converged distance of $r'=31\text{km}$ defines an area larger than the Urban Atlas Urban Core but an area smaller than the FUA, thus emphasising access and distance to the CBD.

2.1 City extents comparisons Urban Atlas and $r'=31\text{km}$



Map 1: Berlin Urban Atlas FUA and Urban Core. Shows the new definition highlights the city core effectively



Map 2: Barcelona Urban Atlas FUA and Urban Core. $r'=31\text{km}$ stresses the importance of access

Acknowledgments

This research was carried out as part of the SCALE-IT-UP project. The SCALE-IT-UP project is funded by the Luxembourg National Research Fund (FNR) under the CORE scheme. See: <http://www.fnr.lu/funding-instruments/core/>

References

- Dijkstra L and Poelman H (2012) Cities in Europe: the new OECD-EC definition. *Regional Focus* 1(2012): 1–13.
- Dijkstra L, Poelman H and Veneri P (2019) The EU-OECD definition of a functional urban area. OECD.
- Dijkstra L, Florczyk AJ, Freire S, et al. (2020) Applying the degree of urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*. Elsevier: 103312.
- Florczyk AJ, Melchiorri M, Corbane C, et al. (2019) Description of the GHS urban centre database 2015. *European*

Commission, Joint Research Centre (JRC) Technical Report - Public Release.

- Gehlke CE and Biehl K (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association* 29(185A). Taylor & Francis: 169–170.
- Lemoy R and Caruso G (2018) Evidence for the homothetic scaling of urban forms. *Environment and Planning B: Urban Analytics and City Science*. SAGE Publications Sage UK: London, England: 2399808318810532.
- Maunier R (1910) The definition of the city. *American journal of sociology* 15(4). University of Chicago Press: 536–548.
- Openshaw S (1984) The modifiable areal unit problem. In: 1984. Geo Abstracts University of East Anglia.
- Rozenfeld HD, Rybski D, Gabaix X, et al. (2011) The Area and Population of Cities: New Insights from a Different Perspective on Cities. *American Economic Review* 101(5): 2205–2225. DOI: 10.1257/aer.101.5.2205.
- Schiavina M, Moreno-Monroy A, Maffenini L, et al. (2019) GHSL-OECD Functional Urban Areas. *European Commission, Joint Research Centre (JRC) Technical Report - Public Release of GHS-FUA*.
- Veneri P (2016) City size distribution across the OECD: Does the definition of cities matter? *Computers, Environment and Urban Systems* 59. Elsevier: 86–94.

Power law, scaling and the “degrowth geographies”.

The case of Italy

Giampiero LOMBARDINI

¹ University of Genova - Department of Architecture and Design

Stradone Sant'Agostino, 37, 16123, Italy

+39 010 209 51 828, giampiero.lombardini@unige.it

Keywords: Power Law, Urban Scaling, Urban Morphology, Self-similarity

The contribution starts from the consideration that Zipf's law applied to urban systems makes it possible to evaluate both the degree of economic maturity reached by a certain city system and its degree of hierarchization. The spatial distribution of populations and cities explained by power laws has normally been used to represent growth conditions, both in demographic and economic terms. In all those situations where territorial development phenomena have been recorded, the power law applied to population conditions has allowed us to grasp interesting allometric relationships between cities and urban systems.

Three of the issues that this theory of the city poses to scholars are taken into consideration: the maintenance of allometric relationships over time, the role played by regional and urban morphology and finally the self-similarity between urban places. The reflections carried out take Italy as a case study.

The first theme addressed by our research concerns the relationship that seems to exist between economic growth and the structure of the urban system. The empirical data would seem to confirm, in the phases of strong economic and social growth, a tendency to maintain over time the Pareto distribution with normally higher exponent values in correspondence with the conditions of greater development. The question that arises is: what happens when we are faced with situations of stasis or even degrowth such as those that many European regions are experiencing? In the Italian case (a country that has entered a phase of demographic stagnation), the paper analyzed whether the same allometric relations that are recorded with a certain constant regularity in growth periods are also maintained in the stabilization phases of urban systems.

It has already been shown that while a certain general constancy is maintained in Zipf's law in the history of urban systems, below this regularity, great changes are actually observed, with significant variations in the rank and economic importance of individual cities. The development paths of a city

or a particular sub-set of cities within a country can diverge from the more general path of economic growth and in some cases present dramatic collapses in terms not only of population but also of economic performance. This fact would indicate that a general economic development, albeit a positive one, can proceed in a spatially unequal way: population and activities move in the geographical space, creating even considerable territorial gaps. The attraction exerted by agglomeration economies often tends to create cumulative cyclical processes to the exclusive advantage of only a group of cities or a type of city, but to the disadvantage of others.

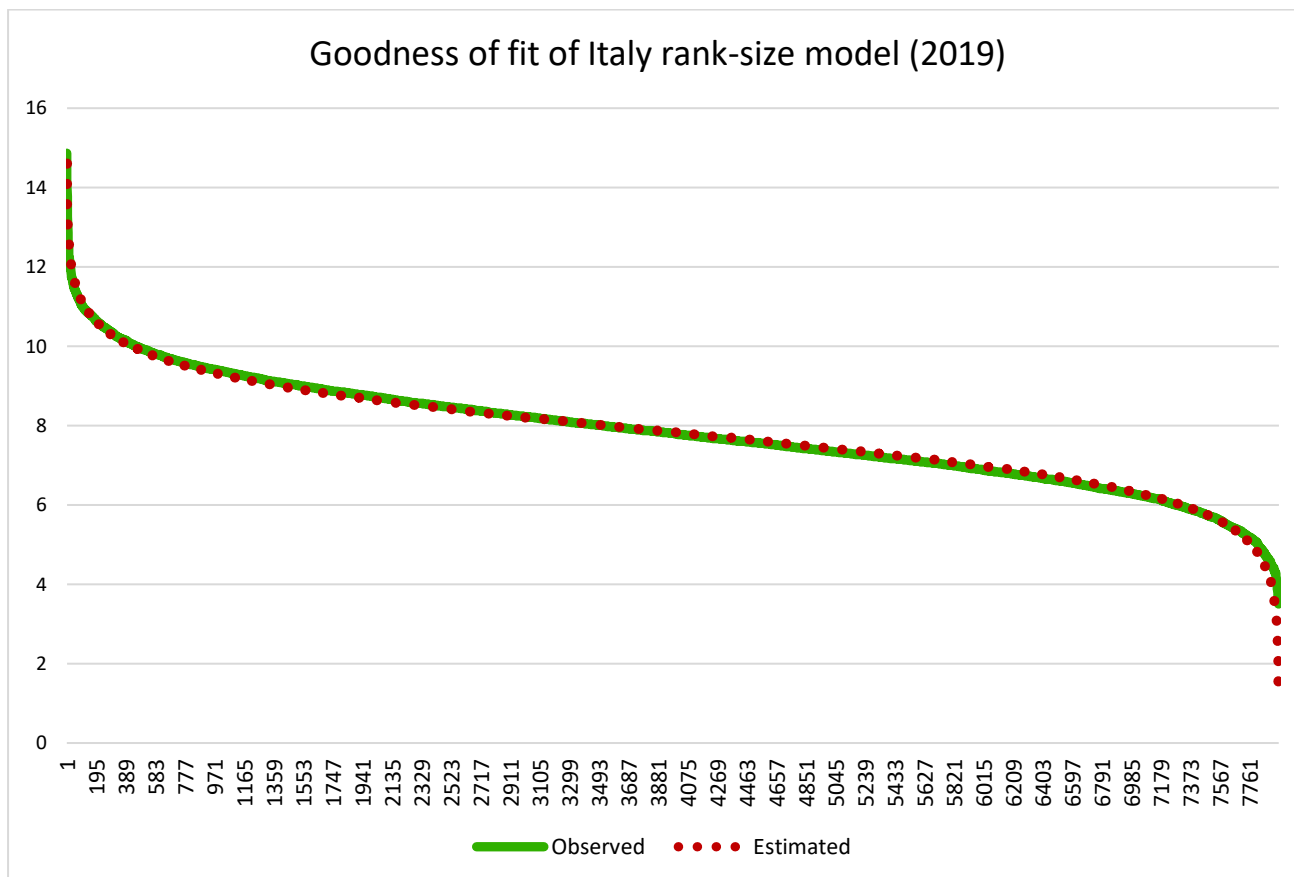


Figure 1: To estimate the population of Italy in function of rank we used a logarithmic functional form for the population and a functional form which is the inverse of a logistic function for the rank.

All this introduces the question of what happens in the historical phases in which a country finds itself in a phase of stagnation or even decrease (or at least of demographic transition). The change can take place according to two main paths: in the first case we can witness homogeneous and equally distributed shrinkage processes in space and which tend, as such, to preserve the original allometric relations. In this case, the main and most dynamic cities tend to keep their rank, even in the presence of a slowdown in growth. However, there is an important problem with regard to small towns: below a certain threshold, they can collapse, as they tend to lose the ability to provide essential services. This

fact generates a further boost to the polarization of the urban system, modifying to some extent the general distribution itself (which will tend to show, in its final stretch, the characters of the log-normal function).

Under these conditions, the problem of how to define a city and how to determine its boundaries also arises. The phenomena of urban fusion and coalescence, in fact, are increasingly frequent in situations of degrowth, so that some small towns, below a certain threshold, tend either to disappear (at least statistically) or to merge into larger urban systems. As an alternative to this homogeneous degrowth process, the case may instead occur in which the dynamics follow selective paths where the main cities tend to concentrate more and more high-ranking functions. It is a top-down change that can also coexist with the previous one. Also in this case there is a change in the law of distribution itself and therefore the centrality relationships within the urban system change.

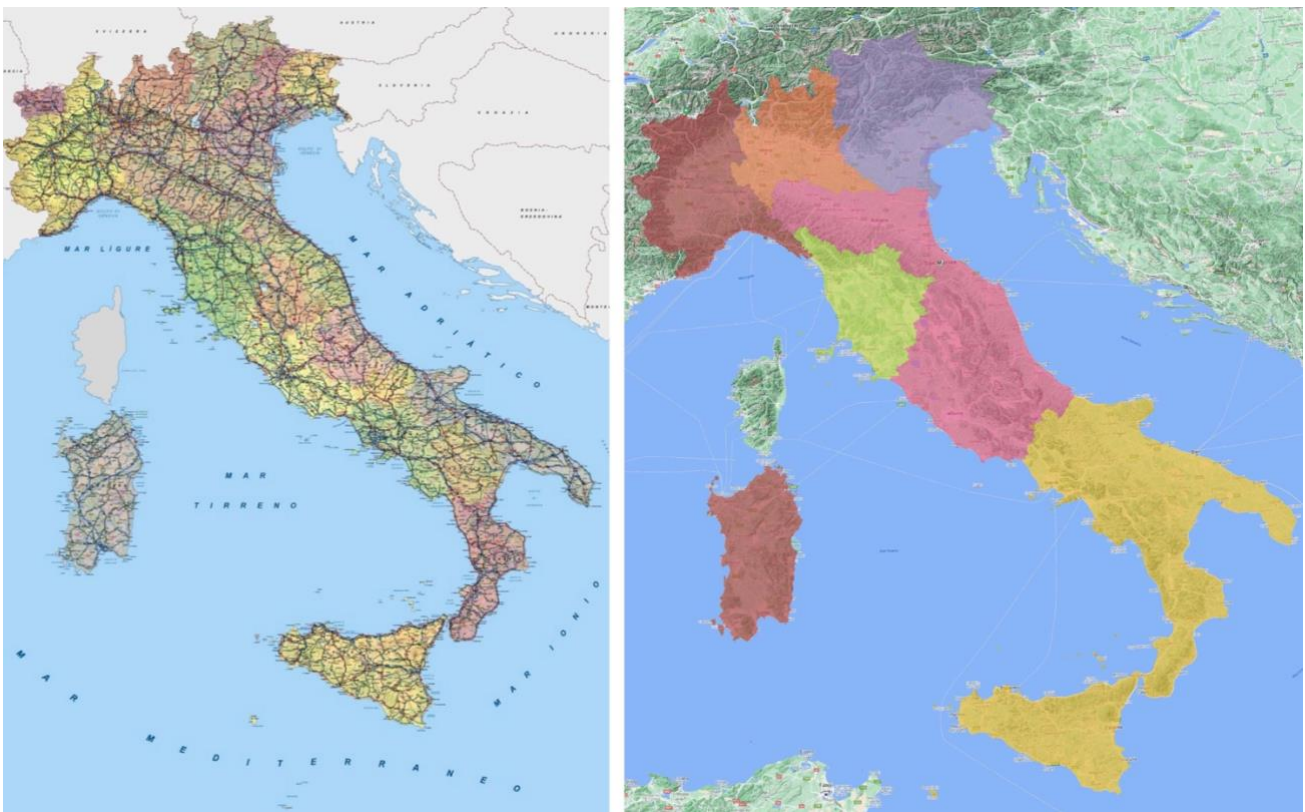
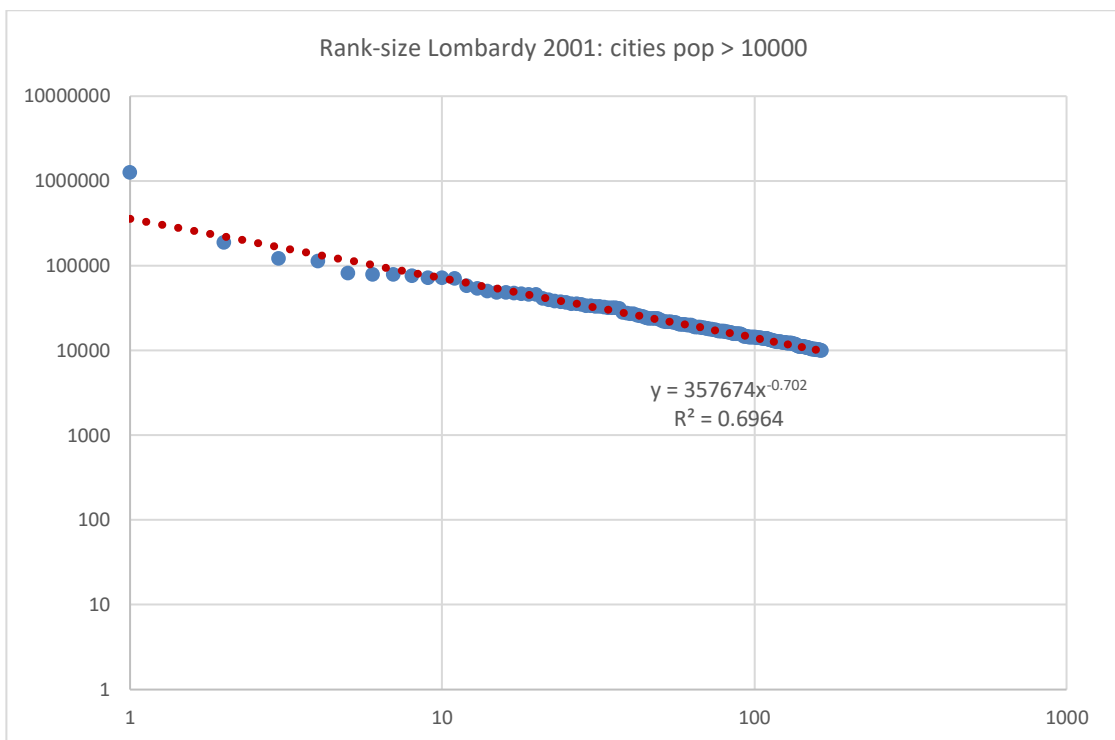
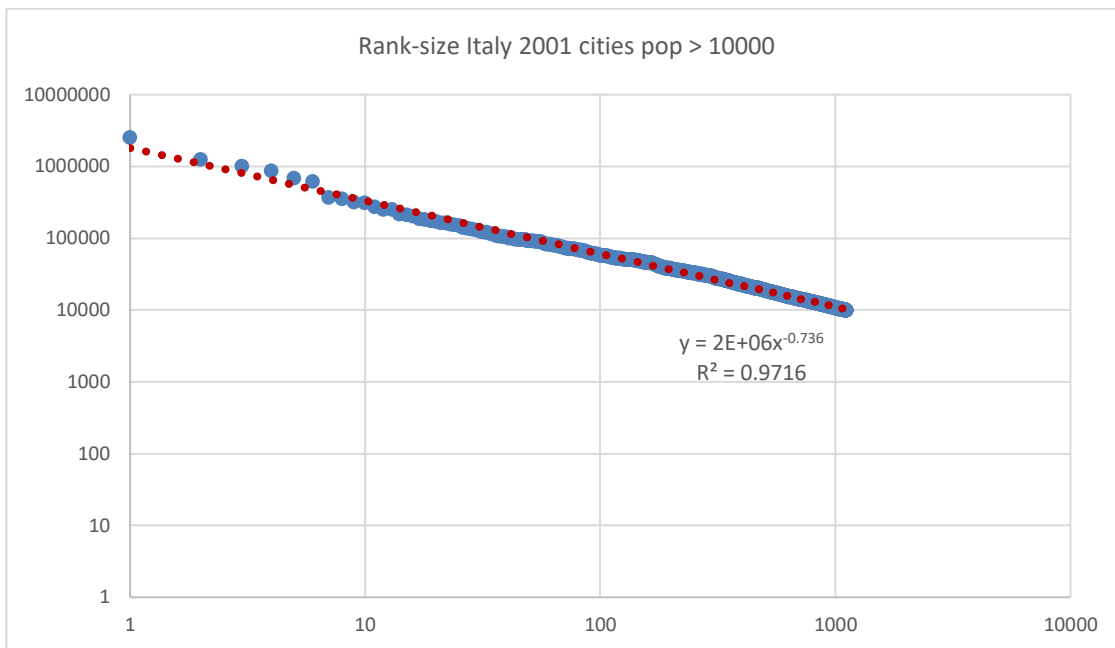


Figure 2: The Italian sub-regions based on historical classification

This introduces the theme of the role played by urban history as well as geography. While on the one hand it has been shown that the rate of development of an urban system does not depend on the initial size of the different centers, on the other hand we must consider that different territorial histories coexist within the same urban system. In our study on Italy we observe how Zipf's law assumes different exponent values depending on whether we consider either the entire nation or if we take into consideration sub-regions that have historical significance. Italy is a newly formed state (1861) and is

a country of late industrialization and urbanization. The fact that the present state is the result of the sum of some ancient previous states is significant in explaining the development trajectories of the different urban sub-systems.



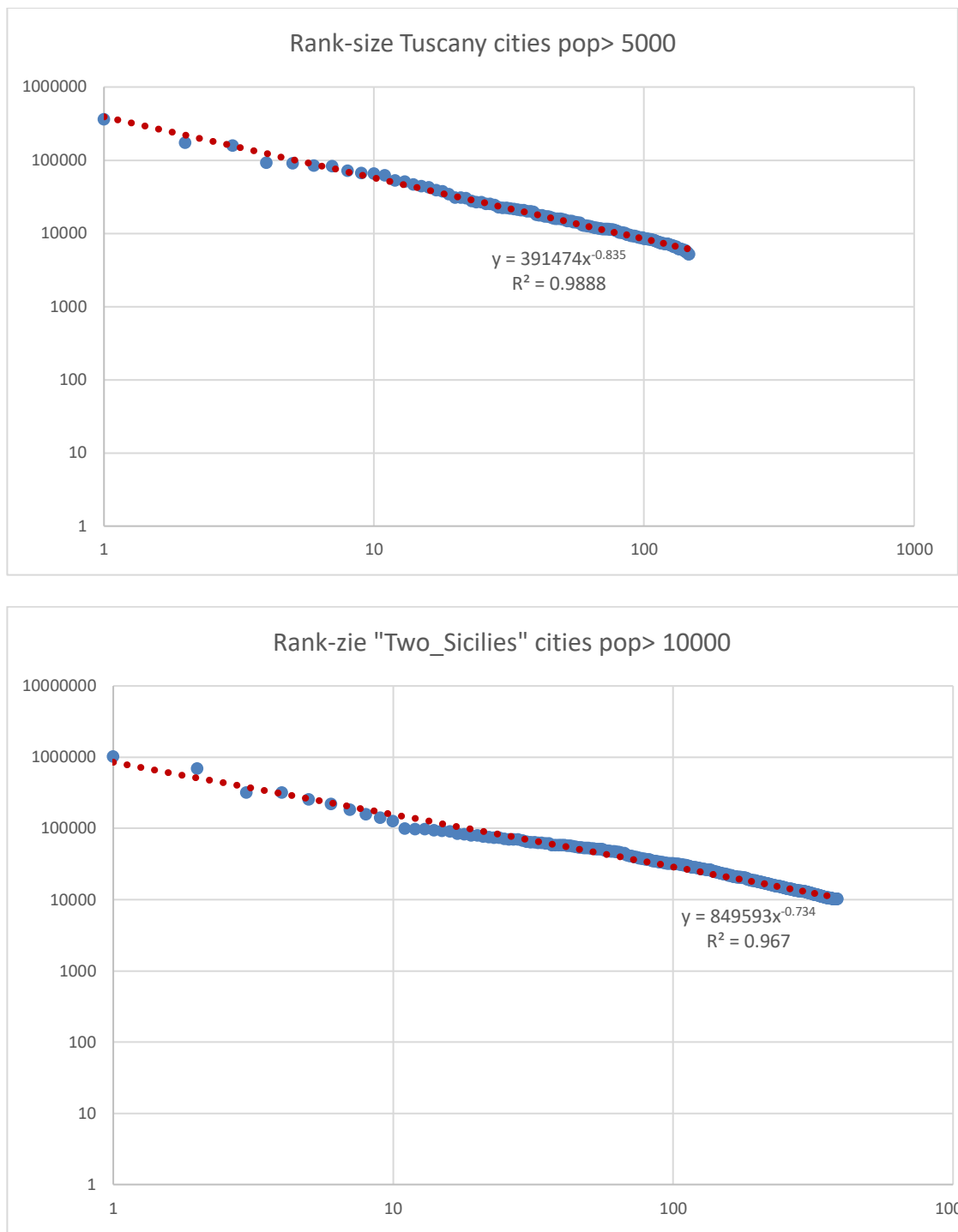


Figure 3: Some rank size distributions for Italy e bub-regions (year: 2019)

In our research, we tried to break down Italy into six large macro-regions corresponding to the historical states of long historical permanence that reached the 1800s. These macro regions all had one or even two autonomous capitals with respect to the rest of the peninsula, their own organization specific administrative, different economic conditions, a different role in the Mediterranean. The post unitarian historical series analysis of the rank-size distribution in these sub-regions shows how they differ significantly from the distribution observed globally over the whole of Italy.

A further consideration then completes the picture of this first analysis of the Italian urban system: self-similarity that seems to emerge as a characteristic factor of the urban phenomenon. Regarding the first of these questions, we started from the demonstrations widely present in the literature, about the fact that cities tend to be all reproductions to scale of each other. This regularity manifests itself in the indices that I assume, for example, two decisive elements: the endowment of infrastructures and urban services, which tend to have scale ratios of less than 1, and the ability of a city to produce or promote highly creative activities. added value (a fact that can be grasped starting, for example, from the production of patents). In this case, a growing city is a city that tends to have this exponent constantly higher than 1. The question asked in the research, also in this case, is what happens when a city begins to decrease. The fact of losing population will lead to a tendential increase in the scale index that links the urban dimension to the infrastructure. This means that the inhabitants of that city will have, at least temporarily, an infrastructural endowment that exceeds their needs. This fact, if it persists over time, can lead to two conditions of change: either a cumulative decrease (which occurs when there are no longer sufficient resources for a city to be able to keep its infrastructural network fully active), or, on the contrary, the higher available asset value can exert an attraction effect precisely for those economic activities that appear to be in crisis (with the index that will tend to drop towards 1). The economic and demographic crisis can therefore lead to a potential bifurcation that can lead the city either towards decline (a city of consumption) or towards economic recovery (the city producing high value goods).

References

- Ausloos M., Cerqueti R. 2016. “A universal rank-size law.” *PloS one* 11.11 (2016): e0166011.
- Batty, M. 2006. “Rank clocks.” *Nature* 444.7119 (2006): 592-596.
- Bettencourt L., Lobo J., Youn H.2013. "The hypothesis of urban scaling: formalization, implications and challenges". *arXiv:1301.5919*
- Bettencourt L., 2021. *Introduction to Urban Science: Evidence and Theory of Cities As Complex Systems*. MIT Press.
- Chen, Y. 2016. “The evolution of Zipf’s law indicative of city development.” *Physica A: Statistical Mechanics and its Applications* 443: 555-567
- Puente-Ajovín M., Ramos A., 2015. “On the parametric description of the French, German, Italian and Spanish city size distributions”. *The Annals of Regional Science* 54.2: 489-509.
- Gabaix X. 1999, “Zipf’s law for cities: an explanation.” *The Quarterly Journal of Economics* 114.3: 739-767.
- González-Val, R., et al. 2015. “Size distributions for all cities: Which one is best?” *Papers in Regional*

Science 94.1: 177-196.

Gabaix, X., and Ioannides Y.M. 2004. “The evolution of city size distributions.” *Handbook of regional and urban economics*. Vol. 4. Elsevier, 2004. 2341-2378.

Special Session SS03_04.1

4 November 9:00-11:00 am (GMT)

Exploration and validation of spatial simulation models

Proponents

Juste Raimbault, Denise Pumain, Eric Koomen and Chris Jacobs-Crisioni

Forecasting Residential Sprawl Under Uncertainty: An Info-Gap Analysis

Dani Broitman¹; Yakov Ben-Haim²

¹Faculty of Architecture and Town Planning

Technion – Israel Institute of Technology, Haifa 3200003, Israel

+972-77-887-5555, danib@technion.ac.il

²Faculty of Mechanical Engineering

Technion – Israel Institute of Technology, Haifa 3200003, Israel

+972-77-887-5555, yakov@technion.ac.il

Keywords: scenario-based forecasting; residential sprawl; Netherlands; uncertainty; info-gap

1. Introduction

Spatial planners attempt to set a clear vision about how a certain area should develop, by addressing environmental, societal and economic issues, and then to define the delivery mechanisms necessary to implement the vision (Taylor 2010). In other words, planners define objectives related to the spatial ordering of a region, and invoke economic and regulative instruments in order to realize the envisioned objectives (Hersperger et al. 2018). If future needs, constraints and political pressures were known, and the defined objectives were feasible, then spatial planners would invariably fulfil their role successfully. But the future is uncertain, and many factors involved in the spatial planning process are unknown in advance. These uncertainties include local parameters (population growth, the type and extent of local economic activity, etc.), national considerations (budget allocations, developmental priorities, etc.), and global factors (climate change, global economic cycles, political upheavals, etc.).

Scenario-based forecasting is a common way to deal with this fundamental uncertainty, since it is well suited for analyses that deal with a wide array of future possible developments. This can be done from two perspectives, one prospective and the other retrospective (Koomen, Rietveld, and de Nijs 2008). The prospective approach involves the definition of a range of plausible future scenarios, based on different values assigned to the most relevant variables assumed to influence future outcomes. These scenarios are alternative views of the future that constrain the planned system. The retrospective approach (also called ex-post analysis) entails trend analysis, based on observed spatial developments, results of policies implemented in the past, and theoretical insights.

Scenario-based forecasting is intended to answer “what-if” questions: What may happen with the issues of interest if particular conditions become a reality in the future? Because of its characteristics, this approach has been used for a wide range of planning issues, since the main objective of scenario-

based forecasting is to offer guidance to decision makers regarding problems that are likely to appear in the future, and possible ways to deal with them in advance (Koomen, Hilferink, and Borsboom-van Beurden 2011).

Most scenario-based forecasting is prospective, envisioning alternative futures related to the analyzed combinations of scenarios and models, while giving little attention to the past. Retrospective analysis of scenario-based forecasting is, in comparison, rare (Oliveira and Pinho 2010). Nevertheless, this perspective can be of great value when new scenarios are imagined, and possible new policies are considered, because retrospective analysis, focused on past policies' degrees of success, may identify the effectiveness of each one (Koomen, Dekkers, and Broitman 2018). Despite the relative scarcity of retrospective analyses in spatial forecasting, there are examples of influential studies performed using this perspective (Alfasi, Almagor, and Benenson 2012; Alterman and Hill 1978; Faludi 2000; Xie, Hou, and Herold 2018).

The main objective of retrospective valuation is to compare the original objective of the spatial forecast, with the observed outcomes, trying to explain the differences between them (Oliveira and Pinho 2010). This allows an observer in the present, with full knowledge about what happened during the whole time-period, to look back at the past and assess the outcome compared with the original forecast. This retrospective knowledge was not available to the original analyst, who had to make do with the models, methods, and reality at the time, and with the uncertain assessments about possible futures. In this sense, retrospective valuation is unfair to the original forecasters, since this approach ignores the uncertainty in which they were mired. But since the future has inherent uncertainties, and the past is well known (assuming data availability), a combination of both approaches could be both informative and productive. Looking backwards it is possible not only to measure performance, but also to evaluate how analysts in the past dealt with uncertainty while imagining possible futures from their point in time. These insights can help to identify potential pitfalls in prospective scenario-based forecasting, highlighting topics that require improvement, not only in the policy field, but also regarding the future scenarios. This paper develops a methodology to combine retrospective analyses focused on past performance with prospective scenario-based forecasting.

Using info-gap decision theory (Ben-Haim 2006, 2010, 2018) as the theoretical and methodological workhorse, and residential sprawl in The Netherlands as a test case, we develop an approach that is summarized in Figure 1. Stage 1 entails retrospective valuation (comparison between the forecast and the observed outcome). Stage 2 entails retrospective assessment of uncertainty and robustness defined by the forecaster in the past. Stage 3 proceeds to prospective scenario-based forecasting, learning from previous experiences regarding the performance and management of uncertainty and robustness.

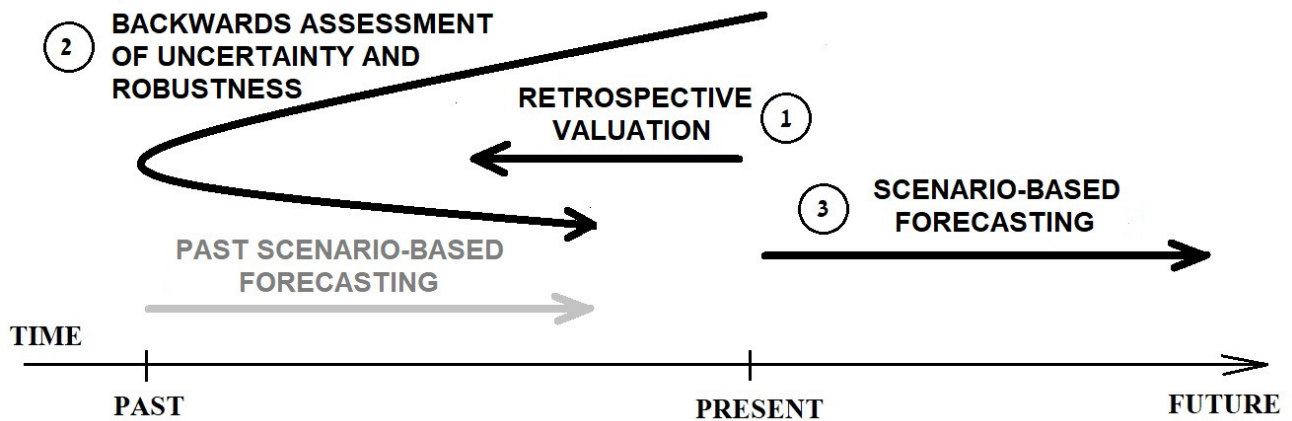


Figure 1: The suggested methodology, assuming that past scenario-based forecasts are available (in grey).

This paper is the first to apply info-gap decision theory to the forecasting outcomes of spatial planning. The main approach is to assess robustness to uncertainty as a function of the required performance. In addition to the robustness function, the info-gap opportuneness function provides a tool for assessing the potential for better-than-anticipated outcomes.

2. General description of the results

The combination of retrospective and prospective analyses allows one to address otherwise inaccessible questions. For example: Did forecasting analysts in the past realistically account for uncertainty? Do forecasting analysts today have a reasonable approach to managing future uncertainty, and how should that uncertainty-management be done? Does the uncertainty-management address both the pernicious as well as the propitious aspects of uncertainty? What can be learned from the analysis of robustness to uncertainty, and opportuneness from uncertainty, in order to improve scenario-based forecasting?

This approach requires data about past forecasting efforts, the actual situation, and plans for the future. Extensive spatial and planning data, in all three categories, are available for Dutch residential sprawl. This data availability allows us to focus on specific regions and compare between them, exploring assumptions, aspirations and points of view. Therefore, the methodology developed in this paper aims to identify potential points of failure in scenario-based spatial forecasting, based on past experiences, the observed status, and the robustness to, and opportuneness from, uncertain possible futures. Shedding light on these issues enables further and focused inquiry. Questions raised by this analysis have the potential to improve scenario-based forecasting, enhancing its response to deep uncertainty about the unknown future.

A detailed description of the info-gap analysis applied to the study of residential sprawl and the obtained results is beyond the scope of an extended abstract. A full paper was submitted for publication to a leading urban science journal, and this extended abstract is based on excerpts of the full version, that will be presented in the conference.

References

- Alfasi, Nurit, Jonatan Almagor, and Itzhak Benenson. 2012. 'The Actual Impact of Comprehensive Land-Use Plans: Insights from High Resolution Observations'. *Land Use Policy* 29(4):862–77. doi: 10.1016/j.landusepol.2012.01.003.
- Alterman, Rachele, and Morris Hill. 1978. 'Implementation of Urban Land Use Plans'. *Journal of the American Institute of Planners* 44(3):274–85. doi: 10.1080/01944367808976905.
- Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd edition, Academic Press, London.
- Ben-Haim, Yakov, 2010, *Info-Gap Economics: An Operational Introduction*, Palgrave-Macmillan.
- Ben-Haim, Yakov, 2018, *Dilemmas of Wonderland: Decisions in the Age of Innovation*, Oxford University Press.
- Faludi, Andreas. 2000. 'The Performance of Spatial Planning'. 2000 15(4):299–318.
- Hersperger, Anna M., Eduardo Oliveira, Sofia Pagliarin, Gaëtan Palka, Peter Verburg, Janine Bolliger, and Simona Grădinaru. 2018. 'Urban Land-Use Change: The Role of Strategic Spatial Planning'. *Global Environmental Change* 51:32–42. doi: 10.1016/j.gloenvcha.2018.05.001.
- Koomen, Eric, Jasper Dekkers, and Dani Broitman. 2018. 'Analyzing and Simulating Urban Density Exploring the Difference Between Policy Ambitions and Actual Trends in the Netherlands'. Pp. 145–71 in *Spatial Analysis and Location Modeling in Urban and Regional Systems, Advances in Geographic Information Science*, edited by J.-C. Thill. Berlin, Heidelberg: Springer. Koomen, Eric, Jasper Dekkers, and Terry van Dijk. 2008.
- Koomen, Eric, Maarten Hilferink, and Judith Borsboom-van Beurden. 2011. 'Introducing Land Use Scanner'. Pp. 3–21 in *Land-Use Modelling in Planning Practice*. Vol. 101, *GeoJournal Library*, edited by E. Koomen and J. Borsboom-van Beurden. Dordrecht: Springer Netherlands.
- Koomen, Eric, Piet Rietveld, and Ton de Nijs. 2008. *Modelling Land-Use Change for Spatial Planning Support*. Springer.
- Oliveira, Vitor, and Paulo Pinho. 2010. 'Evaluation in Urban Planning: Advances and Prospects'. *Journal of Planning Literature* 24(4):343–61. doi: 10.1177/0885412210364589.

- Taylor, Nigel. 2010. 'Commentary: What Is This Thing Called Spatial Planning? An Analysis of the British Government's View'. *TPR: Town Planning Review* 81(2):193–208. doi: 10.3828/tpr.2009.26.
- Walker, Warren E., Marjolijn Haasnoot, and Jan H. Kwakkel. 2013. 'Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty'. *Sustainability* 5(3):955–79. doi: 10.3390/su5030955.
- Xie, Xiaoping, Wei Hou, and Hendrik Herold. 2018. 'Ex Post Impact Assessment of Master Plans—The Case of Shenzhen in Shaping a Polycentric Urban Structure'. *ISPRS International Journal of Geo-Information* 7(7):252. doi: 10.3390/ijgi7070252.

How validation through model exploration empowers theories of spatial complexity: example of urban systems

Juste RAIMBAULT¹; Denise PUMAIN²

¹ Research Fellow, Center for Advanced Spatial Analysis, University College London, Gower Street, London WC1E 6BT, United Kingdom

j.raimbault@ucl.ac.uk

² Emeritus Professor, Université Paris I Panthéon Sorbonne, UMR Géographie-cités, Campus Condorcet, 5, cours des Humanités, 93322 Aubervilliers cedex. France

pumain@parisgeo.cnrs.fr

Keywords : spatial interaction ; urban system ; simulation; model exploration; validation

Submitted to Special Session SS03

Strong inequalities in the size of cities and the apparent difficulty of limiting their growth raise practical issues for spatial planning. At a time when new constraints in terms of limited energy and raw material resources or possible catastrophic events such as pandemics are challenging further urban expansion, it is important to consolidate the theories from various scientific disciplines to estimate to what extent the urban dynamics can be modified.

1 About urban theories and models

While reviewing shortly the contributions to urban theories provided by the new developments in complexity sciences (Batty, 2013), we first advocate for the soundness of urban theories. Although important advances may arise with new insights stemming from recently entering disciplines in the field, we remind how crucial it is of not neglecting the accumulated knowledge that was provided through previous classical methods of investigation (Lobo et al. 2020). Especially, we value concern for a correct adjustment between the selection and definition of urban objects and the conclusions than can be derived from processing the corresponding data (Egidi et al. 2020), we also remind the usefulness of avoiding cultural bias through a careful contextualization of investigations in space (according to scale as well as regions of the world and spatial interaction regimes) and time (in terms of length of considered process and specific historical period) (Pumain et al. 2015).

Second, we develop our original approach considering spatial interaction and evolutionary path dependence as major features in the general behavior of urban entities (Raimbault, Pumain 2019; Pumain, Raimbault 2020). These principles are grounded in an evolutionary theory of urban systems. We explain how our model's ontology was established from comparative research on urban systems evolution and an epistemological questioning related to a geographical frame, i.e. the quest for explanation of urban spatial diversity (Pumain, 2020).

2 New methods for model exploration

Third, and this is the major part of our presentation, we explain how models based on these theoretical premises can help validating and developing urban theories. Four dynamic models of urban growth are experimented. They are complementary, in line with a multi-modelling approach (Cottineau et al 2015). They rely on the same fundamental assumptions and share a common basic structure and formulation: (i) agents are cities, characterized by a main state variable of population; (ii) building on the Gibrat model, they add additional processes to endogenous growth to account for space and spatial interactions; (iii) they simulate not stochastic distributions of city sizes but their averages in time, and are thus specified on these averages only. The interaction between cities, which could be included in the covariance structure for a fully stochastic model, is thus captured by the spatial interaction terms. The first model develops an urban hierarchy with a transportation network (Raimbault 2020); the second and third explain urban development using economic and environmental processes (Raimbault et al. 2020). A fourth model links urban growth with innovation and spatial diffusion processes. The models were benchmarked and applied to several systems of cities worldwide using harmonized empirical data sets in a previous paper by Raimbault, Denis and Pumain (2020). We insist here on the use of the OpenMOLE simulation platform as a tool for exploring the behaviour of models and assessing their relative relevance for reconstructing plausible urban development patterns. Such exploratory experiment do improve to a large extent the usual procedure for hypothesis validation. We illustrate this with the use of genetic algorithms for model calibration, diversity search algorithm applied to model output diversity, and spatial sensitivity analysis. The OpenMOLE workflow system can furthermore be used as a tool to seamlessly integrate and couple models.

3 Two questions brought for discussion

We finally develop two crucial open questions regarding model exploration and validation. First, combining specific technical tools such as automatic code generation and methodological tools with a dedicated niched genetic algorithm to calibrate a large set of these models, Cottineau et al. (2015) were able to produce a population of optimal solutions and statistically evaluate the contribution of diverse mechanisms to model fit, yielding a benchmark of mechanisms in terms of explanatory power. However, this approach is possible only when all models can be integrated into a single framework and directly compared. In the case of Raimbault, Denis and Pumain (2020), such automatic generation and niched calibration could not be achieved since all models do not share all the same core processes, and models were independently calibrated on two objectives and the final Pareto fronts compared. This was still possible, as the models have the same input and output data and a similar structure. How can one couple and integrate models when they are highly heterogeneous, at different scales and on totally different objects? How can one build multi-models of heterogeneous models, ideally automatically towards the idea of ‘model crushing’ introduced by Openshaw (1993)? These questions remain rather open, although initiatives developed within the context of the OpenMOLE platform are built towards the facilitation of tackling such issues (Raimbault and Pumain, 2020).

Secondly, what are the general and robust stylized facts that can be attributed to systems of cities in general, and what are the peculiar geohistorical conditions that lead to a particular trajectory? Beyond the myth of ‘universality’ advocated by physicists, which would imply universal laws derived from the same microscopic processes and which would assume some kind of ergodicity in the system (Pumain, 2012), robust stylized facts such as Zipf’s and

Gibrat's laws are well identified in a first approximation. In that context and within the interaction between models and theories, being able to distinguish between intrinsic model dynamics, robust effects due to space that are observed on similar geographical structures, and effects due to the peculiar initial conditions and geohistorical conditions is a difficult problem. Raimbault, Raimbault et al. (2019) introduced a novel methodological pipeline for sensitivity analysis of spatial models, which makes it possible to test the sensitivity of model outcomes to initial spatial conditions. This requires generating synthetic spatial data that resemble real configurations, in this case population grids at the scale of urban areas. Diverse generators at other scales (microscopic building configurations, macroscopic systems of cities, transportation networks) are developed by Raimbault, Perret and Reuillon (2020). Such work is the basis for a better understanding of how urban dynamics models behave on synthetic systems of cities compared to real case studies. This is essential for the construction of integrated theories in order to understand their level of generality and how they relate to specific geographical situations.

Acknowledgments

This presentation is partly adapted from a book chapter: Raimbault J. Pumain D. (2021), Spatial dynamics of complex urban systems within an evolutionary theory frame, in A. Reggiani, L.A. Schintler, D. Czamanski, R. Patuelli (eds) *Handbook on Entropy, Complexity and Spatial Dynamics. The Rebirth of Theory?* Edward Elgar Publishing Ltd, Cheltenham, UK

References

- Batty, M. (2013). *The new science of cities*. MIT Press, Cambridge, MA, USA.
- Cottineau, C., Chapron, P. & Reuillon, R. (2015). Growing models from the bottom up: An evaluation-based incremental modelling method (EBIMM) applied to the simulation of systems of cities. *Journal of Artificial Societies and Social Simulation*, 18(4), 9.
- Egidi, G., Salvati, L. & Vinci, S. (2020). The long way to Tipperary: City size and worldwide urban population trends, 1950–2030. *Sustainable Cities and Society*, 102148.
- Lobo, J., Alberti, M., Allen-Dumas, M., Arcaute, E., Barthelemy, M., Bojorquez Tapia, L. A., ... & Florida, R. (2020). *Urban science: Integrated theory from the first cities to sustainable metropolises*.
- Pumain, D. (2012). Urban systems dynamics, urban growth and scaling laws: The question of ergodicity. In *Complexity theories of cities have come of age*, Springer, Berlin/Heidelberg, DE. (pp. 91–103).
- Pumain D., Swerts E., Cottineau C. Vacchiani-Marcuzzo C., Ignazzi A., Bretagnolle A., Delisle F., Cura R., Lizzi L & Baffi S. (2015) : Multi-level comparison of large urban systems. *Cybergeo*, 706, <http://cybergeo.revues.org/26730> ; DOI : 10.4000/cybergeo.26730
- Pumain, D., & Raimbault, J. (2020). Conclusion: Perspectives on urban theories, in Pumain D. (ed) *Theories and models of urbanization* (303–30). Springer, Lecture Notes in Morphogenesis..Cham, CH, pp 303-330.
- Raimbault, J. (2021). Modeling the co-evolution of cities and networks. In Rozenblat, C. & Niel, Z. (eds), *Handbook of Cities and Networks*. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.
- Raimbault, J., & Pumain, D. (2019). Methods for exploring simulation models. *Geographical Modeling: Cities and Territories*, 2, 125–50.
- Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F. & Reuillon, R. (2019). Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models. *Journal of Artificial Societies and Social Simulation*, 22(4).
- Raimbault, J., Denis, E. & Pumain, D. (2020) Empowering urban governance through urban science: Multi-scale dynamics of urban systems worldwide. *Sustainability*, 12, 5954.
- Raimbault, J., & Le Néchet, F. (2021). Introducing endogenous transport provision in a LUTI model to explore polycentric governance systems. *Journal of Transport Geography*, 94, 103115.
- Reuillon, R., Leclaire, M., & Rey-Coyrehourcq, S. (2013). OpenMOLE, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, 29(8), 1981-1990.

Bayesian calibration of cellular automata urban growth models from urban genesis onwards

Jingyan YU¹; Alex HAGEN-ZANKER¹; Naratip SANTITISSADEEKORN²; Susan HUGHES¹

¹Department of Civil and Environmental Engineering, University of Surrey, Guildford, Surrey, GU2 7XH, UK
+44 (0)07767940299, jingyan.yu@surrey.ac.uk (correspondent author)

²Department of Mathematics, University of Surrey, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Keywords: Land-use; Urban growth; Cellular Automata; Calibration; Approximate Bayesian Computation

1. Introduction

Cellular Automata (CA) land use change models are widely used to study and understand complex spatiotemporal processes from which urban form and growth emerge (White et al., 2015). The quality of CA-supported land use decisions depends on the efficacy of model calibration (Straatman et al., 2004), which is a long standing challenge. CA calibration is often based on relatively short calibration periods of available data and small geographical areas, providing insufficient land use changes to identify the processes that drive the formation and transition of urban form and problematic signal-to-noise ratios of real changes versus apparent changes (Grinblat et al., 2016, van Vliet et al., 2016). To address the negative effect of short calibration periods, this study investigates a novel approach which extends the calibration to the point of urban genesis, i.e., a landscape void of urban land, and calibrates the model based on emerging urban growth patterns from a simulation starting from urban genesis. These calibrated parameters based on long term urban form dynamics are then applied to project urban growth forward in time from a known urban configuration.

2. Methodology

This approach to calibrate from urban genesis onwards is embedded in a new modelling framework of a Constrained Cellular Automata (CCA) model of urban growth.

The CCA models the neighbourhood effect of interacting urban spatial agglomeration and preservation of scarce non-urban land at multiple spatial scales with four parameters, minimizing the parameter space and requiring no ancillary layers such as zoning, accessibility, and physical suitability.

The model calibration is fully automated and uses a novel application of Markov Chain Monte Carlo Approximate Bayesian Computation (MCMC-ABC) for Bayesian parameter estimation (Haario et al.,

2001, Andrieu and Thoms, 2008, Roberts and Rosenthal, 2009, Toni et al., 2009, Sisson et al., 2018, Ke and Tian, 2019, van der Kwast et al., 2011, Van der Kwast et al., 2012, Verstegen et al., 2014, Verstegen et al., 2016). Comparing to deterministic optimization methods which seek the best fit parameters to an objective function, this Bayesian method estimates the posterior probability distributions of parameters by using Bayes' rule to update the prior probabilities with the observed land use changes.

The minimalistic approach of starting from a non-urban landscape and excluding ancillary layers provides the model no anchor points to achieve location-to-location agreement. The goodness-of-fit (GOF) measure is therefore a pattern-based landscape morphological comparison, inspired by the literature on fractal urban form (Thomas et al., 2008) to evaluate the statistical difference in the distribution of built density using kernel density estimates of urban land.

This method provides an empirical distribution of parameter values that reflects model uncertainty (Figure 1). The validation uses multiple samples from the estimated parameters to quantify the propagation of model uncertainty to the validation measures (Figure 4).

3. Results and conclusions

The framework is applied to two UK towns (Oxford and Swindon) with two distinctive patterns of growth. Urban growth projections with the cross-application of calibrated parameters show that calibrated parameters effectively capture the different urban growth patterns of both towns. For Oxford (Figure 2), the CCA correctly produces the pattern of constrained and scattered growth in the periphery. For Swindon (Figure 3), the pattern of compact and concentric growth. The ability to identify different modes of growth is not only practical but also theoretically meaningful. Existing land use patterns can be an important indicator of future trajectories. Planners are provided with knowledge of urban form characteristics captured by calibrated parameters, an enriched decision space of alternative urban growth trajectories, and the uncertainty of CA models and parameters.

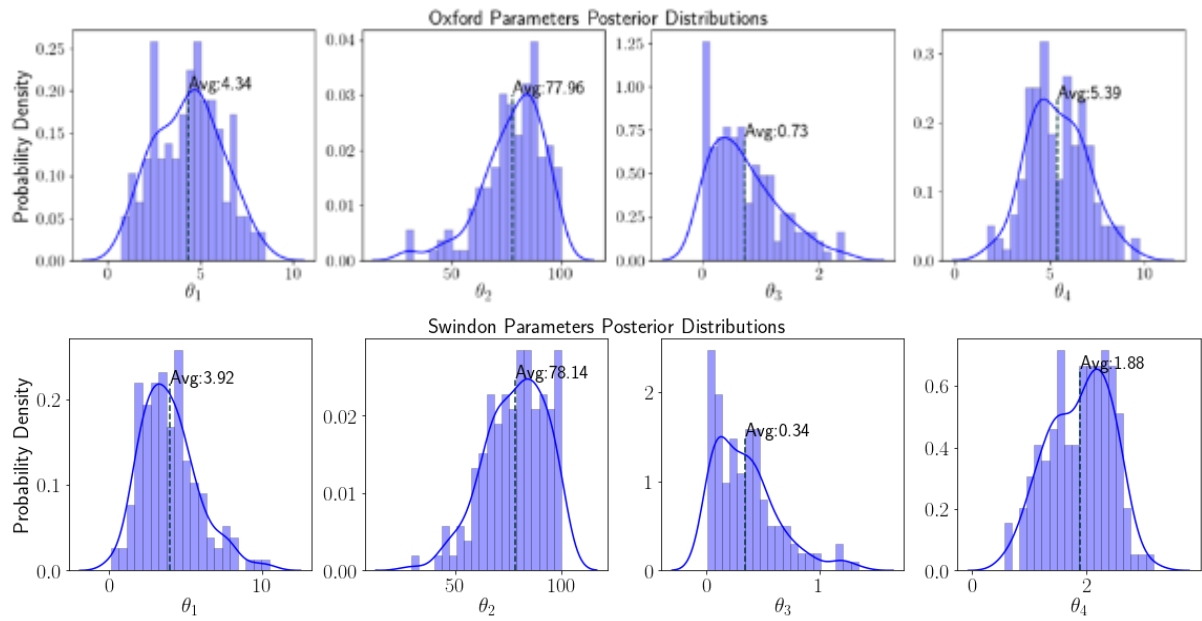


Figure 1: Posterior distributions of calibrated parameters of Oxford and Swindon

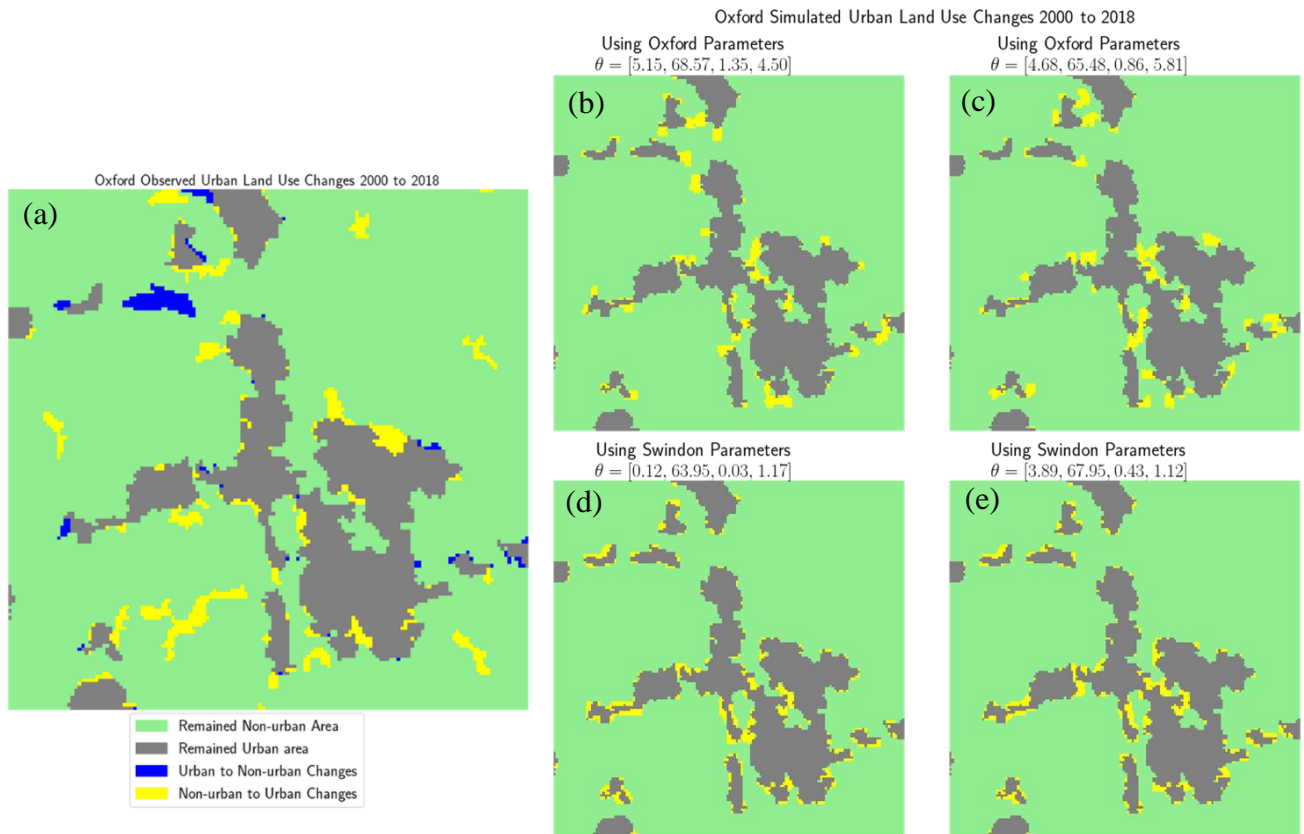


Figure 2: Compare observed and predicted land use changes of Oxford

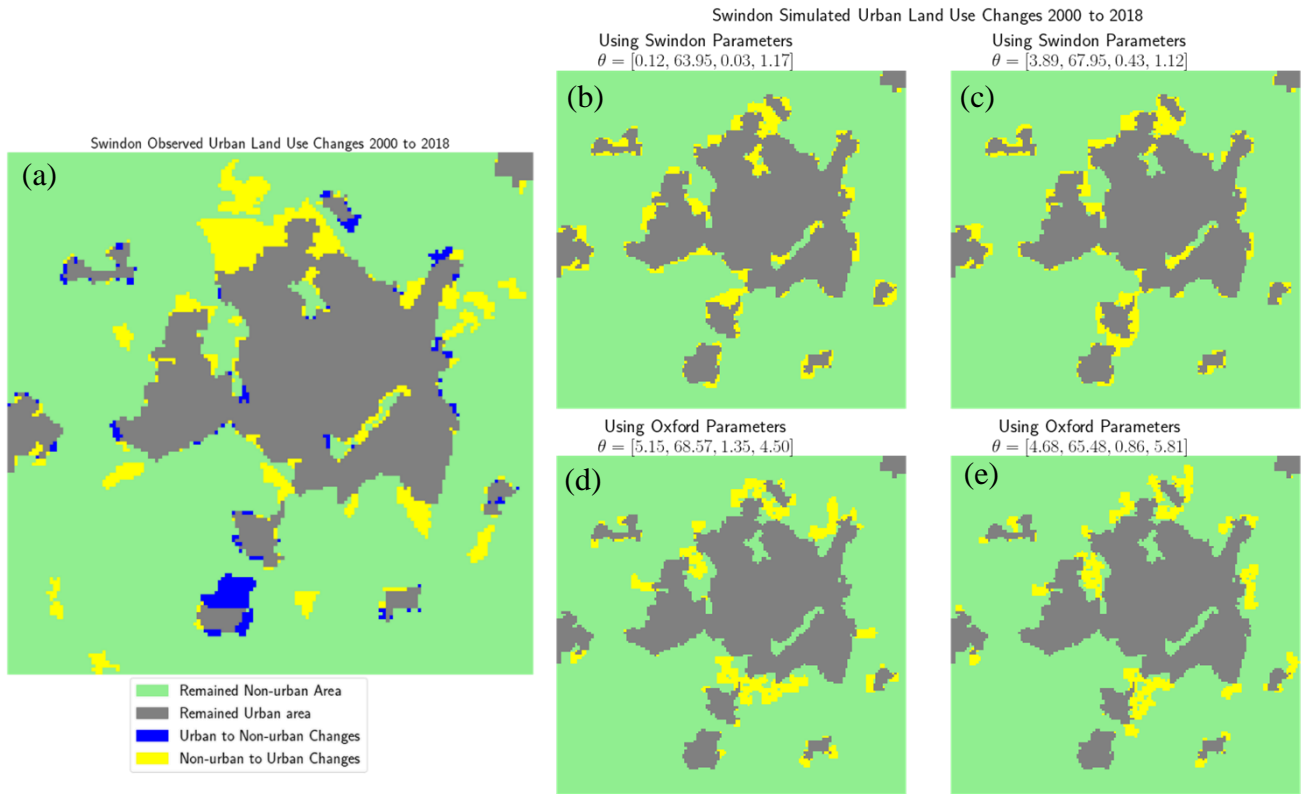


Figure 3: Compare observed and predicted land use changes of Swindon

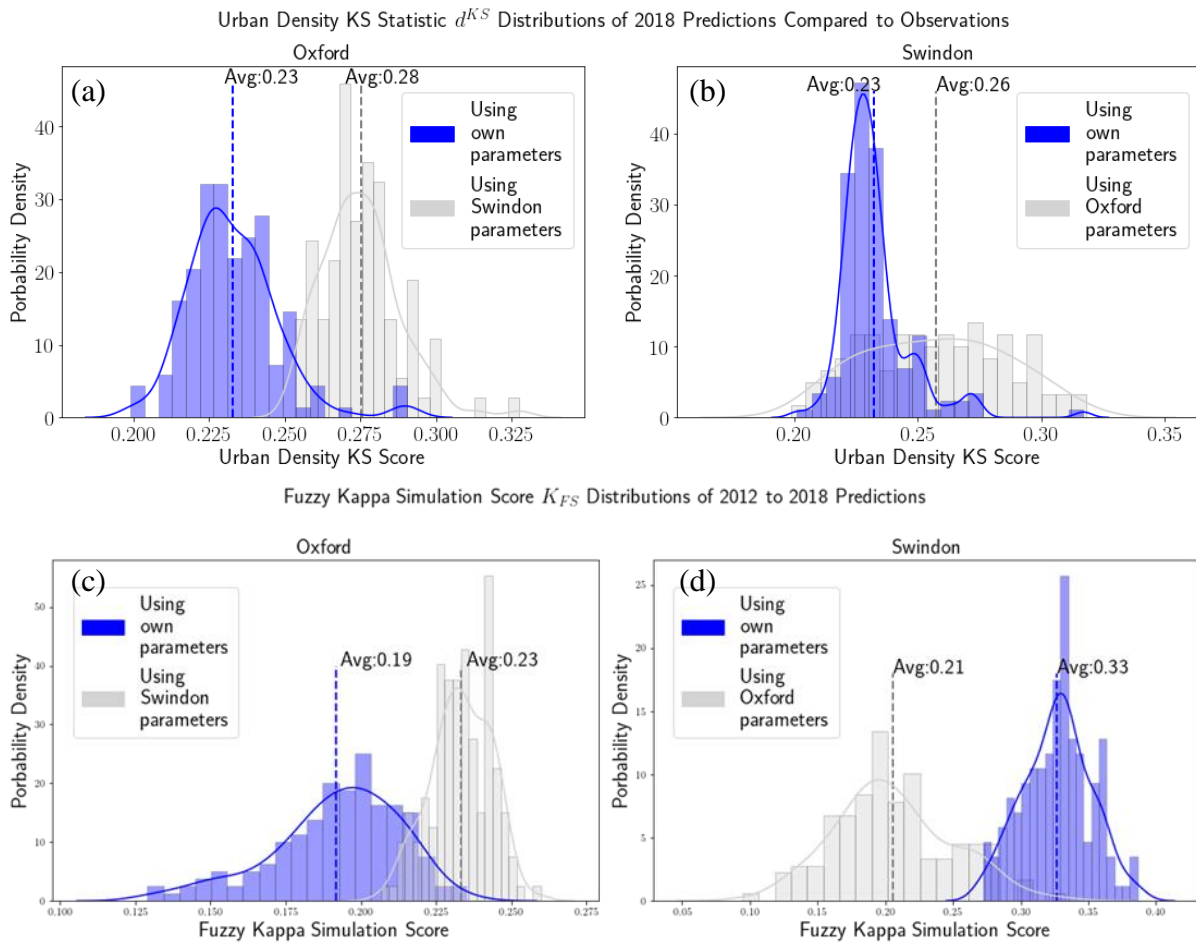


Figure 4: Urban density Kolmogorov–Smirnov statistic and Fuzzy Kappa Simulation score of Oxford and Swindon land use changes validations

Acknowledgments

UKRI, NERC Landscape Decisions programme project number NE/T004150/1.

References

- ANDRIEU, C. & THOMS, J. 2008. A tutorial on adaptive MCMC. *Statistics and computing*, 18, 343-373.
- GRINBLAT, Y., GILICHINSKY, M. & BENENSON, I. 2016. Cellular automata modeling of land-use/land-cover dynamics: questioning the reliability of data sources and classification methods. *Annals of the American Association of Geographers*, 106, 1299-1320.
- HAARIO, H., SAKSMAN, E. & TAMMINEN, J. 2001. An adaptive Metropolis algorithm. *Bernoulli*, 7, 223-242, 20.
- KE, Y. & TIAN, T. 2019. Approximate Bayesian Computational Methods for the Inference of Unknown Parameters. *2017 MATRIX Annals*. Springer.
- ROBERTS, G. O. & ROSENTHAL, J. S. 2009. Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18, 349-367.
- SISSON, S. A., FAN, Y. & BEAUMONT, M. 2018. *Handbook of approximate Bayesian computation*, CRC Press.

- STRAATMAN, B., WHITE, R. & ENGELEN, G. 2004. Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems*, 28, 149-170.
- THOMAS, I., FRANKHAUSER, P. & BIERNACKI, C. 2008. The morphology of built-up landscapes in Wallonia (Belgium): A classification using fractal indices. *Landscape and Urban Planning*, 84, 99-115.
- TONI, T., WELCH, D., STRELKOWA, N., IPSEN, A. & STUMPF, M. P. H. 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society, Interface*, 6, 187-202.
- VAN DER KWAST, J., CANTERS, F., KARSSENBERG, D., ENGELEN, G., VAN DE VOORDE, T., ULJEE, I. & DE JONG, K. 2011. Remote sensing data assimilation in modeling urban dynamics: objectives and methodology. *Procedia Environmental Sciences*, 7, 140-145.
- VAN DER KWAST, J., POELMANS, L., VAN DE VOORDE, T., DE JONG, K., ULJEE, I., KARSSENBERG, D., CANTERS, F. & ENGELEN, G. 2012. Uncertainty analysis and data-assimilation of remote sensing data for the calibration of cellular automata based land-use models.
- VAN VLIET, J., BREGT, A., BROWN, D., VAN DELDEN, H., HECKBERT, S. & VERBURG, P. 2016. A review of current calibration and validation practices in land-change modeling. *Environmental Modelling and Software*, 82, 174-182.
- VERSTEGEN, J. A., KARSSENBERG, D., VAN DER HILST, F. & FAAIJ, A. P. 2014. Identifying a land use change cellular automaton by Bayesian data assimilation. *Environmental Modelling & Software*, 53, 121-136.
- VERSTEGEN, J. A., KARSSENBERG, D., VAN DER HILST, F. & FAAIJ, A. P. 2016. Detecting systemic change in a land use system by Bayesian data assimilation. *Environmental Modelling & Software*, 75, 424-438.
- WHITE, R., ENGELEN, G. & ULJEE, I. 2015. *Modeling cities and regions as complex systems: From theory to planning applications*, MIT Press.

Sensitivity analysis of the MATSim transport model

Juste RAIMBAULT^{1,*}; Michael BATTY¹

¹Center for Advanced Spatial Analysis, University College London
Gower Street, London WC1E 6BT, United Kingdom

* j.raimbault@ucl.ac.uk

Keywords: Transport Microsimulation, MATSim, Sensitivity Analysis

Submitted to Special Session SS03

Agent-based transport models are useful tools to devise policies related to diverse urban issues, such as long-term urban planning, transport system operations, or health and public transport. In such contexts, a knowledge of result uncertainties and a certain confidence at least in qualitative patterns obtained, is crucial. The MATSim framework (Multi-Agent Transport Simulation) is a widely used open-source library for such transport simulations, which has been applied to numerous dimensions of transport systems and on many case studies (Axhausen et al., 2016). Studies providing some validation and sensitivity analysis of transport models based on MATSim, such as (Zhuge et al, 2019), remain however rare. We contribute in this presentation to the effort of validating these models with a sensitivity analysis of an open data implementation for the UK. We build on the four-step multimodal transport model introduced by (Raimbault and Batty, 2021), which combines, for any urban area in the UK, the generation of synthetic population using the SPENSER model (Lomax and Smith, 2017), spatial interaction modeling integrating the QUANT model (Batty and Milton, 2021), and the MATSim framework. We study first the sensitivity of output indicators (trip distance and time distributions, modal shares, congestion) to stochasticity, and find an important variability between runs for a given urban area and the same parameters. We then test the sensitivity to some parameters (in particular the number of agents and modal discrete choice parameters) using first local sensitivity analysis, and global sensitivity analysis methods (Moris method and Sobol indices (Campolongo et al, 2011)). We also obtain significant variability of most indicators, suggesting caution for the application of these methods to real-world problems, and the need for further efforts to systematically explore and validate such large-scale integrated urban models.

References

- Axhausen, K., Horni, A., & Nagel, K. (2016). *The multi-agent transport simulation MATSim* (p. 618). Ubiquity Press.
- Batty, M., & Milton, R. (2021). A new framework for very large-scale urban modelling. *Urban Studies*, 0042098020982252.
- Campolongo, F., Saltelli, A., & Cariboni, J. (2011). From screening to quantitative sensitivity analysis. A unified approach. *Computer Physics Communications*, 182(4), 978-988.
- Lomax, N. M., & Smith, A. P. (2017). Microsimulation for demography. *Australian Population Studies*, 1(1), 73-85.
- Raimbault, J., & Batty, M. (2021). Estimating public transport congestion in UK urban areas with open transport models. *GISRUK 2021 Proceedings*.
- Zhuge, C., Shao, C., Wang, S., & Hu, Y. (2019). Sensitivity analysis of integrated activity-based model: Using MATSim as an example. *Transportation Letters*, 11(2), 93-103.

What metrics don't reveal: process accuracy and the evaluation of land use models

Jasper VAN VLIET¹, Bep SCHRAMMEIJER¹; Yuan WANG^{1,2,3}, and Niels DEBONNE¹

¹Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, the Netherlands, Jasper.van.vliet@vu.nl ²School of Geography and Ocean Science, Nanjing University, 163 Xianlin Road, 210023 Nanjing, China ³Key Laboratory of Coastal Zone Exploitation and Protection, Ministry of Natural Resources, Nanjing 210023, China

Keywords: Urban expansion, model evaluation, process accuracy

Background

Spatially explicit land-use models are predominantly evaluated based on the agreement between model results and reference data. Several papers have addressed the issue of what metric should be used for these evaluations (e.g. Pontius and Millones, 2011; van Vliet, Bregt and Hagen-zanker, 2011). Consistently, several automated calibration procedures use one or several metrics as objective function, such as included in the Idrisi land change modeller and the SLEUTH model. Yet, because these models are readily available, and because many calibration procedures are (semi-) automatic, the model structure itself is rarely questioned, and it is implicitly assumed that applying a model is a matter of finding the right model parameters. Here, we argue that there is also a need to evaluate critically whether the structure of a model itself is appropriate all. We argue that a critical investigation of model outcomes can reveal valuable entry points for model improvement, despite not following a strict methodology. To illustrate this point, we provide three examples, all of which are developed using CLUMondo model (van Asselen and Verburg, 2013).

Case 1: Urban expansion and the competition for land in Cambodia

The application described in (Debonne, van Vliet, and Verburg, 2019) projects land system changes in Cambodia for multiple scenarios of agricultural development. Land change is driven by a demand for agricultural commodities, as well as a demand for urban land, as the latter is expected to increase and affect agricultural land as well. In the model urban land and cropland compete for space, based on the suitability of locations (pixels). Many models use a similar approach, although the implementation of this competition can differ. However, urban land is often located in areas that are very fertile for agricultural production, which is reflected in a very high suitability around existing cities (for example around Phnom Penh). As a result, our model simulated cropland persistence there, automatically allocating urban expansion elsewhere (See Fig. 1). Yet, in reality, we find that in the competition for

land, urban expansion prevails and displaces cropland (van Vliet, 2019). In Cambodia, this means that Phnom is continuing to grow at the cost of cropland (e.g. Tierolf, de Moel and van Vliet, 2021). To solve this, we abandoned the direct competition for land, and instead introduced a two-step allocation in which urban development is simulated first, and agricultural changes are allocated only after that.

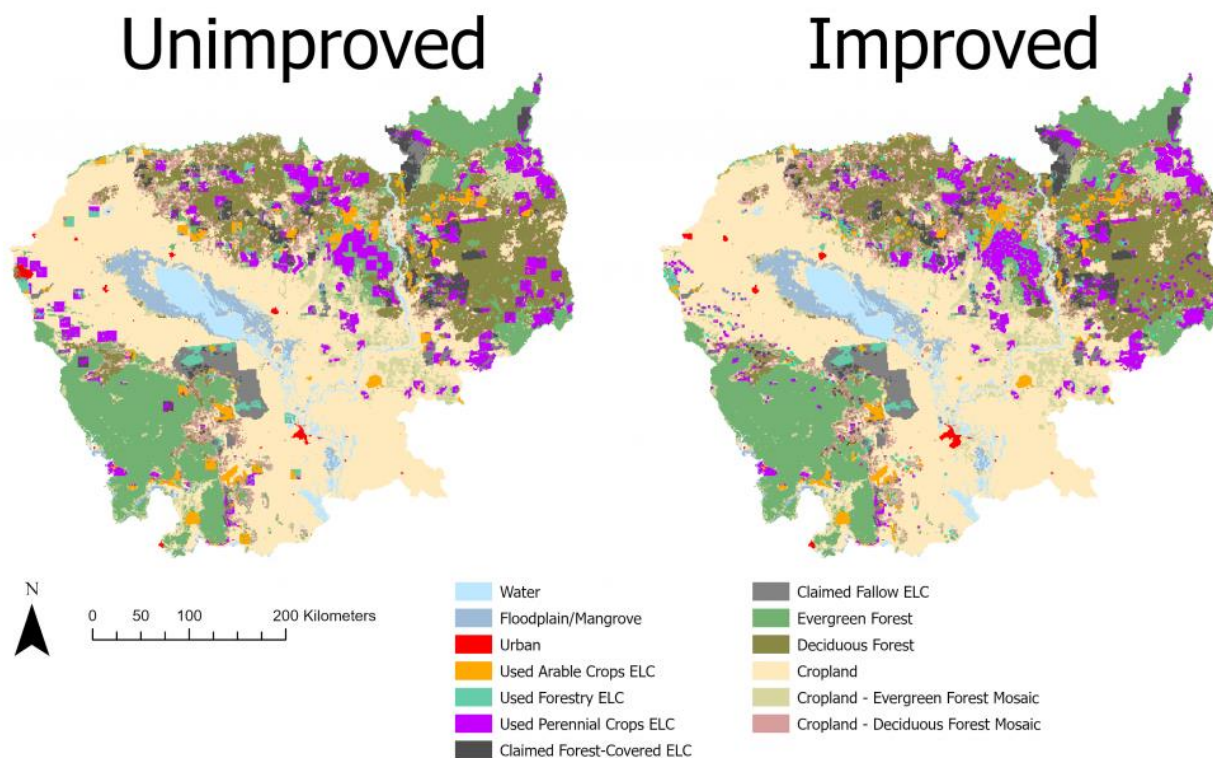


Figure 1: Simulation results before and after taking urban development out of the competition for land. Focus on the large urban area in the western tip of the country in the unimproved result.

Case 2: Village growth in India

In an application of our global land system model we experienced challenges allocating the emergence of village systems and urban systems in India (Wolff *et al.*, 2018). Part of these developments take place at the edge of existing cities and towns already represented in the initial land systems map. Yet, when we compare simulated changes with observed changes (according to the ESA-CCI land cover data) a significant share of the actual development emerged in places that could not be simulated using the usual drivers (neighbourhood effect, various suitability layers), as shown in Fig 2. The missed urban developments, however, were all strongly correlated to small villages and towns already present in the starting year, but not observable in the initial land system map or any of the input layers for suitability analysis. Adding a new layer with “seeds of new settlements”, i.e. the share of built-up land in the initial year from high resolution data or population data could solve this issue, while adjusting existing parameters cannot.

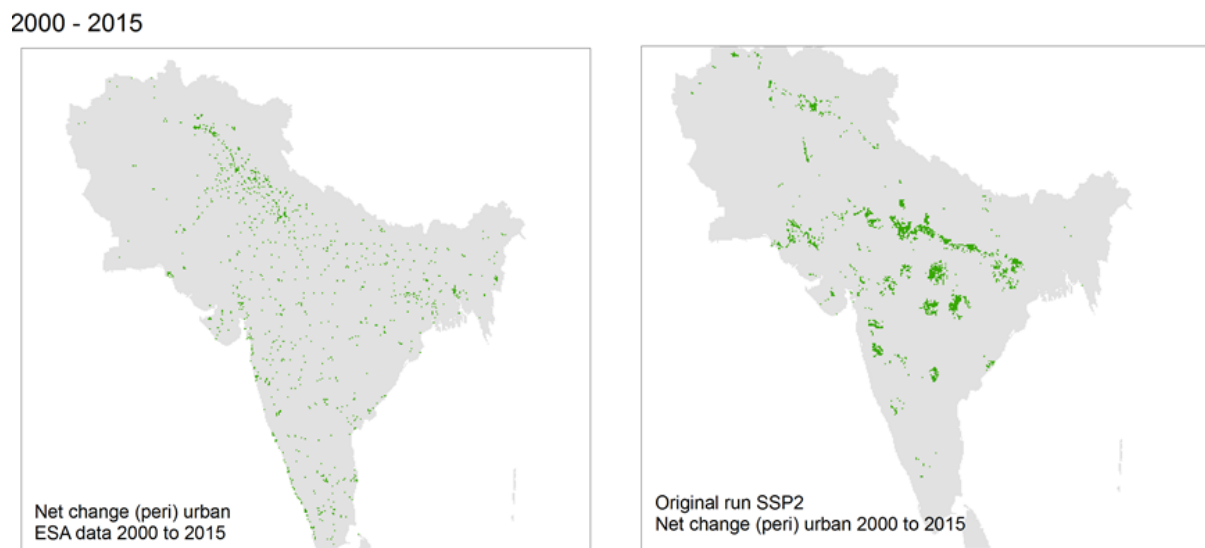


Figure 2: Observed changes in built-up land (left) and simulated changes in village systems and urban systems (right) after calibration of the CLUMondo model.

Case 3: Urban development under population decline in China

The two previous examples dealt with historic land use changes, which allowed for a comparison against observed changes. Yet, an entirely different situation occurs when projecting potential future land use changes. Often, a good fit with historic changes is used to justify extrapolation into the future. However, it is questionable whether this is justified when the change processes itself are likely very different. In an application by (Wang *et al.*, 2021) we simulated urban development under conditions of population decline in China. While this is expected in the next few decades, such developments have not been observed at any meaningful scale in recent times (see Figure 3). Therefore, we could assess whether our model application could replicate historic trends (it did), but extrapolating such trends into the future would lead to a large-scale loss of urban land systems. While there is no precedent at a large scale, similar population trends in Post-socialist Europe do not show massive losses in urban land. As a result, in order to develop realistic scenarios, we had to adjust our model to avoid such results.

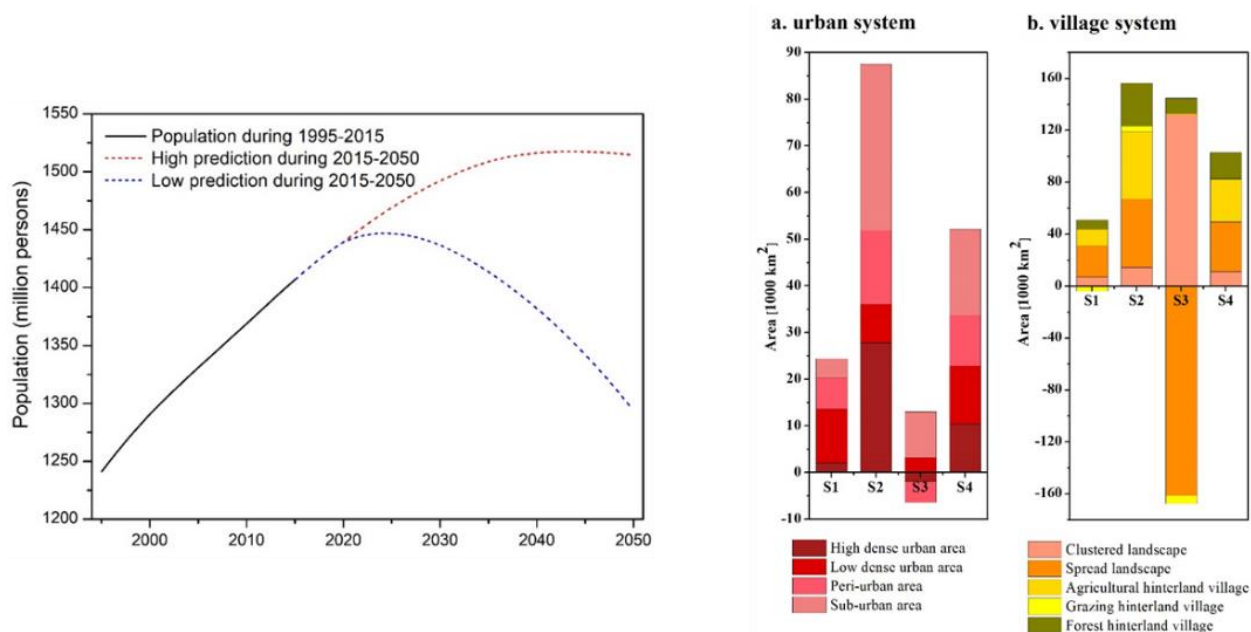


Figure 3: Population trends (left) and simulated land system changes (right) until 2050. Note that scenario S3 and S4 relate to population decline, under conditions of constant population density (S3) and declined population density.

Discussion and implications

The three cases described here illustrate that applying a land use model requires more than simply tuning model parameters. Instead, it also requires knowledge of both the model structure and the actual land change processes being simulated, to assess whether these are applied in a meaningful way. An intuitive way forward would be that we simply need to develop better models, and such calls have been made indeed (Verburg *et al.*, 2019). Yet, we argue here that no model suits them all, as land use change processes differ between locations, and outcomes of interest differ between applications. Therefore, regardless of the need to develop better models, it is also necessary to critically evaluate a model for every new model application. One that has performed well in one context might not perform well in another context.

Our statement that model application requires a critical evaluation of the model structure vis-à-vis the land change process of interest does not mean that other, more formal ways of model evaluation are no longer relevant (e.g. van Vliet *et al.*, 2016). However, we do argue that these metrics should be interpreted cautiously, especially when the purpose of model is to project into the future. While this problem is not unique for land use modelling, it might be more important in this field than in many other fields, due to the complex socio-ecological nature of land use change processes and the often idiosyncratic conditions under which these changes take place.

Acknowledgments

Jasper van Vliet acknowledges the support of The Netherlands Organization for Scientific Research NWO in the form of a VIDI grant (Grant No. VI.Vidi.198.008).

References

- van Asselen, S. and Verburg, P. H. (2013) 'Land cover change or land-use intensification: simulating land system change with a global-scale land change model.', *Global Change Biology*, 19(12), pp. 3648–67. doi: 10.1111/gcb.12331.
- Debonne, N., van Vliet, J. and Verburg, P. (2019) 'Future governance options for large-scale land acquisition in Cambodia: Impacts on tree cover and tiger landscapes', *Environmental Science & Policy*, 94, pp. 9–19. doi: 10.1016/j.envsci.2018.12.031.
- Pontius, R. G. and Millones, M. (2011) 'Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment', *International Journal of Remote Sensing*, 32(15), pp. 4407–4429. doi: 10.1080/01431161.2011.552923.
- Tierolf, L., de Moel, H. and van Vliet, J. (2021) 'Modeling urban development and its exposure to river flood risk in Southeast Asia', *Computers, Environment and Urban Systems*, 87, p. 101620. doi: 10.1016/j.compenvurbsys.2021.101620.
- Verburg, P. H. *et al.* (2019) 'Beyond land cover change: towards a new generation of land use models', *Current Opinion in Environmental Sustainability*, 38, pp. 77–85. doi: 10.1016/j.cosust.2019.05.002.
- van Vliet, J. *et al.* (2016) 'A review of current calibration and validation practices in land-change modeling', *Environmental Modelling & Software*. Elsevier Ltd, 82, pp. 174–182. doi: 10.1016/j.envsoft.2016.04.017.
- van Vliet, J. (2019) 'Direct and indirect loss of natural area from urban expansion', *Nature Sustainability*, 2(8), pp. 755–763. doi: 10.1038/s41893-019-0340-0.
- van Vliet, J., Bregt, A. K. and Hagen-zanker, A. (2011) 'Revisiting Kappa to account for change in the accuracy assessment of land-use change models', *Ecological Modelling*. Elsevier B.V., 222, pp. 1367–1375. doi: 10.1016/j.ecolmodel.2011.01.017.
- Wang, Y. *et al.* (2021) 'Settlement changes after peak population: Land system projections for China until 2050', *Landscape and Urban Planning*, 209, p. 104045. doi: 10.1016/j.landurbplan.2021.104045.
- Wolff, S. *et al.* (2018) 'Meeting global land restoration and protection targets: What would the world look like in 2050?', *Global Environmental Change*, 52, pp. 259–272. doi: 10.1016/j.gloenvcha.2018.08.002.

Special Session SS02.1

4 November 9:00-11:00 am (GMT)

The multitude of spatiotemporal scales in urban systems

Proponents

Janka Lengyel, Seraphim Alvanides, Stephane Roux and Patrice Abry

40 years of sprawl in the USA – Scale and Geographic Effects

Kerry SCHIEL¹; Geoffrey CARUSO¹; Mirjam SCHINDLER²

¹University of Luxembourg

9 Portes des Sciences, L-4366, Luxembourg

+352 466644 9664, kerry.schiel@uni.lu (correspondent author), geoffrey.caruso@uni.lu

²Victoria University of Wellington

Kelburn Parade, 6012, New Zealand

+64 4 463 5645, mirjam.schindler@vuw.ac.nz

Keywords: urban sprawl, scaling, agglomeration

1. Introduction

Urban sprawl in the United States has long been a subject of interest. Its start has been associated with post-war urban development in the 1950s (Barrington-Leigh and Millard-Ball (2015)), as people wanted to live outside of city centres to avoid traffic, noise and to enjoy larger residences (Resnik, 2010). With time, patterns became more complex and an increasingly important topic to research, especially today, given climate change, as sprawl is largely seen as an energy and resource inefficient mode of urban development. It is also a complex topic due to the multidisciplinary nature of sprawl, whereby it not only has a physical element of scattered buildings and the increased area occupied by the city, but its causes relate to economics, demography, sociology, geography and urban planning.

The attempts to define sprawl have generated a huge literature, (recently reviewed by Rubiero-Morollon et al (2020)), highlighting just how multidisciplinary is the topic. The proliferation of definitions of sprawl has also led to several different methods of measurement, each with a slant towards the discipline from which the measurement originated (Jaeger et al, 2010).

The simplest measure of sprawl is density, however, even this method is open to interpretation. For example, it could be calculated using total land area, or just the developed land area. In fact, Galster et al (2001) suggest that “residential density is likely to be a more useful indicator than non-residential development”.

It is also now largely agreed that sprawl is more than just low density. Galster et al (2001) investigated various aspects of the morphology and defined urban sprawl as “a pattern of land-use in an urban area that exhibits low levels of some combination of eight distinct dimensions: density, continuity, concentration, clustering, centrality, nuclearity, mixed use and proximity”.

At a minimum, since Clark et al. (2009) at least, the spatial patterns associated with sprawl quite consistently refer to a scattered mode of urban development at the periphery of cities. In this regard, another common indicator of sprawl is the perimeter/area ratio, which is a measure of the fragmentation or dispersion of urban/residential patches. Higher values indicate higher dispersion, while low values indicate more compact urban zones (Slaev et al, 2017). Along the same idea that sprawl means residential land surrounded by non-developed land, Burchfield et al (2006) calculate the percentage of open space in the 1km² surrounding residential cells and then average those values across an entire city region.

Most of the urban sprawl research is based on case studies, understandably to gain granularity about the ongoing processes and the effect of local conditions. Yet with the progress of computational power and the increased availability of fine-scale land-use data, it is now possible to measure sprawl with precision for large sets of cities. Burchfield et al (2006) pioneered this work and demonstrated it is possible to extract social and economic causes of sprawl from local conditions (first nature). They analysed sprawl for the 275 MSAs in the conterminous United States. Their analysis, though is for 1976 to 1992, thus lacking the last 30 years and is limited to one sprawl index.

Meanwhile, there has also been an increased interest in understanding how the morphology of cities change with their size, as per the scaling literature (Batty (2013) and Bettencourt et al (2007)), where benefits and costs of further agglomeration are under focus. Liaising intra-urban and inter-urban research, Lemoy and Caruso (2018) show urban land profiles are homothetic and thus that there are no fundamental efficiency gains between small and large cities. Their result also means that when cities grow in population, there is both infill closer to the centre and continued sprawl in the periphery. However, they analyse land-use share gradients in a monocentric manner, not the fragmentation/scatteredness, and their analysis is limited to Europe. Thus, in the case of the US, we don't know if increased city size commands further sprawl or infill, and if dispersion goes on in the periphery or if its scatteredness is made of slightly larger clusters (open city-based theory from Caruso et al (2007) or Turner (2005), would suggest it is not).

Our objective here is twofold: first to offer an update on sprawl across the US MSAs, contrasting density, p-a ratio and an adapted Burchfield index and second to identify variations of these measures along city size, i.e. the scaling behaviour of sprawl.

2. Methodology

We calculate sprawl indices using the density and perimeter/area methods, as well as Burchfield's method (although slightly modified into a real focal function) to analyse sprawl in the USA using the most recent data from 2016. Similar to Burchfield et al, we calculate the sprawl indices for the situation at a particular time (static) and for the change in residential development over a time period (dynamic).

2.1 Static Sprawl Indices

First, we will calculate the sprawl indices using the most recent data (2016) from the USGS's National Landcover Database, to assess the current sprawl status, and for 1992, using the same data as Burchfield et al.

We will rank the most and least sprawled MSAs for all the indices (static index for 2016 (S16), 1992 (S92), density 2016 (ρ_{16}), perimeter/area 2016 (P/A16), as well as Burchfield's index for 1992 (B92)) to determine whether there are big differences in the sprawl indices or whether these, despite being calculated using different methods, yield similar results. Through this, we will be able to determine how/if sprawl measurements compare and how much the different measures impact the understanding of sprawl. From the table of rankings, we will also determine if there has been a change in sprawl levels over the last 24 years, or whether it has remained mostly the same.

We will then regress our calculated indices with the variables chosen by Burchfield et al as possible conditions for the development of sprawl, to compare with the results from their B92 index and to see if the same conditions for sprawl remain significant.

To determine if the population of an MSA influences the development of sprawl, we will plot the indices against the population to determine if there is any correlation. Plotting the residuals of the regression model against population will also show whether the variables chosen by Burchfield et al were, in fact, suitable for explaining the presence of sprawl, or if population is also needed as an explanatory variable.

We would like to relate the results from these investigations to the scaling of cities (Lemoy and Caruso, 2018), to determine if the structure of cities is homogenous to population size and whether that has more effect on the development of cities than any other variables shown to contribute to sprawl.

2.2 Dynamic Indices

In this section, we look at the change in residential development over the last 40 years (1976 – 2016), as well as over the last 24 years (1992 – 2016). We will calculate a sprawl index based only on the change in residential cells.

Again, we will use the rankings to determine how this index compares to the static calculation for a particular year.

We will also plot the two dynamic indices against each other to find if any outliers may have more or less sprawl than expected over the last 40 years and endeavour to explain these changes. We will also analyse how the dynamic sprawl index relates to the population size.

Through this research, we hope to update and improve on a previously developed sprawl index (Burchfield et al); determine how our calculation compares to the one calculated by Burchfield et al and to the standard density and perimeter/area measures of sprawl; determine if the variables proposed by Burchfield et al are still significant with our updated sprawl indices; investigate the relationship between the sprawl index and population size and determine how this relates to the scaling of cities. We also calculate indices for the change in development over the last 40 years and determine if the indices yielded indicate many changes in sprawl over the years, or whether the growth and sprawl of a city are endemic to its original development.

References

1. Barrington-Leigh, C. and Millard-Ball A. 2015, 'A century of sprawl in the United States', *Proceedings of the National Academy of Sciences of the United States of America*, volume 112:27, p 8244 - 8249.
2. Batty, M. (2013) 'The New Science of Cities'. Cambridge, Massachusetts: MIT Press.
3. Bettencourt, L.M.A.; Lobo, J.; Helbing, D.; Kuhnert, C. and West, G.B. 2007, 'Growth, innovation, scaling and the pace of life in cities', *Proceedings of the National Academy of Sciences*, volume 104:17, p 7301 – 7306.
4. Burchfield, M.; Overman, H.; Puga, D. and Turner, M. 2006, 'Causes of sprawl: A portrait from space', *The Quarterly Journal of Economics*, p 587 - 633.
5. Caruso, G.; Peeters, D.; Cavailhes, J. and Rounsevell, M. 2007, ' Spatial Configurations in a Periurban City. A Cellular Automata-Based Microeconomical Model', *Regional Science and Urban Economics*, volume 37, p 542 - 567.
6. Clark, J., McChesney, R., Munroe, D. and Irwin, E., 2009. Spatial characteristics of exurban settlement pattern in the united states. *Landscape and Urban Planning*, volume 90:178 - 188.
1. Galster, G.; Hanson, R.; Ratcliffe, M.; Wolman, H. and Coleman, S. 2001, 'Wrestling sprawl to the ground: Defining and measuring an elusive concept', *Housing Policy Debate*, volume 12:4, p 681 - 717.
2. Glaeser, E. and Kahn, M. (2003). 'Sprawl and urban growth'. Henderson, V. Thisse, J.F. (eds) *Handbook of Urban and Regional Economics*, North Holland: Elsevier, p 1 -55.
3. Jaeger, J.A.G.; Bertiller, R.; Schwick, C.; Cavens, D. and Kienast, F. 2010, 'Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl', *Ecological Indicators*, volume 10, p 427 – 441.
4. Lemoy, R. and Caruos, G. 2018, 'Evidence for the homothetic scaling of urban forms', *EPB: Urban Analytics and City Sciences*, volume 47:5, p 870 – 888.
5. Resnik, D. 2010, 'Urban Sprawl, Smarth Growth and Deliberative Democracy', *American Journal of Public Health*, volume 100:10, p 1852 – 1856.
6. Rubiero-Morollon, F. and Garrido-Yserte, R. 2020, 'Recent literature about urban sprawl: A renewed relevance of the phenomenon from the perspective of environmental sustainability', *Sustainability*, volume 12, p 1 - 14.
7. Slaev, A.D. and Nedovic-Budic, Z. 2017, 'The challenges of implementing sustainable development: The case of Sofia's master plan' *Sustainability*, volume 9:15, p 1 - 19.
8. Turner, M. 2005, 'Landscape Preferences and Patterns of Residential', *Journal of Urban Economics*, volume 57:1, p 19 – 54.

Multifractal geometry for integrating several challenges of sustainable metropolitan planning

Pierre Frankhauser¹; Olivier Bonin²

¹Théma, Université de Franche-Comté

32, rue Mégevand, Besançon F-20030, France

+33 3 81 66 54 21, pierre.frankhauser@univ-fcomte.fr

²LVMT, ENPC and Université Gustave Eiffel

6-8 Avenue Blaise Pascal, Champs-sur-Marne F-77420 France

+33 1 81 66 88 68, olivier.bonin@univ-eiffel.fr

Keywords: Sustainable metropolitan development, multifractal approach, urban modelling

The debate about how to manage sustainable development on the scale of metropolitan areas is still ongoing. Residential zones have been developed in what were initially rural areas and therefore far from places of employment and often lack shopping opportunities, public services and public transportation. Hence urban sprawl increased car-traffic flows, generating pollution and increasing energy consumption. New road axes consume considerable space and weaken biodiversity by reducing and cutting natural areas. A return to “compact cities” or “dense cities” has often been contemplated as the most efficient way to limit urban sprawl. However, the real impact of density on car use is less clear-cut. Bouwman, (2000) shows that in the Netherlands the “supposed positive energy related effects of the compact city with regard to its mobility pattern cannot be observed”. Crane, (2000) compares different studies of the relationship between urban form and travel and concluded that the results tend to be contradictory and that “little can be said about the impacts of urban form on travel.” Let us emphasize that moreover climate change will increase, the risk of heat islands on an intra-urban scale, suggesting more nuanced thinking is needed as to how urban patterns should be structured.

Moreover, urban planning cannot ignore social demand. Lower land prices in rural areas, often put forward by economists, is not the only reason of urban sprawl. The quality of the residential environment comes into play, too, through features like noise, pollution, landscape quality, density etc. Schwanen et al., (2004) observe for the Netherlands that households preferring a quiet residential environment and individual housing with a garden will not accept densification, which might even impel them to move to lower-density rural areas even farther from jobs and shopping amenities. Hence a return to a compact city policy may even encourage urban sprawl. Many scholars emphasize the importance of green amenities for residential environments. Some studies show that individuals prefer natural to built-up environments (Lamb & Purcell 1990), others report the importance of easy access to leisure areas (Guo & Bhat 2002). Moreover, vegetation in the residential environment has an important impact on health and well-being. Braubach, (2007) shows that health is directly linked to the quality of the residential environment.

In order to find acceptable and sustainable solutions for future development, it seems important considering different factors by developing spatial arrangements that best reconcile environmental and social issues. As pointed out, just promoting compact cities does not seem any more realistic than accepting completely anarchical and uncoordinated development. The purpose of this paper is to present a planning concept developed in recent years under the name Fractalopolis (Frankhauser, 2019, (Bonin & et al. 2020).

We start by considering social demand by referring to the fundamental work of Maslow who considers human needs from a very broad anthropological standpoint. Max-Neef et al., (1991), referring to this approach, reiterates the basic needs of “Subsistence, Protection, Affection, Understanding, Participation, Idleness, Creation, Identity and Freedom”. He introduces the concept of satisfiers assigned to meet these needs. Food, for example, is not viewed as a need but as a satisfier of the basic need of Subsistence. Satisfiers thus become the link between the needs of everyone and society. They provide goods that meet the needs of individuals or households and thus correspond in our context to the different types of amenities. We consider their importance, their location and their accessibility. Moreover, we rank the needs according to their importance for individuals or households. Certain needs like food shopping must be met every day; others like DIY stores are more rarely used. In order to enjoy a good quality of life and to shorten trips and particularly to reduce automobile use, it seems important for satisfiers of daily needs to be easily accessible. Hence, we take into account the frequency of recourse when reflecting on the implementation of shops which is reminiscent of central place theory but also satisfiers, which are important in themselves like primary care doctor.

The second important feature is taking care of environment and biodiversity by avoiding fragmentation of green space which must benefit, moreover, of a good accessibility, as pointed out. These areas must, too, play the role of cooling areas ensuring ventilation of urbanized areas.

For integrating these different objectives, we propose a concept for developing spatial configurations of metropolitan areas designed which is based on multifractal geometry. It allows combining different issues across a large range of scales in a coherent way. These issues include:

- providing easy access to a large array of amenities to meet social demand;
- promoting the use of public transportation and soft modes instead of automobile use;
- preserving biodiversity and improving the local climate.

The concept distinguishes development zones localized in the vicinity of a nested and hierarchized system of public transport axes. These axes link a hierarchical system of central places. Around the unique main center, we put N smaller sized subcenters each of them being again surrounded by N third order centers. The system is iterated up to 4 to 6 steps.

We compared a compact model consisting of the same number of ranked centers with a fractal one. It turns out that a fractal arrangement optimizes the access to the different types of amenities at least if residents want to benefit to access to green areas.

The highest ranked center offers all types of amenities, whereas lower ranked centers lack the highest ranked amenities. The lowest ranked centers just offer the amenities for daily needs. A coding system allows distinguishing the centers according to their rank.

Each subset of central places is in some sense autonomous, since they are not linked by transportation axes to subcenters of the same order. This allows to preserve a linked system of green corridors penetrating the development zones across scales. This system consists of some large corridors and an increasing number of smaller and smaller corridors up to a very local scale thus avoiding the fragmentation of green areas and ensuring a good accessibility to recreational areas.

These green areas are preserved and no or only a feeble development is possible. For those being closer to the more important centers we usually admit a slightly higher degree of development as for the large zones, farer away from the urbanized areas.

The spatial model is completed by a population distribution model which globally follows the same hierarchical logic. However, we weakened the strong fractal order what allows to conceive a more or less polycentric spatial system. Let us emphasize that is possible to introduce other types of spatial attributes than just development and preserved areas. Hence, e.g. we may distinguish in a similar way individual housing from collective housing, agricultural areas from and forests etc.

We can adapt the concept easily to real world situation without changing the multifractal logic by localizing the development zones according to existing cities and public transport axes. For this aim a software package has been developed by Gilles Vuidel which allows simulating scenarios and evaluate them. The evaluation procedure is based on fuzzy evaluation of distance acceptance for accessing to the different types of amenities according to the ranking of needs. We used for evaluation data issued from a great set of French planning documents like Master plans.

We show a recent application of the model, which has been set up in the frame of research project financed by the ADEME, the French Agency of Ecological Transition.

References

Bonin, O & et al. 2020, *Projet SOFT sobriété énergétique par les formes urbaines et le transport*, ADEME, Angers.

Bouwman, ME 2000, 'Changing mobility patterns in a compact city: Environmental impacts', in *Compact cities and sustainable urban development. A critical assessment of policies and plans from an international perspective*, Ashgate Publishing, pp. 229–240, retrieved February 18, 2021, from <<https://www.taylorfrancis.com/chapters/changing-mobility-patterns-compact-city-environmental-impacts-bouwman/e/10.4324/9781315189369-18>>.

Braubach, M 2007, 'Residential conditions and their impact on residential environment satisfaction and health: results of the WHO large analysis and review of European housing and health status (LARES) study', *International Journal of Environment and Pollution*, vol. 30, pp. 384–403.

Crane, R 2000, 'The Influence of Urban Form on Travel: An Interpretive Review', *Journal of Planning Literature*, vol. 15, no. 1, pp. 3–23.

Frankhauser, P 2019, 'Complexity and fractal urban planning',.

Frankhauser, P, Tannier, C, Vuidel, G & Houot, H 2018, 'An integrated multifractal modelling to urban and regional planning', *Computers, Environment and Urban Systems*, vol. 67, no. 1, pp. 132–146.

Guo, J & Bhat, C 2002, *Residential Location Modeling: Accommodating Sociodemographic, School Quality and Accessibility Effects*, University of Texas, Austin TX.

Lamb, RJ & Purcell, AT 1990, 'Perception of naturalness in landscape and its relationship to vegetation structure. Landscape and Urban Planning', *Landscape and Urban Planning*, vol. 19.

Max-Neef, MA, Elizalde, A & Hopenhayn, M 1991, *Human scale development: conception, application and further reflections*, The Apex Press, New York.

Schwanen, T, Dijst, M & Dieleman, FM 2004, 'Policies for Urban Form and their Impact on Travel: The Netherlands Experience', *Urban Studies*, vol. 41, no. 3, pp. 579–603.

A dominance tree approach to systems of cities

Thomas LOUAIL¹ ; Marc BARTHELEMY^{2,3}

¹ CNRS, Géographie-cités

Campus Condorcet, 5 cours des humanités, FR-93322 Aubervilliers, France

thomas.louail@cnrs.fr

(corresponding author)

² Institut de Physique Théorique, CEA, CNRS-URA 2306

FR-91191, Gif-sur-Yvette, France

³ CAMS (CNRS/EHESS)

54 Avenue de Raspail, 75006 Paris, France

marc.barthelemy@ipht.fr

Keywords: spatial hierarchies; marked point processes; Voronoi tessellation; systems of cities; US and France historical city populations (1870-2010)

Understanding the organization of urban systems has always been a central challenge in geography and economics [1–3]. On one hand the population sizes of cities have been extensively discussed [4–6] and have been shown to follow a broad distribution: there is a hierarchy of cities characterized by many small cities, a few medium cities and a very few large ones, whose population sizes are much larger than the rest. While this distribution has been fully characterized and discussed [6,7], the spatial distribution of cities has also been an important subject of debate [8–12]. Cities are obviously not scattered at random but follow some logic based on geographical constraints, economical considerations and historical path dependency. Central place theory [8] sought to explain the spatial distribution of cities of different sizes based on the idea that settlements function as ‘central places’ that provide services to surrounding areas. The result of consumers’ preferences is then a system of centers of various sizes, forming different levels of a hierarchy. A consequence obtained by Christaller is that the most efficient pattern to serve areas without overlap is a triangular or hexagonal arrangement of settlements. Although this idea has been very inspirational to many scientists, few works have evaluated it empirically, thanks to a shared and well-defined quantitative method. An important contribution to this problem is due to Okabe and Sadahiro [13] who showed that random (uniform) arrangements of cities could match several of Christaller’s findings. To reach this result they defined a non-parametric procedure to capture the spatial dominance of a city on another, and in this talk we will elaborate on this idea.

More generally, the difficulty of characterizing the spatial organisation of urban systems appears in many problems of spatial statistics [14, 15] where the points (cities in this case) have a location and are described by at least one quantity, such as their population size. In spatial statistics this general setting is referred as a ‘marked point process’ [14], in which each point of the process carries extra information called a mark. This setting is relevant for many fields ranging from ecology (positions of plants of different species) to epidemiology (locations of infected individuals). Standard tools developed for the analysis of point processes usually consist in measuring spatial autocorrelations [16], or testing the null hypothesis of complete spatial randomness with the K or L statistics which summarize the deviations from specific distributions [14, 15]. These tools were then extended to marked point processes and describe deviations from well-defined cases, or intensity and moments measures [17]. A tool that would go beyond these measures and provide a more precise characterization of these processes would then be extremely useful for a wealth of different problems.

To further characterize marked point processes it is possible to start from geometrical structures constructed on top of the point pattern. These ‘secondary structures’ [14] comprise in particular tessellations and networks [18]. The Voronoi tessellation is one of the most relevant structure in computational geometry, and is of major importance in the resolution of many problems, in particular in location science [19]. Networks can also be constructed over a set of points, and useful tools include proximity graphs (such as the random geometric graph) or excluded volume graphs (such as the Gabriel graph). Measures on these secondary structures can then characterize the point process itself, and constructing a spatial network on top of the point process enables to import all the networks knowledge into spatial statistics. However, approaches based on secondary structures are also often used to estimate deviations from uniformity, or the importance of some correlations. Here, starting from the Voronoi tessellation, we will present a tool that is useful to understand the multiscale structure of a system of marked points, and to compare it to other systems. Such a tool for characterizing the organization of this type of systems should thus encode both the spatial information and the population size (in more general terms both the positions and the marks of the points). The purpose of this tool is not to describe the deviation from a uniform distribution, but to compare two systems (such as urban systems in different countries for example) and to understand the temporal evolution of these systems from a non-local, high-level perspective.

In this talk we will first describe the dominance tree, introduced in [13], that will constitute the basis of our analysis. We will then introduce different measures to extract information from this tree. In

particular, we will show that the height of a node in this tree encodes both its mark and space-related information. We will illustrate this method on toy models, and then on empirical data, to compare the evolution of the French and the US urban systems over the 20th century.

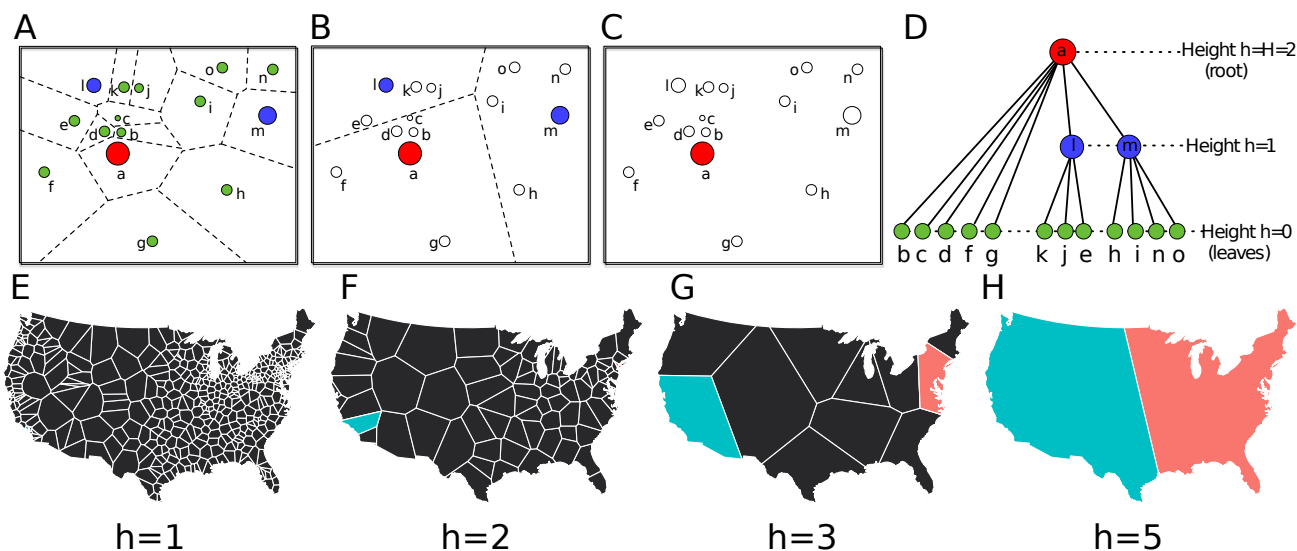


Figure 1: Construction of the dominance tree. (A) Each point is characterized by a mark/value represented here as the size of the circle. We first construct the Voronoi tessellation (shown with dotted lines) for these points and we observe that there are 3 local maxima: a , l and m , in red and blue. These are the points which do not have any neighbor whose mark is larger than their own. (B) From step (A) we keep the local maxima only, we construct the Voronoi tessellation over this set and determine the local maxima (which is here the node a). (C) At the end of the process, we are left with the node with the largest mark. (D) The entire process can be described by the 'dominance tree': leaves correspond to the points which are not local maxima, and the children of a given node are the points that belong to its attraction basin. For example, as can be seen in (B), nodes k , j and e belong to the attraction basin of node l , which is a local maximum at level/height 1. (E-H) Successive Voronoi tessellations obtained when applying this 'decimation' process for US cities marked with their 2010 population sizes. After $H=6$ iterations, we obtain a single city. By construction, the node with the largest mark will always be the root of the dominance tree, but for the other nodes their height in this tree will not only depend on their mark but also on their location.

References

- [1] Helsley, R. W., & Strange, W. C. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and urban economics*, 20(2), 189-212.
- [2] Sanders, L., Pumain, D., Mathian, H., Guérin-Pace, F. & Bura, S. (1997). SIMPOP: a multiagent system for the study of urbanism. *Environment and Planning B: Planning and design*, 24(2), 287-305.
- [3] Rozenblat, C., & Pumain, D. (2007). Firm linkages, innovation and the evolution of urban systems. *Cities in globalization: Practices, policies, theories*, 130-156.

- [4] Zipf, G. K. (2016). Human behavior and the principle of least effort: An introduction to human ecology. Ravenio Books.
- [5] Gabaix, X. (1999) Zipf's law for cities: an explanation. *The Quarterly journal of economics* 114, 739-767.
- [6] Verbavatz, V. & Barthelemy, M. (2020) The growth equation of cities. *Nature* 587(7834): 397-401.
- [7] Cottineau, C. (2017). MetaZipf. A dynamic metaanalysis of city size distributions. *PloS one*, 12(8), doi.org/10.1371/journal.pone.0183919
- [8] Christaller, W. (1933) Die zentralen Orte in Süddeutschland : eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen. PhD thesis, Erlangen (Gustav Fischer, Jena)
- [9] Ullman, E. (1941). A theory of location for cities. *American Journal of sociology*, 46(6), 853-864.
- [10] Haggett, P. (1965) *Locational analysis in human geography*. Edward Arnold Publishers, London.
- [11] González-Val, R. (2019) The Spatial Distribution of US Cities. *Cities* 91:157-64. <https://doi.org/10.1016/j.cities.2018.11.015>.
- [12] Strano E., Simini F., De Nadai M., Esch T. and Marconcini M. (2020) The agglomeration and dispersion dichotomy of human settlements on Earth. arXiv:2006.06584.
- [13] Okabe A. & Sadahiro Y. (1996) An illusion of spatial hierarchy: spatial hierarchy in a random configuration. *Environment and Planning A* 28(9):1533–52.
- [14] Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008) *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons.
- [15] Jacobsen, M. (2006) *Point process theory and applications: marked point and piecewise deterministic processes*. Springer Science & Business Media.
- [16] Moran, P. A. (1950) Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17-23.
- [17] Haenggi, M. (2012) *Stochastic geometry for wireless networks*. Cambridge University Press.
- [18] Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. (2000) *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams (Second Edition)*. John Wiley & Sons.
- [19] Laporte G, Nickel S, Saldanha da Gama F (eds.) (2015). *Location Science*, Springer.

Radial analysis and scaling of urban land use and sprawl

Rémi LEMOY¹; Geoffrey CARUSO^{2,3}; Paul KILGARRIFF³

¹University of Rouen Normandie, IDEES Laboratory UMR 62 66 CNRS,

Bâtiment n°7, 17 rue Lavoisier, 76130 Mont-Saint Aignan, France

+33 2 35 14 60 65, remi.lemoy@univ-rouen.fr (correspondent author)

²University of Luxembourg, Department of Geography and Spatial Planning,

Maison des Sciences Humaines, 11 Porte des Sciences, L-4366 Esch-Belval, Luxembourg

+352 46 66 44 66 25, geoffrey.caruso@uni.lu

³Luxembourg Institute of Socio-Economic Research,

Maison des Sciences Humaines, 11 Porte des Sciences, L-4366 Esch-Belval, Luxembourg,

p.kilgarriff1@gmail.com

Keywords: urban land use, urban sprawl, urban scaling laws, radial analysis, European cities

Cities are at the heart of many challenges to our modern societies, including climate change, air pollution, urban heat islands, and the growing scarcity of different resources (arable land, rare metals, fossil fuels...). In this context, defining cities in an accurate way and understanding them is a crucial challenge. In particular, as cities exist in a wide range of sizes, it is important to understand and characterize the relation between the size and the structure and evolution of cities.

In this work, we study cities in a radial manner : distance to the center is the first determinant of many variables characterising cities like land use, population density, land and housing prices, air pollution etc. The simple lens of radial analysis allows us to study a large number (hundreds) of cities simultaneously, across a wide range of sizes, and to make sense of their radial profiles.

The present work is based on previous research (Lemoy and Caruso, 2020, 2021) which showed that the radial profile of artificial land use in cities scales in the two horizontal dimensions like the square root of total population, and that artificial land use decreases exponentially with the distance to the center.

1. Data and methods

1.1 Data

The data used for this work is similar to previous work (Lemoy and Caruso, 2020), namely the 2006 Urban Atlas dataset for a precise description of land use in the 300 largest functional urban areas (FUAs) in Europe (representing approximately 40% of European population) and the 2006 Geostat

population grid for the computation of the total population of these cities. A second part of this work uses additional land use data from another work on 30 large Chinese cities (Jiao, 2015).

1.2 Methods

For all European cities, we choose the location of the (main or historic) city hall as the location of the center, and average land use in concentric rings of fixed width 141m, which yields radial profiles $s_N(r)$ of artificial land use, functions of the distance r to the center and of total city population N . Then following Lemoy and Caruso (2021), we perform a simple exponential (non-linear) fit of these radial profiles according to $s_N(r) \sim \exp(-r/l_N)$, where the unique parameter l_N gives a characteristic distance (radius) of the artificial land use gradient, that is, a measure of the spatial scale of the considered city in terms of land use.

2. Results

2.1 Urban land use and City Mass Index (CMI)

The results of this analysis, shown on Fig. 1, correspond to the expectations of Lemoy and Caruso (2020): the characteristic distance l_N scales with the square root of total population N . This means that artificial land use per capita is independent of city size. And coastal cities (shown in colors on Fig. 1) are the main exception to this rule, since part of the land is covered by water bodies.

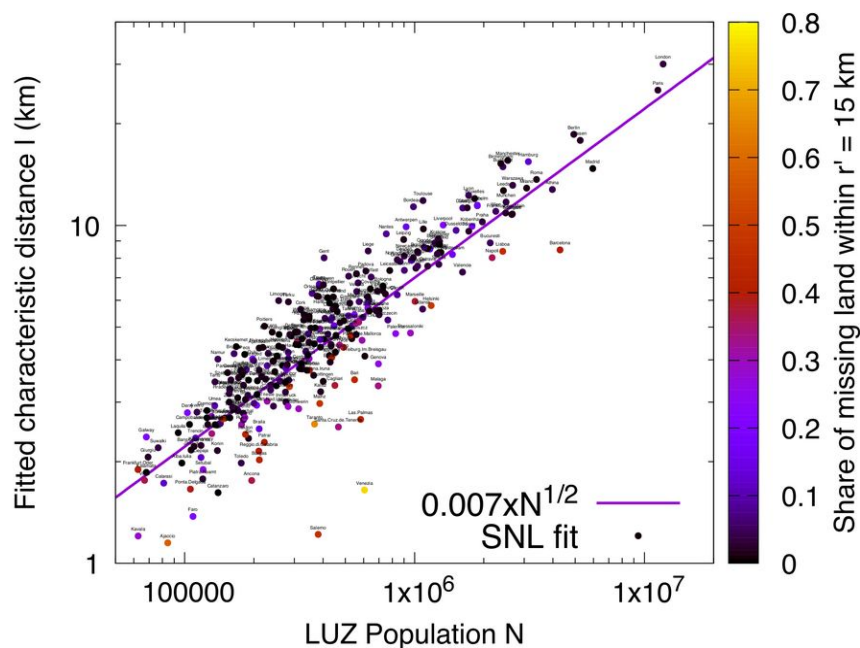
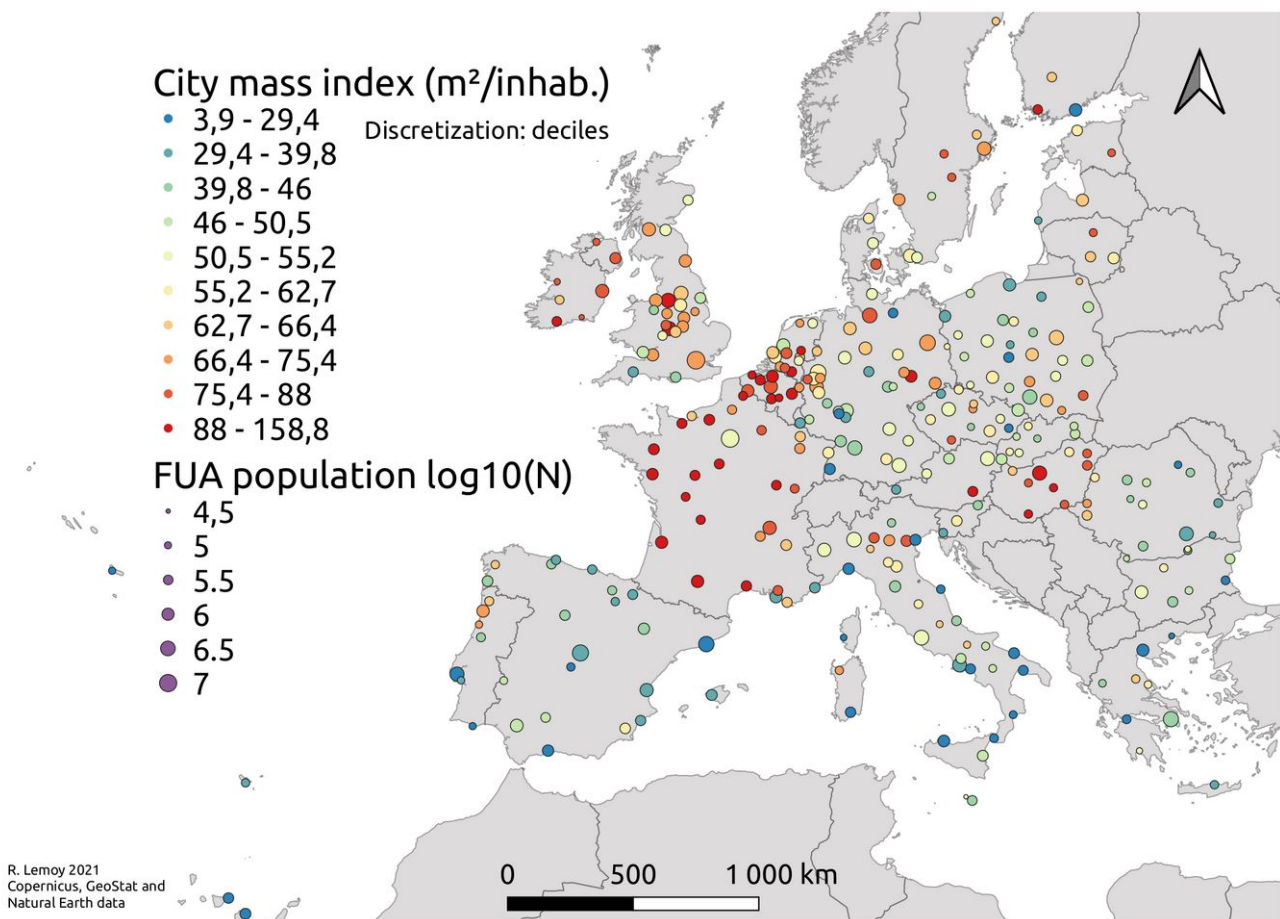


Figure 1: Estimated characteristic decrease distance l_N of the simple (one parameter) exponential/non-linear fit (SNL) of artificial land use profiles, displayed as a function of the total population N . The straight line gives the expectation of Lemoy and Caruso (2020): a scaling with the square root of total population.

This result suggests to define a city mass index, which would be similar to the body mass index (or Quetelet index), but adapted for the social body of cities : the body mass index of humans is defined as $mass/length^2$, and the present study suggests a city mass index computed by $population/length^2$, where the size is measured by the characteristic distance l_N of decrease of the artificial land use radial profile. This index gives a characteristic population density in the city's artificial core. If l_N is measured in kilometers, our analysis shows that this city mass index would have on average a value of nearly 20 000 inhabitants/km² in our European dataset. Higher values would correspond to compact cities, while lower values would indicate sprawled cities (not forgetting that coastal cities need to be treated separately). Actually, we prefer to use the inverse indicator l_N^2/N , which has more intuitive variations. It corresponds to the characteristic artificial area per capita, and would give on average nearly 50 m² in our dataset. Higher values indicate sprawled cities while lower values would indicate compact ones.

We then define the city mass index CMI of a city of total population N as $CMI = l_N^2/N$.



Map 1: the city mass index CMI is showing a clear national effect. The size of symbols indicates the decimal logarithm of FUA population N .

This is a measure of the residual of the scaling relationship displayed on Fig. 1, which is presented on Map 1. In addition to the fact that coastal cities have low values of the CMI, as discussed before, this map shows a clear national effect: for instance, French and Belgian cities tend to have high values of the CMI, meaning that they are quite extended in terms of artificial land use compared to their total population, while Spanish ones have low values on average. This phenomenon must be studied in more detail in order to understand and improve national planning practices.

2.2 Urban sprawl

We also study the evolution in time of both total population N and of artificial land use profiles, measured by the characteristic radius l_N defined above. We perform this study on the previous European dataset using the 2006 and 2012 versions of the Urban Atlas, and on 30 large Chinese cities studied by Jiao (2015) for the years 1990, 2000 and 2010. We note that city definitions and land use data are different between both datasets, but the methods should not be too sensitive to this fact (although this remains to be tested).

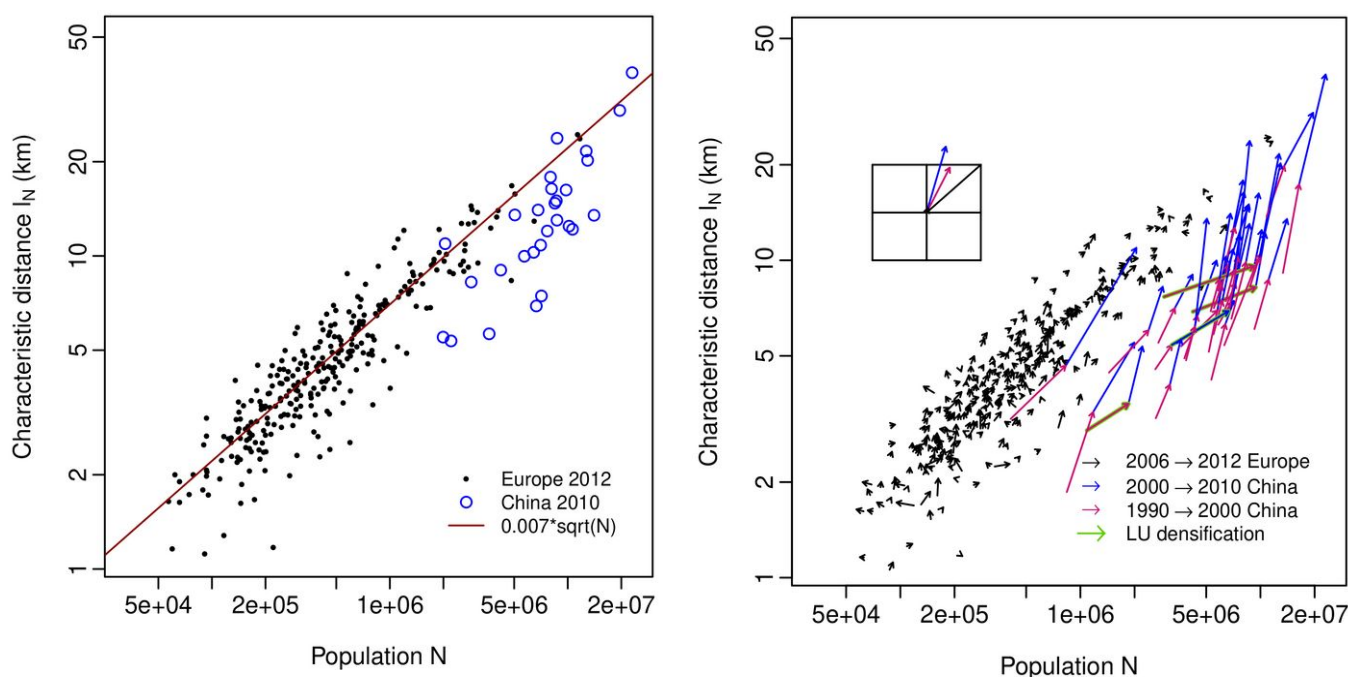


Figure 2: Characteristic radii l_N of urban areas in terms of artificial land use against population size N . Left panel: European (2012) and Chinese (2010) cities. The \sqrt{N} scaling law's line is shown as a guide to the eye. Right panel: for Chinese cities, evolution between 1990, 2000 and 2010 (and for European cities in black, 2006-2012) of these characteristic radii l_N against their evolution in population size N . All 28 Chinese cities have an artificial land area increasing faster than their population, except Haerbin, Xining and Shijiazhuang between 1990 and 2000, and Nanning between 2000 and 2010 (densification in terms of land use). The inset shows the average evolution over all cities for each of the two periods.

The results show that cities are mostly sprawling, both in Europe and in China, in the sense that their artificial area increases faster than their population. The evolution in Europe is much slower than in China (European arrows are much shorter on Fig. 2). Chinese cities in 1990 seem much more compact than European ones, but they are sprawling enough to be nearly as sprawled as European ones in 2010.

We conclude that the tools developed here give us new insights on urban land use and the challenging worldwide phenomenon of urban sprawl. They facilitate comparisons between cities of different sizes and could help urban planning and policy regarding land use and sprawl.

Acknowledgments

The research was supported, in part, by funding from FNR Luxembourg with the SCALE-IT-UP project within the CORE program. We heartily thank Limin Jiao for sharing his land use profiles data from Chinese cities, and acknowledge stimulating discussions with V. Viguié, D. Rybski and M. Barthélémy.

References

- Jiao, L. (2015) 'Urban land density function: A new method to characterize urban expansion', *Landscape and Urban Planning*, 139, pp. 26–39.
- Lemoy, R. and Caruso, G. (2020) 'Evidence for the homothetic scaling of urban forms', *Environment and Planning B: Urban Analytics and City Science*, 47(5), pp. 870–888.
- Lemoy, R. and Caruso, G. (2021) 'Radial analysis and scaling of urban land use', *under review*.

Geovisualization of artificial land use in European cities in 2006, with urban scaling laws

PECHERIC Axel¹; LEMOY Rémi²; LE TEXIER Marion³; CARUSO Geoffrey⁴

¹University of Rouen, IDEES Laboratory UMR 6266 CNRS, Mont-Saint-Aignan, France (correspondent author).

+33 06 58 98 46 50, axel.pecheric@univ-rouen.fr

²University of Rouen, IDEES Laboratory UMR 6266 CNRS, Mont-Saint-Aignan, France.

+33 2 35 14 60 65, remi.lemoy@univ-rouen.fr

³University Paul Valéry Montpellier 3, LAGAM, Montpellier, France.

marion.le-texier@univ-montp3.fr

³University of Luxembourg, Department of Geography and Spatial Planning, Maison des Sciences Humaines,

Luxembourg

+352 46 66 44 66 25, geoffrey.caruso@uni.lu

Keywords: geovisualization, land use, urban scaling laws, radial analysis

Almost all cities, even those with a decreasing population, continue to spread out with a growing urbanization. Current urbanization tends to consume a lot of space (suburban areas, business parks), combined with strong de-densification in rural areas (less attractive), which leads to a significant increase in soil artificialisation on the outskirts of large cities. The constant growth of cities reveals sustainability issues in the housing and transport sectors by challenging the spatial organization of cities, for example by increasing travel times and car use. Cities are major sources of pollution, even more in the context of climate change and the health of city residents. They are also responsible for the depletion of many energy resources. Moreover, heavy urbanisation and the constant increase in artificial land use create profound environmental consequences (impacts on fauna and flora, urban heat island...). And urban sprawl is often an irreversible process.

The desire to have a more sustainable management of urban space, for example by reducing the artificial land use, have been at the heart of urban development policies for decades now.

This web application presented here focuses on artificial land use in 2006 with urban scaling laws, in 305 European urban areas with more than 100,000 inhabitants, described by harmonized land use databases. We use the Copernicus Urban Atlas 2006 land use database at 5 m resolution and the Geostat population grid at 1 km resolution. Artificial land use refers to all land use categories in the Urban Atlas except water areas, agricultural spaces, urban green spaces and forests.

The objective is to analyze how the urban structure (more specifically land uses) changes with distance from the main center.

Urban scaling laws (which allow to transform a set of objects from one spatial scale to another without changing their structure) prove themselves to be a great tool to compare cities (Batty, 2015). Indeed, after rescaling distances from City Hall, European cities are positioned at similar levels on the issue of land use (figure 1). The large corpus of cities to be treated and the objective of comparison on a European scale reinforces the relevance of the use of urban scaling laws. Thanks to rescaling, comparison becomes possible for cities of different sizes and geography makes it possible to provide spatial comparability of territories.

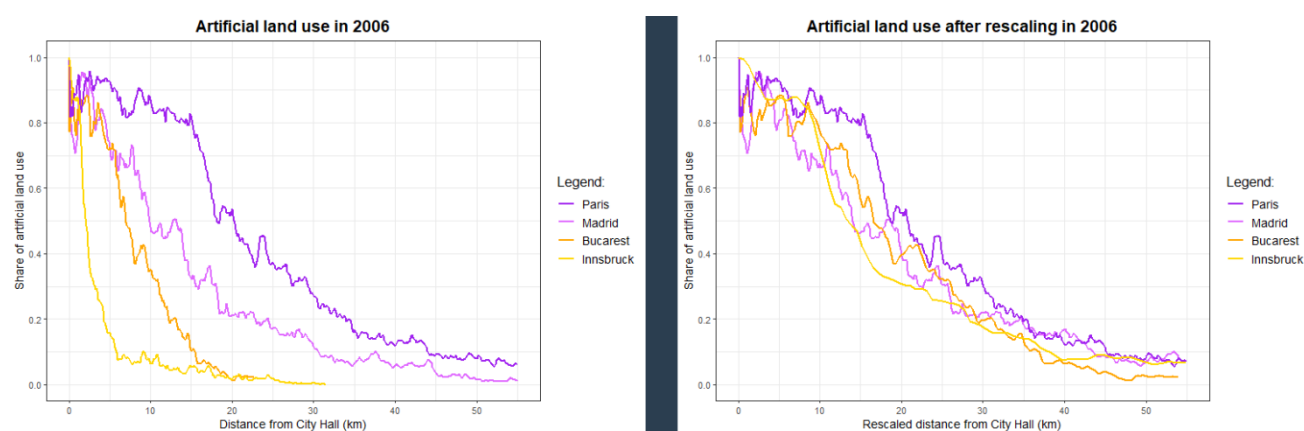


Figure 1. The regularity of land use at the European level with urban scaling laws

The methodological choice behind this work is the use of radial analysis to describe the cities. The radial analysis allows the complete distribution of land and people in the city. This is based on Alonso's model, the city is considered as circular and monocentric. The reference center for the city corresponds to the city hall in each case.

The city does not only represent the urban core, but the peripheral spaces must also be considered in the analysis of urban scaling laws (and thus the whole urban environment is taken into account). The city is studied morphologically, focusing on the organization of the succession of land uses. Using a distance-based approach allows us to better grasp the phenomenon of urban sprawl and how to better control it.

The main result of this work is that the organization of land use successions occurs at the same rate if we cancel out the size effect of urban areas (Lemoy, 2020). Larger cities are larger in terms of area but are not more parsimonious in their land use per capita.

A web platform was developed using the Shiny package of R to favour interactivity. The R language is well suited to the use of interactive cartography. Thematic tabs allow the user to move easily from one visualisation tool to another and to access specific methodological descriptions. This web app is organized in several tabs: Interactive map (figure 2), Methodology (to go deeper into urban scaling laws and the datasets), European examples (to see cities before and after rescaling with graphs and maps on typical cities), France versus Spain (to see differences on urban fabric, industries and roads land use), "And my city, how does it stand?" (statistical and cartographic comparisons of the cities selected by the user).

The application has been designed as a "step-by-step" application, starting with the discovery of the interactive map and ending with the visualization of similar European cities according to the city chosen by the user. For the moment, it is in its infancy.



Figure 2. Estimate of the size of cities required to obtain a share of land artificialization found on average on a disc surrounding London by 50 km.

The web platform can be described as an intuitive analysis interface thanks to the R Shiny language which is an open-source technology and can allow the application to be displayed on a browser, making it accessible to a wide audience through geovisualization (Keim and al., 2008). This web app, which can free us from the sizes of the cities, also allows us to compare urban forms while having the indication of the rescaled distance to the city center.

Visualization tools (Keim and al., 2008) are useful to operationalised the use of scaling laws for managing and comparing cities, and to better understand this complex phenomenon. With this comparative tool, such as a didactic and cartographic platform, cities can assess their own situation but also to compare themselves to other similar cities in 2006 (figure 3). A new point of view on urban areas is thus possible, by including cities in a broader than regional debate on land use.

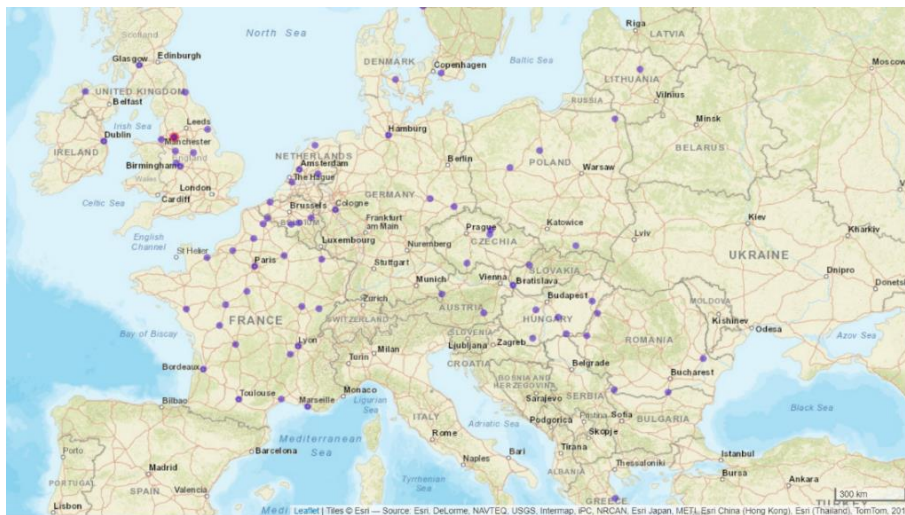


Figure 3. European cities that are more than 95% similar in land use succession to Manchester in 2006

Nevertheless, urban scaling laws can't explain everything either: there are some cities that stand out, and some variations that appear at the margin. In particular, two large countries stand out compared to the average European city: French cities are more artificial than the average while Spanish cities are less (figure 4), more specifically in the first 10-30 kilometers (rescaled distance with London as reference). It is therefore interesting to study the differences in artificial land uses and to analyse which types of artificial spaces can be found in excess in France.

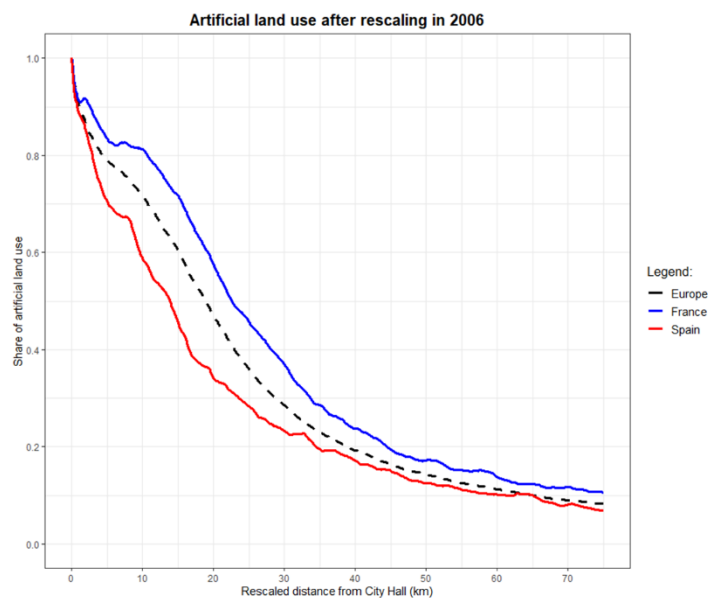


Figure 4. Differences in the share of artificial land uses between French and Spanish cities

According to first analyses on artificial land uses, it is mostly the proportion of urban fabric in France which is more important (after rescaling) than in Spain. The proportions of industries and roads land

use are also higher in France. But the differences are smaller in relation to urban fabric land use, which means that in Spain there are more roads in comparison to available urban fabric than in France.

Urban contexts (topographical, historical and cultural effects) and planning policies are different between France and Spain, which explains some differences in artificialization. The urbanization in Spain contains more small compact clusters (Laborde, 1984) between agricultural lands, while in France at equivalent size (after controlling the size effect through rescaling) one observes more low density peri-urban areas. In Spain, until 1960, urban development choices were made to the use the land already available within the cities and the construction of high buildings, which may explain why Spanish cities are more compact. Then, in order to cope with the arrival of new city dwellers, satellite cities were built near large metropolises (Díaz, 2007).

References

- Batty, M. (2015). Competition in the Built Environment: Scaling Laws for Cities, Neighbourhoods and Buildings. *Nexus Network Journal* 17, pp. 831-850
- Díaz, H. and al. (2007). Évolution du processus urbain espagnol dans la seconde moitié du XXe siècle. *Sud-Ouest européen*, tome 23, pp. 89-106
- Keim, D. and al. (2008). Visual Analytics: Definition, Process and Challenges. *Information Visualization - Human-Centered Issues and Perspectives*, Springer, pp. 154-175
- Laborde, P. (1984). Espaces ruraux et villes-centres en Espagne. *Revue géographique des Pyrénées et du Sud-Ouest*, tome 55, Péninsule Ibérique et Pyrénées, pp. 495-505
- Lemoy, R. and Caruso, G. (2020). Evidence for the homothetic scaling of urban forms, *Environment and Planning B*.

Parallel Session PSB3

4 November 9:00-11:00 am (GMT)

Networks

The fallacy of the closest antenna: Towards an adequate view of device location in the mobile network

A. OGULENKO¹; I. BENENSON^{*1}; M. TOGER²; J. ÖSTH^{2,3}; A. SIRETSKIY²

* Corresponding author

¹Department of Geography and Human Environment

Porter School of the Environment and Earth Sciences, Tel Aviv University

Ramat-Aviv, Tel-Aviv, 6997801, Israel

+972-36409896, alekseyo@tauex.tau.ac.il, bennya@tauex.tau.ac.il

²Department of Social and Economic Geography, Uppsala University

P.O. Box 513, 751 20 Uppsala, Sweden

+46-73-9384911, marina.toger@kultgeog.uu.se

³Department of Civil Engineering and Energy Technology, Oslo Metropolitan University

P.O. Box 4 St. Olavs plass, N-0130, Oslo, Norway

+46-70-1679105, john.osth@kultgeog.uu.se

Keywords: mobile phone positioning, call details record, Voronoi tessellation, GPS

1. Introduction

The Voronoi tessellation of the Mobile Phone Network (MPN) service area is an iconic starting point of numerous papers that exploit mobile phone data for studying human mobility. This tessellation was introduced by the MPN developers for presenting technical ideas behind the mobile telephony (Baccelli et al, 1997). However, the students of mobile phone data made an additional step and, explicitly or implicitly, assumed that the mobile device is served by the closest MPN antenna, and thus, at the moment of connection, the user is located within the Voronoi polygon of the closest cell tower (Calabrese et al., 2014; Wang et al., 2018; Huang et al., 2019).

Only recently this consensual assumption was questioned (Ricciato et al., 2017; Bachir et al., 2019; Tennekes et al., 2020; Ogulenko et al., 2021). Based on the monthly distributions of distance to devices, Ogulenko et al. (2021) demonstrate that, on average, the size of the area that comprises 95% of monthly calls of tower's antennas is twice as large as the tower's Voronoi polygon.

In this study, we directly investigate how far may the mobile device be from the cell tower at the moment of connection. For this purpose, we analyze a series of the consenting volunteers' GPS locations, aligned in time with the Call Detail Records (CDR) from the mobile operator.

2. Experimental data

The GPS dataset we used in this study describes the mobility of two volunteers on occasional days during the years 2018–2021, 130 days in total, with the majority of GPS locations (68%) recorded in Uppsala, Sweden. Based on the exact time of connection, 90% of the GPS data were matched to the CDR records obtained from the secure MIND database established for research purposes by one of the largest MPN operators in Sweden. Overall, the volunteers' dataset contains data on connections to 1124 antennas belonging to 480 cell towers. The CDR data are aggregated over the 5-minute intervals. The CDR consists of an anonymized user ID, 5-minute timestamp, IDs of antennas used for the connection during these 5 minutes. We distinguish between the cases when during the 5-minute interval a device was connected to a Single Cell Tower (SC) or switched between Multiple Cell Towers (MC).

3. Comparison of the CDR- and GPS-based device locations

We estimated the distance between GPS location and the cell tower in two ways - based on the VPs ring neighborhoods and the Euclidean distance (Figure 1). For the SC case, we can determine the closest to and the farthest to the cell tower GPS location during the 5-minute interval. In the MC case the exact time of connection inside the 5-minute interval is unknown, and we can consider the

minimal distance only. That is, in the MC case, the distance between the GPS location and the cell tower is underestimated.

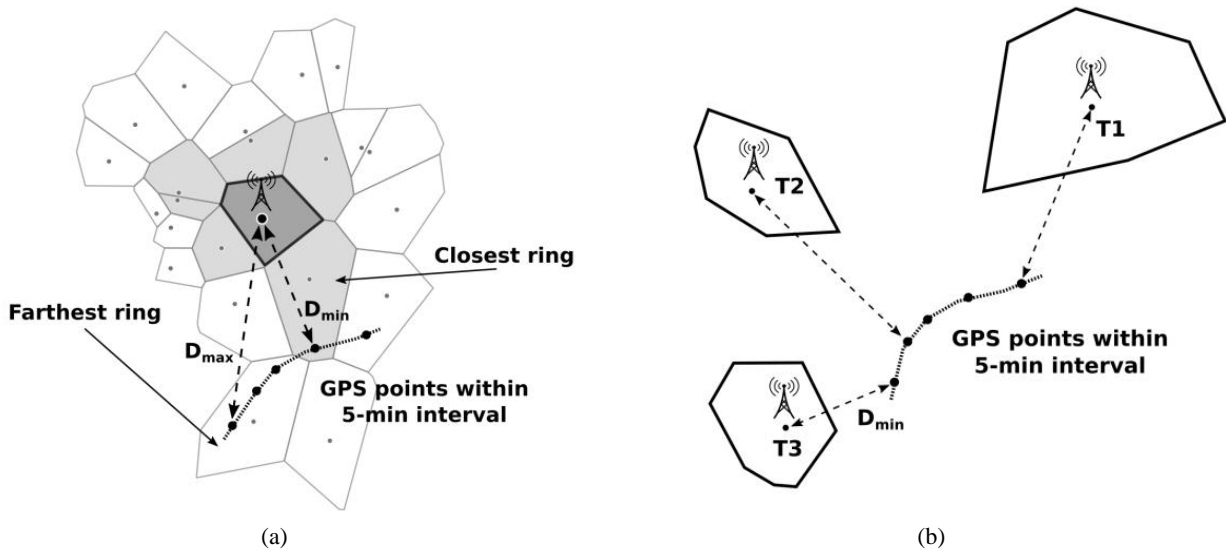


Figure 1. The rings of neighborhoods of the 1st and 2nd order around the cell tower’s Voronoi polygon and the estimates of the distance between the mobile device and cell tower(s) that served the device during the 5-minute interval. (a) SC case; (b) MC case.

Table 1 presents the statistics and Figure 2 the distribution of the minimal and maximal possible distance between the GPS location of the mobile device during a 5-minute interval and the cell tower(s) that served this connection. In the MC case, the average distance (Table 1) as well as the share of distant locations (Figure 2a, 2b vs Figure 2c, 2d) is higher than in the SC case, even for the maximal distance. The latter is easily explainable: For the device that is far away from all surrounding towers the possibility to switch between them is higher.

Number of cell towers that served the connections during 5-min interval	Average minimal distance		Average maximal distance	
	In meters	In rings	In meters	In rings
SC	2490 m (SD = 3480 m)	0.73 (SD = 1.11)	3670 m (SD = 4830 m)	1.12 (SD = 1.36)
MC	3970 m (SD = 4190 m)	1.23 (SD = 1.24)	-	-
SC + MC	3540 m (SD = 4050 m)	1.00 (STD = 1.23)		

Table 1. Minimal and maximal possible distance between the GPS location of the mobile device during a 5-minute connection and the cell tower(s) that served this connection.

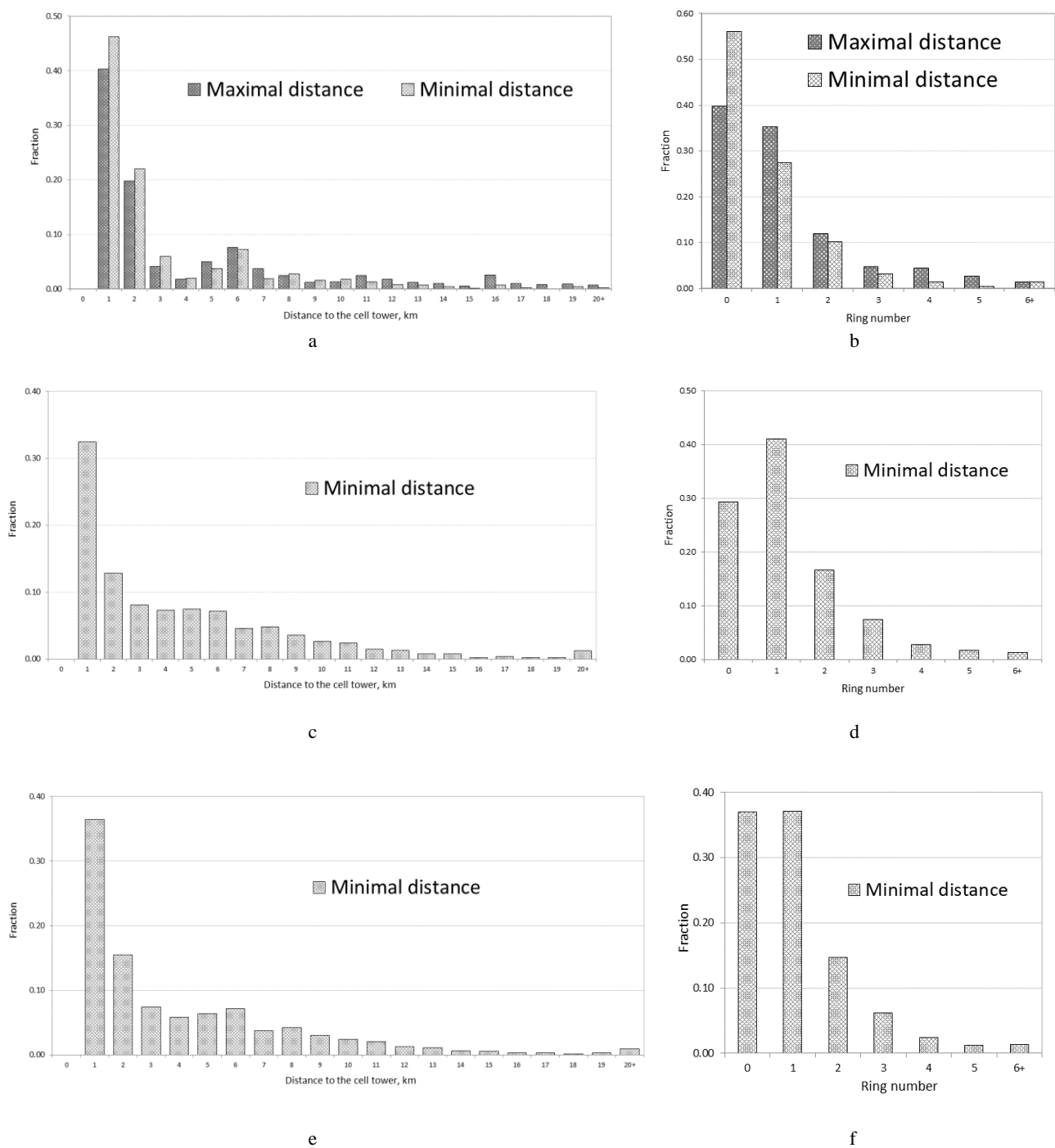


Figure 2. Minimal and maximal distances between the GPS location of the mobile device during a 5-minute connection and the cell tower that served this connection: (a) in kilometers for SC case; (b) in VP rings for SC case; (c) in kilometers for MC case; (d) in VP rings for MC case; (e) in kilometers for the entire dataset; (f) in VP rings for the entire dataset.

Figures 2e, 2f aggregates our main results: In 63% of all cases, the device was located beyond the VP of the serving tower. In 37% of cases the device was situated within the first ring; in 25% of cases

beyond the first ring. To comprise 90% of the possible locations of the device one has to consider the tower's VP together with its 1st and 2nd rings. 95% area should include the 3^d ring (Figure 2f).

We thus claim that the Voronoi-based view of the MPN service is wrong and must be revoked.

For connections established by volunteers when they were at work are excessive and we analyzed separately. These connections were served by three cell towers: VP of one of them covers the work building, while two others belong to its 1st ring of this VP and are not the closest ones (Figure 3).

Analysis of volunteers' connections at home and work stress two facts: (1) Antennas of several cell towers can serve connections from the same location; (2) Connections from a certain location are not necessarily served by the closest antennas.

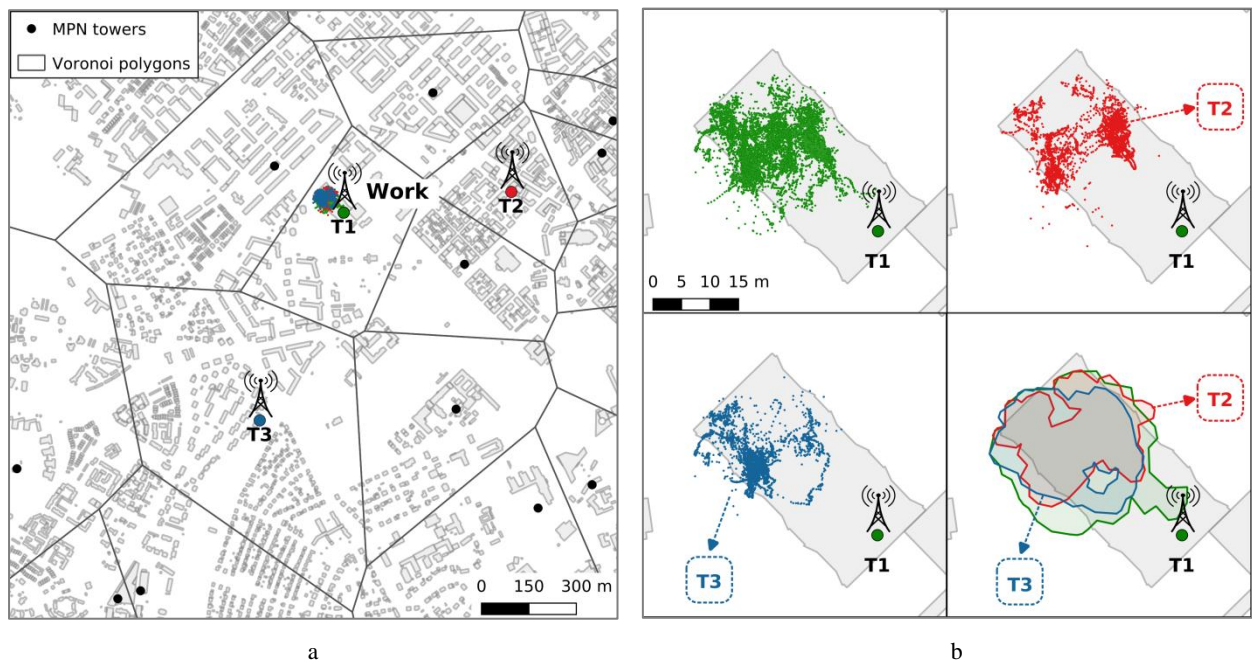


Figure 3. Volunteers' GPS locations at "Work" that can be related to their connections. (a) Cell towers and their VPs of the orders 0 and 1 around the "Work" location, T1, T2, and T3 denote antennas that served users' connections, black dot denote cell towers that didn't serve the connections; (b) GPS sequences for the periods that match to the connections, colored by the tower of connection.

5. Discussion

Our study clearly demonstrates that possible area of the device location is substantially larger than the cell tower's VP. This area can be roughly approximated by the neighborhoods of the 2nd or 3rd

order of the servicing tower's VP. This uncertainty is an inherent property for the MPN and entails essential overlap between the cell towers' service areas. The account for this uncertainty demands a change of the entire paradigm of the mobile phone data use in mobility studies and location privacy.

Acknowledgments

The research of the Tel-Aviv team was funded by the Blavatnik Interdisciplinary Cyber Research Center at Tel Aviv University and Ministry of Science & Technology, Israel, and The Ministry of Science & Technology of the People's Republic of China, Grant No. 3-15741. The research of the Uppsala team was funded by the FORMAS project grant 2018-00255.

References

1. Baccelli, F., Klein, M., Lebourges, M., Zuyev, S. 1997, Stochastic geometry and architecture of communication networks, *Telecommunication Systems* 7, 209–227.
2. Bachir, D., Khodabandelou, G., Gauthier, V., El Yacoubi, M., Puchinger, J., 2019, Inferring dynamic origin-destination flows by transport mode using mobile phone data *Transportation Research Part C*, 101, 254- 275.
3. Calabrese, F., Ferrari, L., Blondel, V.D., 2014. Urban Sensing Using Mobile Phone Network Data: A Survey of Research. *ACM Comput. Surv.* 47, 25:1-25:20.
4. Huang, H., Cheng, Y., Weibel, R., 2019. Transport mode detection based on mobile phone network data: A systematic review. *Transportation Research Part C: Emerging Technologies* 101, 297–312.
5. Ogulenko, A., Benenson, I., Omer, I., Alon, B., 2021. Probabilistic positioning in mobile phone network and its consequences for the privacy of mobility data. *Computers, Environment and Urban Systems* 85, 101550. <https://doi.org/10.1016/j.compenvurbsys.2020.101550>
6. Ricciato, F., Widhalm, P., Pantisano, F., Craglia, M., 2017. Beyond the “single-operator, CDR-

only” paradigm: An interoperable framework for mobile phone network data analyses and population density estimation. *Pervasive Mob Comput* 35, 65–82.

7. Tennekes, M., Gootzen, Y., Shah, S.H., 2020. A Bayesian approach to location estimation of mobile devices from mobile network operator data.
8. Wang, Z., He, S.Y., Leung, Y., 2018. Applying mobile phone data to travel behavior research: A literature review. *Travel Behaviour and Society* 11, 141–155.

The Northern Cities of the United Kingdom: firm ownership links, productivity and interconnection

Natalia ZDANOWSKA^{1,2}; Robin MORPHET²

¹Luxembourg Institute of Socio-Economic Research

11, Porte des Sciences

L-4366 Esch/Alzette

natalia.zdanowska@liser.lu

²Centre For Advanced Spatial Analysis

University College London

90 Tottenham Court Road

W1T 4TJ London

Keywords: Decentralisation, Northern Powerhouse Strategy, cities, interactions, ownership networks of firms, productivity, United Kingdom

1. Introduction

Despite many decentralisation initiatives proposed by the UK government to devolve power to city regions (Sandford2020), sub-regional governance remains subordinate to a centralised model of decision-making. Notwithstanding the relative weakness of local governance in the Northern cities, there remains the necessity of local business to adapt to and exploit local conditions and advantages with or without central or local government support. The Northern cities are not alone in having lower productivity than their European peers but preeminently, London shows higher productivity (OECD 2020). The Northern Powerhouse Strategy (NPS) was launched in 2016 to create a more integrated economic unit with agglomeration, scale and multiplier effects sufficient to offer improved productivity and a counterweight to London. Productivity assessments tend to reflect conditions at a regional or sub regional scale whereas a more localised assessment would allow a more finely targeted policy response.

In this paper, we test the use of the Bureau Van Dijk ORBIS and AMADEUS databases, which define multi-level ownership linkages between firms to explore this question of local performance, and we analyse Manchester's interactions between 2010 and 2018 with Liverpool, Leeds and Newcastle. We compare it with the ONS NUTS1 and NUTS2 GFCF (Gross Fixed Capital Formation) estimates (2000-2016) to capture the evolution of productivity.

2. Literature Review

Local embeddedness has tended to conflate network and territorial dimensions in focussing on the nature of a firm's linkages to, and interactions with, other local actors, whether firms or cities (Clark et al2018).

Strengthening economic inter-urban ties is one of the NPS's objectives, although no clear assessment of this policy in these terms has been made since 2010. NESTA's report (NESTA2016) highlights only the insufficient boosting of links between the city and entrepreneurs and recommends expansion of international transport linkages to strengthen global business connections. This paper seeks to fill this gap in the literature by conducting a temporal assessment of the NPS within the system of cities framework where preferential interactions between certain cities can drive the emergence of national sub-systems (Pumain1997). The productivity effects of different forms of ownership have concentrated largely on the foreign direct investment of multinationals, where some skills improvement has been noted and may outweigh the productivity effects of R&D transfer and exploitation (Siler et al2003).

3. Methodology & case study overview

The study is exploratory as the datasets themselves have been used relatively infrequently to date and may be subject to error. We seek to examine the effect of subsidiary firm ownership on productivity by location and sector to determine the potential for more localised estimates of productivity. We also seek to identify polarisation and peripheralization effects (Kuhn2015; Benedek2015) as evidenced by interactions within the NP economy since 2010, by analysing data on the inter-urban ownership links of firms in Manchester and Northern England, using methods derived from multivariate statistical analysis and network analysis. Literature on ownership networks is generally very scarce, because of the lack of precise geo-located data at city level. An original aspect of this analysis is to construct ownership links between cities where companies are localised and to decompose their ownership links into capital control chains at several levels. The hypotheses that we seek to test are the following. Are changes in productivity reflected in changes in the size and extent of ownership links? Is there a discernible economic sub-system across the Northern cities as evidenced by inter-urban interaction between firms?

4. Data analysis

Data used:

- The population and boundaries of UK agglomerations in 2018 defined as 138 Functional Urban Areas with a common and harmonized definition of cities from the Global Human Settlement (GHS) database produced by the European Commission;
- The ownership links of capital between firms in the UK in 2010, 2013 and 2016 at city level from the ORBIS database and from the AMADEUS database for 2018, produced by the Bureau Van Dijk (BVD). The exact geolocation, turnover and activity codes of companies (both owners and owned) in all sectors are available;
- The three-level constructed subnetwork of ownership networks in the UK contains 2,312 firms and 1,562 ownership linkages. An index of intensity of the revenues generated by these links at city level has been calculated (Śleszyński2015). An aggregation into 21 first level NACE

classification of activities has been applied. Correspondence analysis was applied to characterise the economic ownership specialisation of cities;

- ONS NUTS1 and NUTS2 GFCF (Gross Fixed Capital Formation) estimates 2000—2016 to capture de evolution of productivity.

5. Discussion

The data offers limited support for the relationship between productivity and the levels of subsidiary ownership and this aspect will need further development.

The analysis firstly showed that the geographical scope of ownership of firms located in Manchester is mainly oriented towards firms in the city itself or in London. Local links towards other UK cities outside London represent only 18% of total links in 2016.

Secondly, the analysis of the evolution of these local links between Manchester, Liverpool, Leeds and Newcastle has revealed a lack of increase of interactions between since 2010. The latter were polarised on Liverpool, Manchester and Leeds leading to a peripheralization of Newcastle in 2016.

Thirdly, findings revealed that Liverpool and Manchester have switched from a traditional manufacturing base towards a specialisation in professional and scientific activities within the ownership networks in this same period. Despite the modernisation of the Manchester and Liverpool industrial structure, it would seem that the NPS has had little or no impact on economic interaction between Northern cities since 2010.

6. Conclusions

This paper suggests that the Northern cities studied do not yet represent an integrated sub-system as envisaged by the NPS, at least from the economic perspective of ownership links between firms and to this extent, the NP remains a spatial imaginary. This is unsurprising given central government ambivalence to the concept, the lack of decentralised powers and the paucity of central government investment.

However, the evolution of the economic specialisation of Liverpool and Manchester in the global value chain of production reveals that this process of creating a viable economy of the North is still ongoing. These findings are relevant for policy-makers who may consider that an initiative such as the NPS should be continued and reinforced in the future.

A future extension of this research could include transportation and commuting interactions between cities of the NPS, as they may well exhibit patterns of interaction that differ from those of the ownership networks. This research suggests that one may think differently about the NPS from a network perspective with a focus on inter-urban interactions rather than focussing on interactions within a city.

Acknowledgments

The authors acknowledge support from the Urban Dynamics Lab under the EPSRC Grant No. EP/M023583/1

References

- Benedek, J. and Moldovan, A. (2015). 'Economic convergence and polarisation: towards a multi-dimensional approach', *Hungarian Geographical Bulletin*, 64(3), pp. 187–203.
- Clark, J., Harrison, J., and Miguelez, E. (2018). 'Connecting cities, revitalizing regions: the centrality of cities to regional development', *Regional Studies*, 52(8), pp. 1025–1028.
- Kuhn, M. (2015). 'Peripheralization: Theoretical concepts explaining socio-spatial inequalities'. *European Planning Studies*, 23(2), pp. 367–378.
- NESTA (2016). *City initiatives for Technology, Innovation and Entrepreneurship, The Northern Powerhouse Analysis*. Available at: https://media.nesta.org.uk/documents/citie_northern_powerhouse-2016.pdf (Accessed; 12 February 2020).
- OECD (2020). *Enhancing Productivity in UK Core Cities: Connecting Local and Regional Growth*. OECD Urban Policy Reviews, Paris, OECD Publishing, Available at: <https://doi.org/10.1787/9ef55ff7-en> (Accessed; 8 September 2021).
- Pumain, D. (1997). 'Pour une théorie évolutive des villes', *L'Espace géographique*, 26(2), pp. 119–134.
- Sandford, M. (2020). *Devolution to local government in England*. House of Commons Library, (07029), Available at: <https://researchbriefings.files.parliament.uk/documents/SN07029/SN07029.pdf> (Accessed; 8 November 2020).
- Siler, P., Wang, C., & Liu, X. (2003). 'Technology transfer within multinational firms and its impact on the productivity of Scottish subsidiaries', *Regional studies*, 37(1), pp. 15–25.
- Śleszyński, P. (2015). 'Economic control functions in Poland in 2013', *Geographia Polonica*, 88(4), pp. 701–708.

“VKontakte” social network behavior patterns: case study of Vologda region

Nikita SINITSYN¹

¹Lomonosov Moscow State University

1, Leninskie Gory, Moscow, 119234, Russia

+7 925 704 7066, nicksinus@yandex.ru (correspondent author)

Keywords: VKontakte, Vologda region, Russia, graph clustering, urban cultures

The study is based on the “VKontakte” open data. The personal data of users from Vologda region cities, whose age is between 15 and 70 years old, was collected with the help of “VKontakte” API. From the one side, it’s impossible to register in “VKontakte” for children. From the other side, older generations are not active in social networks. Several filters were developed to exclude fake users. The first is a filter of two weeks. Many fakes are created only for certain project. They become inactive after the finish of project. So, we can exclude such fake users by the time of the last visit. The second is a filter by the number of subscribers. Fakes usually have huge amounts of subscribers. The third is a filter by the content of status. Fakes often use status for advertising. The last step of data collection was to get subscriptions of users with open pages (table 1)

Table 1. Data volume at different data collection steps.

Data collection step	Numbers of people or users, ths people	Proportion, %
Citizens of the Vologda region aged 15-70	768	100
List of users by cities	616	72 %
Filtration from fakes	381	44 %
Collection subscriptions from open pages	277	32 %

A complex users’ characteristics was constructed. It reflects its tastes and interests according to his or her subscribes for social network communities (publics). Topics of “VKontakte” publics cover significant part of human’s everyday life. Most important trends of Russian society discourse are represented there. According to the studies of Levada-center (Volkov, Goncharov, 2019), “VKontakte” is still the most popular social net in Russia (fig. 1).

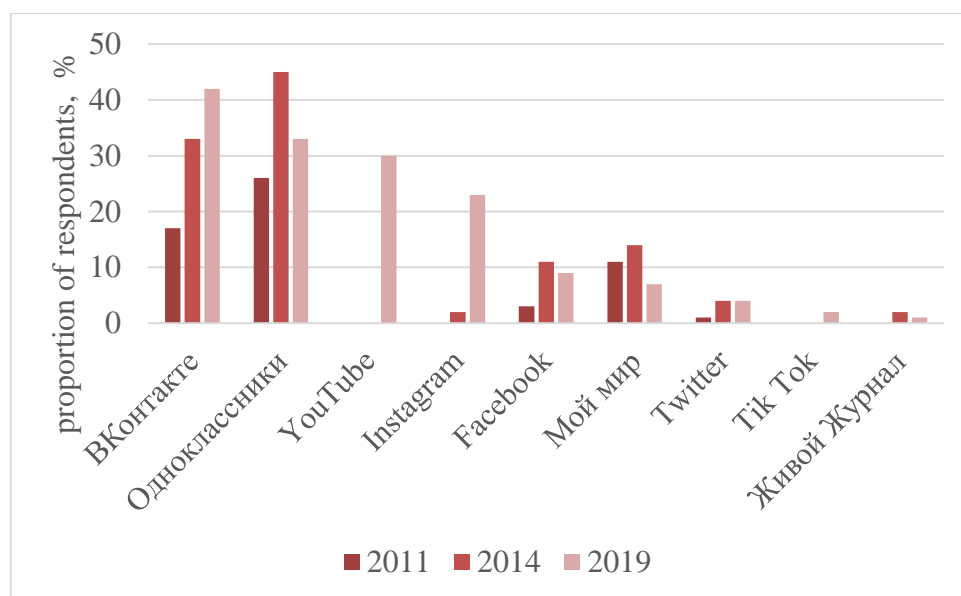


Figure 1. Proportion of respondents, which use different social networks in Russia

A group of users, which tends to subscribe to a certain set of communities is called a pattern of social network behavior. The patterns were defined using the developed method of graph clustering which is based on the force layout of graph (Martin et. al., 2011). Cutting long edges in the OpenOrd algorithm allows to express multiple identity of people in modern society. Eleven obtained patterns of social network behavior were divided in 2 groups: age-sex and thematic (table 2).

Table 2. Portraits of social network behavior patterns

Name of pattern	Average age, years	Male/Female ratio, в %	Main topics of the most popular publics	Category of cities' population
Women	30-40	20:80	Women's humor, housekeeping, cooking, healthy lifestyle, relationships,	big
Pensioners	50-60	10:90	Women's humor, cooking, pensioners, housekeeping, dacha, news, health,	middle
Men	30-40	90:10	Men's humor, relationships	everywhere is equal
Pupils	15-18	50:50	Teen's humour	small

Young women	17-21	10:90	Young women's humor, pop culture, beauty, aesthetics, astrology.	big
Guys	18-25	80:20	Youth's humor, football, news, moralizing publics.	big
Creative class	30-40	80:20	Popular science, sports, news, music, design, programming	big
Gamers	20-30	90:10	Computer games, computers, cybersport	big
Anime fans	15-18	10:90	Anime, drawing, astrology, k-pop.	big
Footballers and rappers	17-25	90:10	Football, cynical humor, moralizing publics, rap	big
Street racers	15-25	90:10	Cars and music	it doesn't depend on cities' population

Communities of age-sex patterns have no common subject, they have large number of users, they contain a lot of humorous resources. Communities of thematic patterns have one or two common subjects, they are much less populated, they contain a few number of humorous resources. City's structure of age-sex patterns depends on its population. City's structure of thematic patterns is also affected by the composition of its economy. The diversity of city's social network behavior patterns is directly proportional to its population. The diversity is associated with the role of the services in the local economy for cities with a comparable population.

References

Volkov, D., Goncharov, S. (2019). *'Russian media-landscape 2019: tv, press, internet and social networks'*. Moscow: Levada-center

Martin, S., Brown, W., Klavans, R., Boyack, K. (2011). 'OpenOrd: An Open-Source Toolbox for Large Graph Layout, in Visualization and Data Analysis 2011', *Visualization and Data Analysis (VDA)*: 7868-06

Leveraging GeoIPs as a source of understanding and comparing infra-regional spaces

Robin CURA¹

¹UMR Géographie-cités, CNRS

5, cours des Humanités, Bâtiment de recherches Sud, 93322 Aubervilliers Cedex, France

robin.cura@parisgeo.cnrs.fr (corresponding author)

Keywords: GeoIP, big data, user-generated data, geolocation technologies, spatial analysis

1. The IP address, a pervasive spatial data source

Many scholars have seized the opportunity to study dynamical aspects of spatial structures and individuals' mobility patterns thanks to sensor-based location technologies. GPS data, geolocated social media data or volunteered geographic information have proved to be of great use to modern geographers. Yet among the different sources of geographical big data, a very pervasive one seems yet to be rarely used for geographical studies: the IP address in itself, which is the main identification source when any digital device communicates on the internet.

According to Kitchin (2014), "it is becoming increasingly more difficult to take part in daily life without leaving some trace of participation due to the mediating role of digital technologies, software and the use of indexical identifiers (Kitchin and Dodge 2011). [...] Even if a person uses an anonymous username on social media, their IP and MAC addresses are recorded. We thus routinely leave a trail of data in our wake, though we often have little control over its form, extent, or how it is used". Although IP addresses datasets are widely available, and even more when directly working with any digital service, geographers seem to avoid using this data source. It is remarkable that Kitchin (ibid.) does not mention IP addresses as one of the numerous "sources of big data" that he identifies. Nonetheless, through various "IP geolocation techniques" (Poese et al., 2011), IP addresses can be geolocated, resulting in geographical coordinates being attributed to each IP address. Many different commercial providers offer such geolocation techniques, and most internet services do use the IP-location of their customers at many steps: geo-filtering, geo-targeting, geographical market analysis etc. In the following we will refer to this set of techniques as GeoIP.

2. An erroneous reputation: GeoIP can be accurate

The scarcity of geographical (and wider social sciences) literature using GeoIPs is thus rather intriguing, and might be explained by the reputation of this data source : GeoIP would be highly unreliable and could only be used on large scales (Siwipersad et al., 2008; Poese et al., 2011; Dan et

al., 2016). For Poesse et al. (ibid.), “Geolocation databases can claim country-level accuracy, but certainly not city-level”.

This for sure is a problem for a wide use of GeoIPs, but there’s also a way to still be able to use such data as part of geographical analyses. Along with the estimated geolocation, GeoIPs providers include an accuracy measure. For example, the provider MaxMind (www.maxmind.com) associates to each IP address an estimated radius of precision: if the accuracy is estimated to be 5km, it means that the device might have connected in a radius of 5km around the given coordinates. Figure 1 shows that MaxMind estimated accuracy is quite coherent with a test ground-truth dataset of surveyed peoples, and that it’s highly dependent on factors like the type of connection used (a cellular data connection in France is most of the time not locatable apart from country-scale). Figure 2 illustrates an example of the communicated accuracy for a sample of IP addresses in different countries: this suggests that for most countries in this sample, over 40% of the IPs can be geolocated within a 10km radius, and over 50% within 20km, which remain smaller than a typical NUTS3 area. That means that among a big IP dataset, the logs can be filtered using this accuracy, resulting in a somehow more precise geographical dataset for fewer users, but in all cases, for a substantial amount of people given the easy-to-gather nature of the IP information, in a time where there’s “a computer in every thing” (Arribas-Bel and Reades, 2018).

3. How to handle GeoIPs: an example study

This talk will demonstrate that once geolocated, user-generated IP-data are an interesting source of data for spatial analysis, even at sub-national scales. I will show that the estimated accuracy that is provided by IP-geolocation services is an important information that may be used to filter the IP data. Building upon a case study (see below) I will show that IP addresses allow a consistent geolocation of a fair share of the users of a music streaming company. This will let me discuss how geolocated IP addresses, once properly filtered, can be of a great use to study common geographical issues, such as population dynamics or mobility patterns.

Case study

This research takes place in an interdisciplinary program (records.huma-num.fr) that aims to better understand content consumption on streaming platforms. Thanks to anonymized individuals’ listening history data provided by one of the major streaming services in France, we have access to large volumes (approx. 10^{6-7}) of daily logs that include the IP address, for a large number of users. Additionally, for thousands of volunteer users, we also have access to self-declared user data, including the city of residence. In this talk, I will show how such IP logs can provide additional

information not only about the users themselves, but also about the places they live in or frequent.

Relations between Maxmind GeoIP accuracy and user-declared residences locations for a sample of streams
 Sample of ~1000 french Deezer users that answered a survey in October 2020.

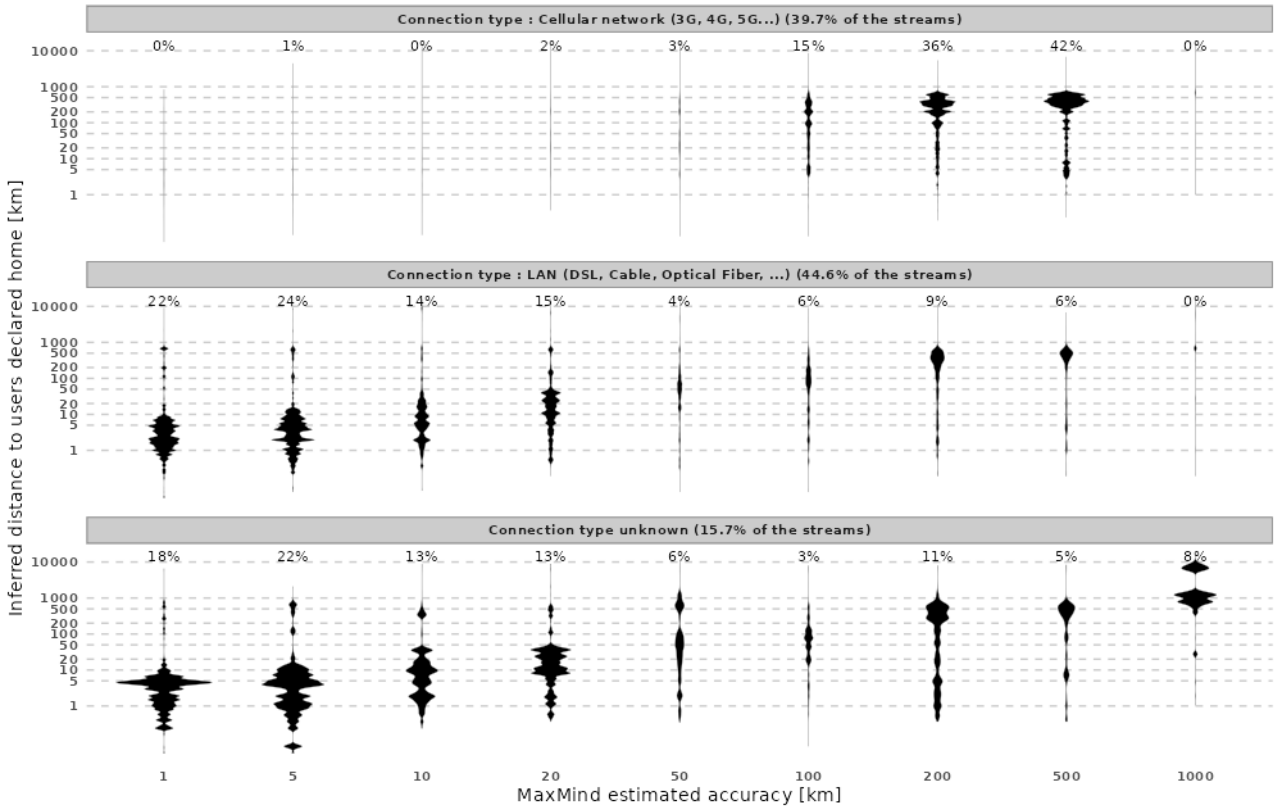


Figure 1: GeoIP accuracy according to the users' connection types. For a sample of surveyed people that declared their home location, we show the distance between this location and each of their streams IP locations (Y-axis), along with MaxMind estimated accuracy (X-axis) depending on the connection type.

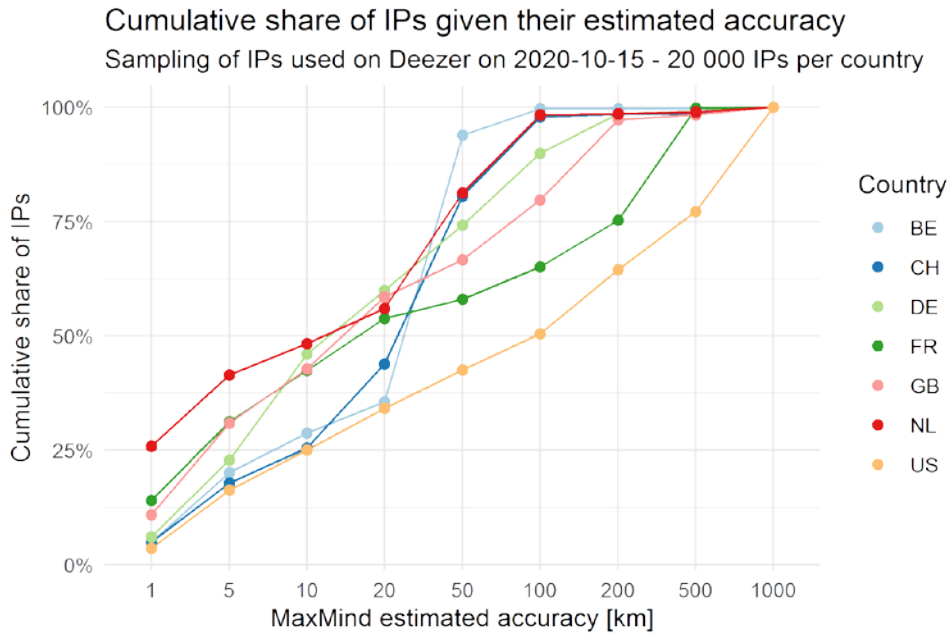


Figure 2: GeoIPs accuracy distribution.

Acknowledgments

This research work is being done within the RECORDS project (<https://records.huma-num.fr>). This project is funded by the French national research agency (ANR) under the reference ANR-19-CE38-0013.

References

- Arribas-Bel, D. and Reades, J. (2018) ‘Geography and computers: Past, present, and future.’ *Geography Compass*, 0(0) p. e12403.
- Dan, O., Parikh, V. and Davison, B. D. (2016) ‘Improving IP geolocation using query logs.’ *In Proceedings of the ninth ACM international conference on web search and data mining*, pp. 347–356.
- Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. 1 Oliver’s Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd.
- Poese, I., Uhlig, S., Kaafar, M. A., Donnet, B. and Gueye, B. (2011) ‘IP geolocation databases: unreliable?’ *ACM SIGCOMM Computer Communication Review*, 41(2) pp. 53–56.
- Siwipersad, S. S., Gueye, B. and Uhlig, S. (2008) ‘Assessing the Geographic Resolution of Exhaustive Tabulation for Geolocating Internet Hosts.’ *In Claypool, M. and Uhlig, S. (eds) Passive and Active Network Measurement*. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science), pp. 11–20.

The economic cost of urban sprawl approximated by the cost of maintaining road infrastructure in Monterrey

Roberto Ponce-Lopez^{1, 2}; D. Fabian Lozano-Garcia¹; Sebastian Fajardo¹; Alicia Landin¹;
Pedro Pacheco¹

¹ Tecnológico de Monterrey

Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849

dflozano@tec.mx (correspondent author)

² Institute for the Future of Education

Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849

Keywords: urban sprawl, cost of infrastructure, pavement and asphalt

Abstract

This study seeks to answer the question “is the current growth model of the Monterrey Metropolitan Area (MMA) economically sustainable?” This question arises from numerous discussions in different forums on the growth pattern of the MMA, presenting an expansion territorial 2.7 times in the period from 1990 to 2019; with a population growth of just under 2.5 million inhabitants in the same period; which resulted in a low-density territorial expansion (with a decrease of 32% for the period studied). To answer the question posed, this study focuses on the analysis of the costs of the replacement of roads in the MMA. To do this, a geographic database was built in which data was included on: Existing roads, urban properties and the presence of sidewalks. The processing of these data resulted in a little more than 195 million square meters of paved surface in the MMA in 2019.

The results show that the annual maintenance cost of the roads in the MMA is around 3.9 billion pesos. Comparing with the annual budget of the municipalities of the MMA, this amount represents 17% of the annual budget dedicated to public works.

1. Introduction

The Monterrey Metropolitan Area (MMA) is the second largest city in Mexico, spanning over 16 municipalities (Abasolo, Apodaca, Cadereyta, El Carmen, García, Gral. Escobedo, Gral. Zuazua, Guadalupe, Juárez, Monterrey, Pesquería, Salinas Victoria, San Nicolás de los Garza, San Pedro Garza García, Santa Catarina, and Santiago). Between 1990 and 2020, we observe a 39% drop in population density, where it went from 7,377 to 4,451 inhabitants per km². At the same time, the population of the Monterrey metropolitan area increased from 2.6 million in 1990 to 4.5 million in 2020. Therefore, the population increased 1.7 times, while the urbanized surface increased 2.8 times. The city is

consuming more land than before, and it is sprawling. The aim of this research is to quantify the cost of this pattern of land consumption and urban growth, by answering the question: Is the current growth model of the AMM economically sustainable?

2. Literature Review

There is a rich literature that focuses on measures and dimensions of urban sprawl (Galster *et al.*, 2001; Ewing and Hamidi, 2015), and on the effects of sprawl on social outcomes (Freeman, 2001; Ewing, 2008) and on public health (Zhao and Kaestner, 2010). There is also a large body of research on the costs of sprawl (Carruthers and Ulfarsson, 2003; Burchell *et al.*, 2005; Najafi *et al.*, 2006; Fregolent and Tonin, 2016). The problem is that the estimated costs of sprawl associates with economies of scale in the provision of public services and such costs tend to radically differ from place to place. Methodologically speaking, it is too complex to measure the cost of sprawl associated with each type of public service and several weak assumptions are made in the literature to quantify an overall cost. We rather select a single public service, the streets, and use a mixed methods approach to explore deeper into the cost of maintaining this infrastructure. We select the cost of streets among other services because it is a service that has a lifecycle of repair and maintenance, and it is inherently tied to sprawl.

3. Methodology

3.1 Quantifying the surface of streets

For this purpose, we proceeded to build a geographic database that included the center lines of existing roads, polygons representing parcels, and curve lines delimiting the sidewalks. The database was built using the "Road_Widths" tool within ArcMap. This tool defines the width of the roads, by means of lines perpendicular to the centerline that represents a road, up to the limits of a sidewalk or lot. The processing of this data resulted in the MMA having a little more than 195 million square meters of paved surface in 2019.

Additionally, roads were classified into three basic categories (based on their intensity of use). This phase of the analysis was carried out to differentiate the roads and to be able to make a more accurate estimate of the cost of maintaining them. The first step was to classify the roads into three basic categories according to their intensity of use. Two basic criteria were used for this purpose:

1. The type of road according to the category assigned by its nomenclature, i.e. Avenue, Road, Street, etc.

2. The estimated width of the roads, dividing them into three categories base on their width: primary, secondary, and tertiary

Three basic categories were used to categorize the road infrastructure:

- a) Primary. Roads identified as Avenues and Main, and/or whose width exceeded 44 m.
- b) Secondary. Roads identified as Avenues, Causeways, Ring Roads, Extension,
- c) Tertiary. Roads identified as Streets, Closed, Private, Pedestrian, etc. and/or whose width was less than 10m.

Additionally, the following classes were considered:

- d) Roads. Roads identified as such in their nomenclature, and which function as connectors of disjointed urban centers.
- e) Rural Roads. Roads that connect small settlements or individual buildings with population centers.
- f) Railroad. Railroad tracks within the metropolitan area.

3.2 Quantifying the lifecycle and cost of maintenance for streets.

Once the pavement surface were precisaly estimated by type of road, our team conducted 17 in-depth interviews with civil servants in charge of public infrastructure in the municipal and state government to learn about common practices and the costs related with road maintenance in the MMA. Four professionals from the two largest private construction companies in the area validated the collected information among the public servants. Three types of preventive maintenance processes were identified: a) Surface patching and crack sealing; b) Deep patching and crack sealing; and c) crack sealing for hydraulic concrete. In addition, the specialists mentioned ten pavement surface maintenance treatments (modified asphalt cutback, conventional asphalt cutback, modified asphalt layer replacement, conventional asphalt layer replacement, layer removal, recarpeting, seal watering, micro surfacing, cutback or milling, and hydraulic concrete rehabilitation). The specialists assigned a cost per m² to each type and frequency of the intervention, depending on the type of road: primary secondary and tertiary. Finally, we selected the most average type of maintenance, with it estimated cost, and with the average frequency by type of road to estimate a total cost of maintenance for the streets of the MMA.

3. Results and Discussion

Table 1 shows the results of the estimation of paved and sidewalks area per municipality, as well as data on total population (for the 2020 census).

Table 1. Paved areas and per capita estimates for the 16 municipalities.

Municipality	Total Population (2020)	Total Area (m ²)	Paved Area			Per Capita Paved Area			Percent of Urbanized Paved Area			Percent Total of paved surface
			Urbanized Area (m ²)	Streets Area (m ²)	Sidewalk Area (m ²)	Total Paved Area (m ²)	Per Capita Streets Area (m ²)	Per Capita Sidewalk Area (m ²)	Per Capita Paved area (m ²)	Percent of streets	Percent of Sidewalks	
Abasolo	2,974	46,863,758	1,583,965	200,523	43,924	244,447	67.4	14.8	82.2	12.70%	2.80%	15.40%
Apodaca	656,464	224,610,433	123,101,173	24,153,848	5,487,699	29,641,547	36.8	8.4	45.2	19.60%	4.50%	24.10%
Cadereyta	122,337	1,140,186,364	33,763,556	6,103,785	981,900	7,085,685	49.9	8	57.9	18.10%	2.90%	21.00%
El Carmen	104,478	104,235,138	10,724,499	3,389,339	170,208	3,559,547	32.4	1.6	34.1	31.60%	1.60%	33.20%
García	397,205	1,031,895,332	85,160,981	14,572,401	1,292,569	15,864,970	36.7	3.3	39.9	17.10%	1.50%	18.60%
General Escobedo	481,213	148,874,290	88,488,166	18,944,645	2,811,314	21,755,958	39.4	5.8	45.2	21.40%	3.20%	24.60%
General Zuazua	102,149	184,405,729	28,885,353	2,758,603	1,145,546	3,904,149	27	11.2	38.2	9.60%	4.00%	13.50%
Guadalupe	643,143	117,643,418	93,023,472	20,241,769	5,996,734	26,238,503	31.5	9.3	40.8	21.80%	6.40%	28.20%
Juárez	471,523	247,176,240	72,514,140	17,115,535	1,783,943	18,899,478	36.3	3.8	40.1	23.60%	2.50%	26.10%
Monterrey	1,142,994	324,724,084	186,424,677	43,698,195	13,429,591	57,127,786	38.2	11.7	50	23.40%	7.20%	30.60%
Pesquería	147,624	322,623,461	20,166,064	3,518,636	176,239	3,694,875	23.8	1.2	25	17.40%	0.90%	18.30%
Salinas Victoria	86,766	1,666,704,252	36,347,232	5,460,395	240,571	5,700,965	62.9	2.8	65.7	15.00%	0.70%	15.70%
San Nicolás de los Garza	412,199	60,146,050	58,716,302	13,422,913	4,744,959	18,167,872	32.6	11.5	44.1	22.90%	8.10%	30.90%
San Pedro Garza García	132,169	70,758,947	39,299,710	7,331,317	2,066,913	9,398,230	55.5	15.6	71.1	18.70%	5.30%	23.90%
Santa Catarina	306,322	915,570,986	55,783,646	9,824,923	2,027,077	11,852,000	32.1	6.6	38.7	17.60%	3.60%	21.20%
Santiago	46,784	738,921,156	39,135,435	5,089,234	1,145,631	6,234,866	108.8	24.5	133.3	13.00%	2.90%	15.90%
Totals	5,256,344	7,345,339,638	973,118,373	195,826,064	43,544,816	239,370,879	37.3	8.3	45.5	20.10%	4.50%	24.60%

As a result, the applied and ideal preventive maintenance cycles were identified, as well as the costs of such processes per square meter, varying from \$1,127 pesos (\$ 56.36 US Dls) to \$ 70 pesos (\$ 3.5 US Dls) per m². The total annual maintenance cost estimate ranges from \$ 2,447 million pesos (\$ 122 million US Dls) for primary roads; \$ 728 million pesos (\$ 36.4 million US Dls) for secondary roads; and \$ 817 million pesos (\$ 40.85 US Dls).

The total costs were validated against the public finance data of the ZMM municipalities. National Bureau Census' information was consulted regarding municipal expenditures by item, and it was calculated that municipalities exercised in 2019 a public works budget in the public domain of 3.8 billion pesos (INEGI 2020). This expenditure item includes spending on roads (including pothole patching) and public space. Our estimate is slightly higher than the spending reported by the Census Office. The explanation for these variations may be twofold: 1) in reality, maintenance of primary roads does not always occur with a periodicity of five years, but rather six or seven; 2) some primary roads have stretches of federal or state highways, whose maintenance is the responsibility of the federal and state governments. However, the proximity between both numbers is an indicator of the validity of the maintenance scenarios we are assuming for this study.

We can generally state that the annual cost of pavement replacement in the ZMM is 3.9 billion Mexican pesos in 2019. To dimension the magnitude of this expense, it represents 17% of the municipal

expenditures of the ZMM. In other words, the calculated scenario indicates that municipalities will have to dedicate an average of 17% of their annual budget only to maintain roads according to the base scenario stipulated by the experts.

Acknowledgments

This research was funded by Fundación FEMSA.

References

- Burchell, R. *et al.* (2005) *Sprawl costs: Economic impacts of unchecked development*. Island Press.
- Carruthers, J. I. and Ulfarsson, G. F. (2003) 'Urban sprawl and the cost of public services', *Environment and Planning B: Planning and Design*, 30(4), pp. 203–522. doi: 10.1068/b12847.
- Ewing, R. H. (2008) 'Characteristics, causes, and effects of sprawl: A literature review', *Urban ecology*, pp. 519–535.
- Ewing, R. and Hamidi, S. (2015) 'Compactness versus Sprawl: A Review of Recent Evidence from the United States', *Journal of Planning Literature*, 30(4), pp. 413–432. doi: 10.1177/0885412215595439.
- Freeman, L. (2001) 'The effects of sprawl on neighborhood social ties: An explanatory analysis', *Journal of the American Planning Association*, 67(1), pp. 69–77.
- Fregolent, L. and Tonin, S. (2016) 'Local Public Spending and Urban Sprawl: Analysis of This Relationship in the Veneto Region of Italy', *Journal of Urban Planning and Development*, 142(3), p. 05016001. doi: 10.1061/(ASCE)UP.1943-5444.0000318.
- Galster, G. *et al.* (2001) 'Wrestling sprawl to the ground: defining and measuring an elusive concept', *Housing policy debate*, 12(4), pp. 681–717.
- Najafi, M. *et al.* (2006) 'The Fiscal Impacts of Alternative Single Family Housing Densities : Infrastructure Costs The Fiscal Impacts of Alternative Single Family Housing Densities : Infrastructure Costs', (May).

Zhao, Z. and Kaestner, R. (2010) 'Effects of urban sprawl on obesity', *Journal of Health Economics*, 29(6), pp. 779–787.

Parallel Session PSB4

4 November 9:00-11:00 am (GMT)

Transport and Urban Form

Estimating individual journeys from collective transfers at bus stops

Romain LOUP¹; Guillaume GUEX²; François BAVAUD^{1,2}

¹Institute of Geography and Sustainability, Faculty of Geosciences and Environment, University of Lausanne, Switzerland. Corresponding author: +41 21 692 43 76, romain.loup@unil.ch

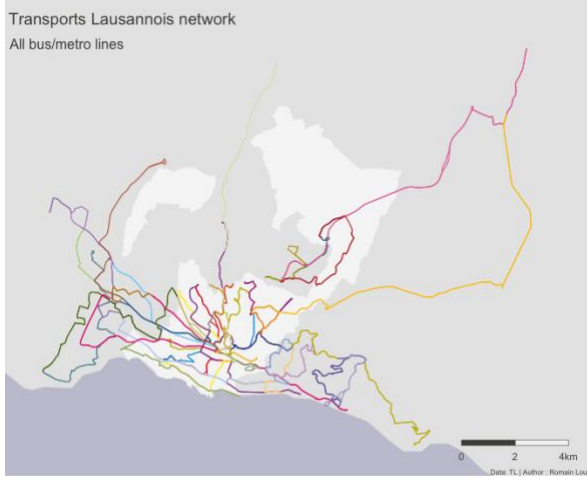
² Department of Language and Information Sciences, Faculty of Arts, University of Lausanne, Switzerland

Keywords: journey estimation, transportation data, Markov chains, directed gravity model

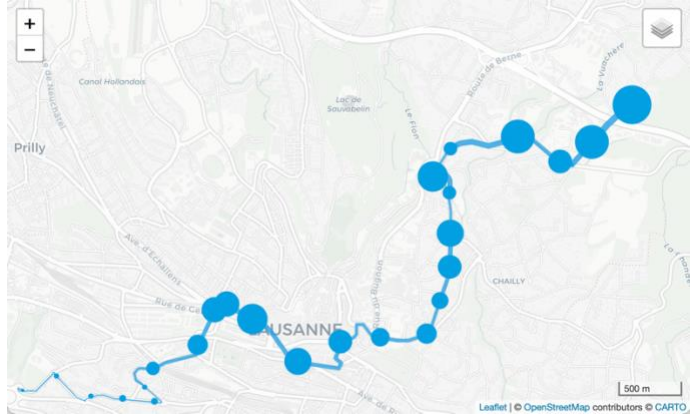
Analysing a transportation network and predicting its flows is a recurring issue in the field of mobility. Digitalisation of society and the availability of large datasets offer renewed possibilities in the statistical analysis of transportation. The aim of this communication is to present a model that predicts individual journeys and the associated origin-destination probabilities of a public transport user knowing only the aggregated collective transfer (ingoing and outgoing) intensities for each bus.

As a case study, this paper uses data from the *Transports Lausannois* (TL), the operator of the public transportation system of the city of Lausanne in Switzerland. The dataset includes 48 lines of buses (or subways) and more than a thousand stops. Each data row refers to a bus stop i characterised by the number of passengers getting in (x_i) and out (y_i), on bus line L and direction (back and forth) for a given date and time.

The complete data (all in- and out-transfers at each bus stop for each day) represent over than 37 million rows for a year. Moreover, they are spatially located. Stops and segments between stops are mappable by using shapefiles specifying the routes and the stops of the different bus lines. We consider the year 2019 only to avoid the loss of attendance due to COVID pandemic. In our first analysis, data is aggregated over the year yielding a single entering count x_i and a single leaving count y_i for each stop of each line in a given orientation. Map 1 displays the TL network comprising 48 lines, where the white area represents the municipal territory of Lausanne.



Map 1: Transports Lausannois network



Map 2: Line n°6, node size represent the aggregated entering counts and segment size the aggregated attendance.

We first perform an analysis of a single line among the complete TL network. Current work addresses the commuting issue as a multiline approach within the whole network.

1. Single oriented line

In the first part of this work, we consider a single oriented ($O =$ direct or return trip) bus line $L = 1, \dots, m$, comprising $i = 1, \dots, l$ ordered stops (nodes) that are connected by segments (edges). Let x_i and y_i be, respectively, the number of passengers getting in and out at station i . The quantity $n_{i,i+1}$ denotes the number of passengers carried from i to $i + 1$. The goal is to estimate N_{ij} , the number of passengers going from i to j . By construction, $N_{ij} = 0$ for $j \leq i$, $y_1 = 0$ and $x_l = 0$.

Our approach for solving this problem is probabilistic and makes a minimum assumption about passengers' behaviour: after travelling at least one segment, all passengers "forget" how many stops they already made and have the same probabilities of going out of the bus. This is typically modelled by a Markovian framework with expresses as follows.

Denote by p_i^{in} and p_i^{out} the probabilities for a passenger to get in, respectively to get out at a bus station. They can be estimated by

$$p_i^{\text{in}} = \frac{x_i}{x_{\bullet}} \quad \text{where } x_{\bullet} := \sum_{i=1}^{l-1} x_i$$

$$p_i^{\text{out}} = \frac{y_i}{n_{i-1,i}} \quad \text{where } n_{i-1,i} := \sum_{1 \leq k \leq (i-1)} (x_k - y_k)$$

Applying an absorbing Markov chain modelling (figure 1) with transition matrix $W = (w_{ij})$ (with non zero components for $i < j$ only) yield joint probabilities f_{ij} (the probability for a passenger to get in a bus at the stop i and get out at the stop j) of the form

$$f_{ij} := p_i^{\text{in}} z_{i+1,j} p_j^{\text{out}}$$

where

$$Z := (I - W)^{-1}$$

is the fundamental matrix of the Markov chain, and yields an estimated “directed gravity flow” of the form $N_{ij} = N a_i b_j I(i < j)$, where $I(i < j) = 1$ if $i < j$, and $I(j > i) = 0$ otherwise. The pull parameter a_i and the push parameter b_j can be estimated at once, unless some segments transport no passengers, in which case the Markov chain becomes reducible. The absence of distance-deterrence parameter can be justified by the shortness of trips. The possible unbalance resulting of the comparison of the estimated (daily aggregated) flows of the two orientations of a line can presumably be primarily explained by the altitudinal slope of the sections, often important in Lausanne.

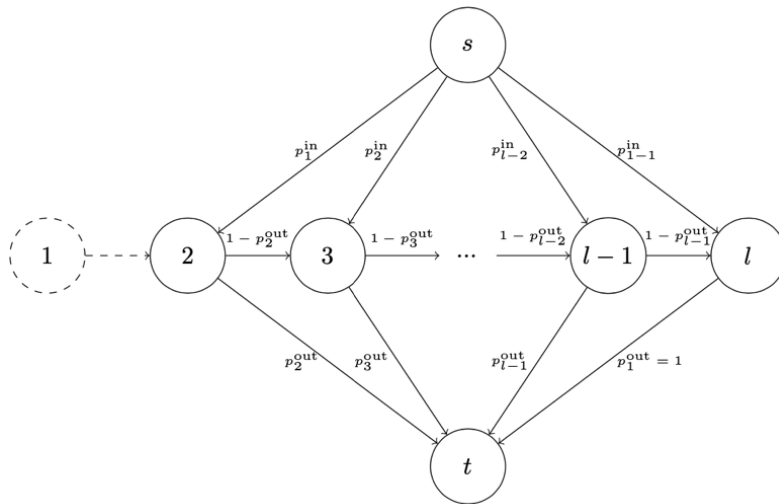


Figure 1: Absorbing Markov chain W , i.e., all transition probabilities from node t are null. Input probabilities are shifted to impose entering passengers to travel at least one segment.

As an example, map 2 displays an [interactive visualisation](#) of the x_i and the y_i for the line number 6 of the TL network¹. From this data, we apply the Markov model on this bus line that comprises 25 stops for the going trip and 26 stops for the coming trip. This line starts on the south of the city, runs through the centre, and ends in the north of the city. It carries over 2.7 million people in 2019 in direction south to north and over 2.2 million people in the opposite direction. Thanks to this model, it

¹ The visualisation of the complete network is available at [this link](#)

is possible to find the probabilities that a user of the lane $n^{\circ}6$ goes from stop i to stop j . For example, the most likely origin-destination pair f_{ij} is from $i = \text{“Sallaz”}$ to $j = \text{“Praz-Séchaud”}$ ($f_{ij} = 5.5\%$) for the going trip and the opposite with the coming trip ($f_{ji} = 6.1\%$). “Sallaz” is a transportation hub and “Praz-Séchaud” a densely populated area. In addition, the model shows an expected user trip of approximately 5 stops, which corresponds to the length of the trip between the aforementioned stops.

2. Multiline approach

The ultimate aim of this work, which is still under progress, consists in estimating M_{st} , the number of passengers entering the public bus network at source s and leaving it definitely at target t , possibly by using different bus lines. The crucial quantities, the between-lines transfers at each stop i , can be estimated as proportional to the total number of shortest-paths trips (available from the associated GIS) with source, respectively target masses $M_{s\bullet}$ and $M_{\bullet t}$. Those margins can be re-evaluated in turn by the former scheme, yielding an iterative algorithm, hopefully converging towards sensible network flows, and allowing classical networks characterization, such as nodes and edges centralities. Further comparisons with the optimal transportation solution, materializing the fiction of identical, exchangeable travellers, seem promising.

Evaluation and prioritization of urban traffic bottlenecks

Efrat BLUMENFELD LIEBERTHAL¹; Nimrod SEROK¹; Shlomo HAVLIN²

¹Azrieli School of Architecture

The Yolanda and David Katz Faculty of the Arts

Tel Aviv University

Ramat Aviv, 6997801, Israel

972-3-6405556, efratbl@tauex.tau.ac.il

²Department of Physics

Faculty of Exact Sciences

Bar Ilan University

Ramat Gan, 5290002, Israel

972-3-5318436

Keywords: big data, urban traffic, traffic bottleneck

The increasing urbanization process in the last decades is accompanied by a significant increase in traffic congestion in cities worldwide. This results in people spending an enormous time on roads leading to significant money waste and air pollution. Here, we present a new methodology for identification, cost evaluation, and thus, prioritization of the congestion sources i.e., the jam bottlenecks. It extends existing methods as it is based on network analysis of the entire road network and can be applied to different traffic models. Our results show that the macro-stability, presented by scaling characteristics of the traffic bottlenecks, overshadows the existence of meso-dynamics, where the bottlenecks change their location in time and space. This means that to plan and manage traffic jams in different locations and at different times, it is required to implement a framework, as developed here, that tracks traffic and evaluates the relative effect of each evolving bottleneck on the entire road network.

1. Background

The increasing urbanization process in the 21st century is accompanied by the growing usage of vehicles, leading to a significant increase in traffic congestion in cities around the world (Helbing 2003; Batty 2008; Barthélemy 2011). The price of congestion is enormous (Currie & Walker 2011; Pishue 2017), and it is increasing significantly with urban population growth.

There exists extensive work in various disciplines aimed at reducing traffic congestion generally, and in urban areas in particular (Batty 2008; Youn, Gastner & Jeong 2008; Li et al. 2010). There are two main approaches to reducing congestion: (1) increasing road supply by developing traffic-management solutions and car-sharing systems, and (2) reducing the demand for car usage by implementing taxes or fees on road usage.

The objective of this work is thus to develop a new framework to identify and prioritize traffic bottlenecks, based on big data (retrieved in real-time). This methodology can be implemented in planning transportation systems and reduce urban traffic congestion, by combining both of the above approaches. I.e., increasing road supply by means of traffic planning and management systems and reducing the demand by means of a dynamic road-pricing tool.

Our present work extends (Li et al. 2015) who developed a method to identify traffic jams bottlenecks based on the percolation process while using big data, retrieved in real-time, of traffic speeds. We, however, propose to identify traffic bottlenecks based on (Ban, Chu & Benouar 2007) who suggested that if a bottleneck causes its upstream to be congested, it must be congested prior to it. Hence, for the definition of a bottleneck, time is as important as space.

1.1 Methodology: Identification of traffic bottlenecks

We developed an innovative methodology for the identification and prioritization of traffic jams bottlenecks based on big data retrieved in real-time. The prioritization of jams' bottlenecks is based on the analysis of the global urban road network, rather than on local or adjacent junctions (as commonly used in available technologies).

We converted three datasets of two urban areas (London, Tel Aviv, and the center of Tel Aviv – without the Ayalon Highway) to dynamic, directed traffic networks where each node represents a junction, and each link represents a street segment between two junctions. The direction of the link represents the allowed traffic direction, and its weight is the relative speed at each time (in comparison to the maximal measured speed). These cases were chosen as they represent cities of different scales, different public transportation systems, and different regulations that influence driving behavior. We collected from Google Directions API the duration between every two adjacent junctions, every 15 minutes, over a week.

For each street segment, we extracted from the data the free-flow speed and calculated the density and the flow, based on a generalized car-following model that bridges microscopic and macroscopic models (May 1990).

We construct for each time t a weighted network, where $W'_{ij}(t)$ is the cumulative time each link has been considered as jammed at t (figure 1) and used the following process to create tree-shaped clusters of jammed links:

1. We identified the street segments that have been congested for the longest time and define them as trunks (see the red segment in figure 1, which has been loaded for the 12 time units).
2. We connected each trunk to adjacent streets that have been jammed for shorter periods and refer to them as branches of the jammed tree (JT) (see for example the orange street in figure 1).
3. To avoid false identification of branches that have not been influenced by the trunk, we defined a parameter that reflects the maximal duration a jammed street is considered as the cause for the jam in its upstream. In our case, we set this duration as 2 time-units (that represent 30 minutes).
4. We continue until there are no more jammed streets that can be connected to the tree and go back to the first stage. Then, we look for streets that have been jammed for the longest but are not connected to any of the identified trees.



Figure 1: Clusters of JTs. (A) All the colored streets are part of one JT (B) Two JTs (represented by red and blue colors). The red JT does not include the street that has been loaded for 2 measurements, as the time gap between this street and its adjacent one is larger than the pre-defined threshold (see the upper green circle). The blue JT cannot be considered as part of the red JT, as the duration of its trunk is longer than that of its adjacent street in the red JT (see the lower green circle).

We calculate the cost (in human hours units) caused by each street segment $C_{ij}(t)$ in comparison to its cost at free-flow speed. Then, we calculate the cost for the entire JT as the sum of the cost of its branches, and the cumulative cost that represents the cost of a JT from the moment it was created to the current time.

We found, that although some universal power-laws distributions that appear daily, govern the macroscopic spatio-temporal behavior of traffic jams, there are also unique behaviors that indicate that local attributes affect traffic dynamics as the same traffic bottlenecks usually do not reappear on different days. In other words, the macro-stability, presented by the scaling characteristics of the traffic bottlenecks that represent the seeming regularity of traffic load both in time and space, overshadows the existence of meso-dynamics, where the bottlenecks that create these jams, change their location in time and space. This means that in order to manage traffic jams in different locations and at different times and determine priorities, there is a need to implement unique solutions that track traffic and evaluates the relative effect of each bottleneck on the entire road network. This method can assist in identifying and prioritizing the bottlenecks based on their cost (in this case - in human-hours units, but it can also be based on other characteristics such as transportation equity).

Acknowledgments

Financial support for this study was provided by a grant from the Center for Innovative Transportation for supporting this research

References

- Ban, X., Chu, L. & Benouar, H., 2007, "Bottleneck identification and calibration for corridor management planning," *Transportation Research Record*, (1999).
- Barthélemy, M., 2011, *Spatial networks*, *Physics Reports*, 499(1–3).
- Batty, M., 2008, *The size, scale, and shape of cities*, *Science*, 319(5864).
- Currie, J. & Walker, R., 2011, "Traffic congestion and infant health: Evidence from E-Zpass," *American Economic Journal: Applied Economics*, 3(1).
- Helbing, D., 2003, "A section-based queueing-theoretical traffic model for congestion and travel time analysis in networks," *Journal of Physics A: Mathematical and General*, 36(46).
- Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H.E. & Havlin, S., 2015, "Percolation transition in dynamical traffic network with evolving critical bottlenecks," *Proceedings of the National Academy of Sciences of the United States of America*, 112(3).

- Li, G., Reis, S.D.S., Moreira, A.A., Havlin, S., Stanley, H.E. & Andrade, J.S., 2010, "Towards design principles for optimal transport networks," *Physical Review Letters*, 104(1).
- May, A.D., 1990, *Traffic flow fundamentals*, Prentice-Hall, Inc., New Jersey.
- Pishue, B., 2017, "US Traffic Hot Spots: Measuring the Impact of Congestion in the United States."
- Youn, H., Gastner, M.T. & Jeong, H., 2008, "Price of anarchy in transportation networks: Efficiency and optimality control," *Physical Review Letters*, 101(12).

Post-Metro Mode Choice Changes: from Macro- to Micro-Analysis

Anna IBRAEVA¹; Gonçalo Homem de Almeida CORREIA²; Cecília SILVA³; António Pais ANTUNES¹

¹ CITTA, Department of Civil Engineering, University of Coimbra
R. Luis Reis dos Santos, Polo II, 3030-788, Portugal
ibraevanna@student.uc.pt

²Faculty of Civil Engineering and Geosciences, Department of Transport & Planning, TU Delft
P.O. Box 5048, 2600 GA Delft, the Netherlands
G.Correia@tudelft.nl

³CITTA, Faculty of Engineering, University of Porto
Rua Dr Roberto Frias, s/N, 4200-465 Porto, Portugal
ccsilva@fe.up.pt

Keywords: longitudinal, transit-oriented development, panel data, micro scale, mode choice

Transit-oriented development (TOD) is a manifold concept that intends to address current urban challenges like heavy traffic and associated pollution by providing dense, mixed-use, and lively neighbourhoods around transit stations. Neighbourhoods with dense street network, ease of access to transit and a variety of local retail and services are supposed to discourage the use of a private vehicle. However, due to the complexity of modern cities, many factors (like socio-demographics, built environment characteristics at the destination, etc.) may condition the success of TOD in reducing car usage. In the last ten years, accompanying the implementation of TOD projects in many countries around the globe, numerous attempts have been made to analyse and evaluate TOD's contribution to car use reduction (Ibraeva et al., 2020). Notwithstanding major progress made in this direction, several issues remain understudied.

More precisely, comprehensive studies at a micro-level (assuming by micro-level units of analysis smaller than the smallest administrative division units) are still rare. While there are numerous works developed on individual/household scale (Bardaka and Hersey, 2019; Chatman, 2013; Cervero and Radisch, 1996; Handy et al., 2005; Van Acker et al., 2007; Cervero and Day, 2008; Oлару and Curtis, 2015; Cao et al., 2007; Brown and Werner, 2008; Van de Coevering et al., 2016) or on a neighbourhood scale (Cervero and Gorham, 1995; Cervero and Kockelman, 1997; Cervero, 2007; Griffiths and Curtis, 2017; Nasri and Zhang, 2014), they are typically limited in terms of the sample size (being based on survey data) and/or geographically narrowed to specific pre-selected station areas. City or region-wide

micro-scale studies that would rely on comprehensive census data are seemingly unavailable yet expanding the area of analysis may provide stronger evidence and, owing to greater variability, disclose phenomena that may not manifest in a neighbourhood-specific study, while using census data allows to overcome the limitations of a reduced sample size.

In this study, we aim to further explore the link between travel behaviour and TOD, addressing some shortcomings of the aforementioned works: this is a detailed analysis of changes in the number of car trips for work/study at a macro- and micro-level (civil parish and census tract level, respectively) applied to a region that received a new metro service during a 10-year period of analysis (2001-2011). We examine and compare the effect of metro on the number of car trips per 1000 inhabitants first on a macro-, then micro-level of analysis, moving from initial parish-level evaluation to a more detailed census tract level. This approach allows to accurately estimate not only the direct effect of metro implementation but also its spillover effect. Besides, as station types vary depending on the surrounding environment, using section-level data allows us to compare the performance and spillover effects from different station types present in the metro network: TOD, TAD (transit-adjacent development) and P&R (Park&Ride).

1. Methodology

The main data source is the Census data for 2001 and 2011 provided by Statistics Portugal (INE). In between these years, the metro system was introduced. Seven municipalities served by Metro do Porto (Gondomar, Maia, Matosinhos, Porto, Póvoa de Varzim, Vila do Conde and Vila Nova de Gaia) form the region of the analysis. These municipalities are divided into parishes, the smallest administrative units in Portugal, which are further divided into census tracts. Information in both years was aggregated into the same units of analysis forming two panel datasets, one with 120 parishes and the other one with 1561 census tracts.

1.1 Modelling approach

To compare the results between different scales of analysis and highlight details that only a more refined dataset can detect, an essentially similar approach was used in both cases (fixed effects panel data model). While both models are based on the same equation, spatial structures at a macro and micro scale are naturally different (Figure 1) and need to be accommodated using different spatial specifications for the parish-level model and the census tract-level.

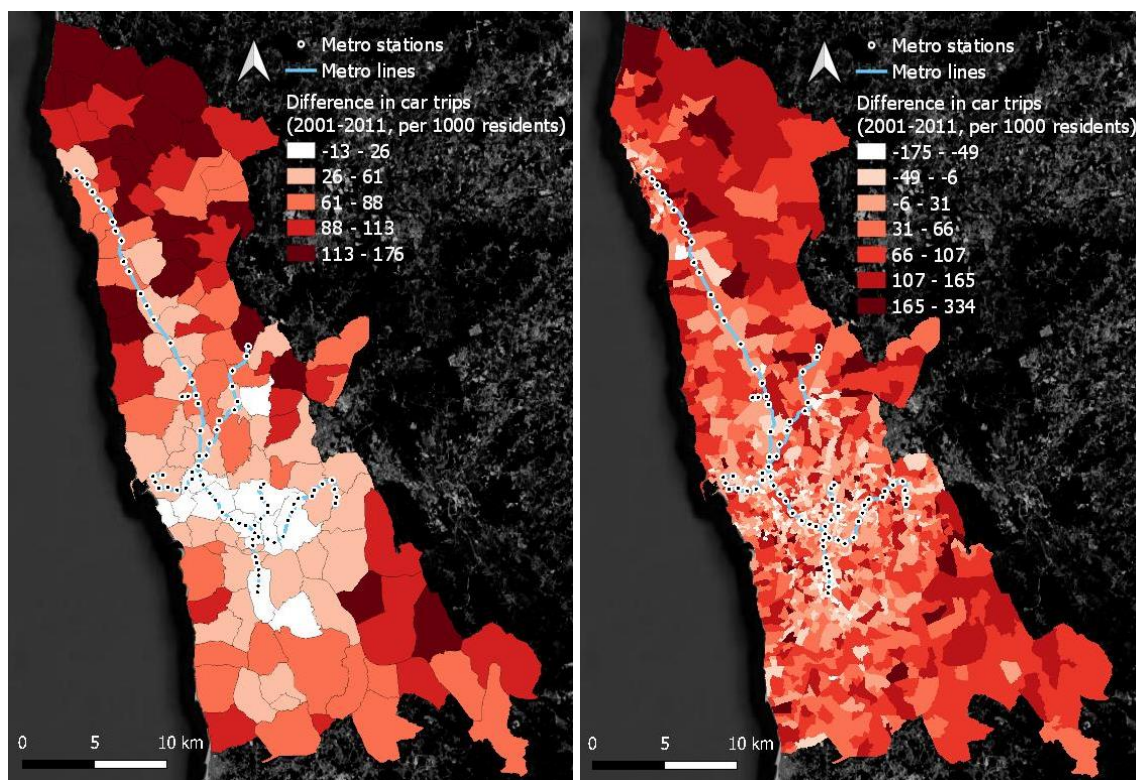


Figure 1: Difference in car trips at macro (parish) and micro (census tract) level in 2001-2011

Robust Lagrange Multiplier tests that allow detecting spatial error autocorrelation in the presence of a spatially lagged dependent variable and vice versa (Anselin et al., 1996) were performed to identify the source of spatial dependence. At the macro-level, RLM tests found appropriate the use of a spatial lag model, while at the micro-level specification with spatially correlated errors was the most appropriate.

1.2 Variables

Both macro- and micro-level models contain a set of control variables that account for socio-economic characteristics and built environment characteristics. However, metro-related explanatory variables vary depending on the scale of the model. For the aggregated macro-level model, a parish was considered metro-served if it was covered by a 400-meter buffer from a metro station. Since the spillover effect of metro was visible already on a macro-scale, first-order neighbours of the metro-served parishes were categorized as metro spillover parishes (“mp_spill”).

To analyse the spillover effect of metro at a micro-level, census tracts whose centroids are located within a 400-m buffer and a series of consecutive distance ranges (400m – 800m, 800m - 1200m, 1200m - 1600m, 1600m - 2000m) from the metro stations were identified and categorized accordingly. The distance for the last buffer (2 km) was selected to cover the first-order neighbours of the metro-served parishes thus ensuring coherence between the macro and micro-level models.

Finally, to evaluate the effect and the spillover magnitude for different station types, a qualitative analysis was performed to identify TOD, TAD and P&R stations. Network distance between metro stations and census tracts' centroids was calculated, finding the closest metro station for each census tract. Census tracts were then grouped into the categories that reflect the distance to the closest station and its type.

2. Results

2.1 Parish-level model

The results of the parish-level model are provided in Table 1. Judging by the high value of the R-squared, the model has quite strong explanatory power. The autoregressive term (λ) is statistically significant and positive. The effect of metro in the directly served parishes as well as metro spillover effect for the first-order neighbours of the directly served parishes is visible already on the macro scale. For the directly served parishes, metro implementation was associated with a decrease of 29 car trips per 1000 inhabitants. Metro spillover effect for the first-order neighbours of metro-served parishes was much weaker (though still significant): about 8 car trips less per 1000 inhabitants.

Table 1: Parish-level model estimation results

Variable	Estimate	Std.Error	t-value	p-value
lumix	-0.0038569	0.0243602	-0.1583	0.874197
up13	-0.0868063	0.1212061	-0.7162	0.473876
plus65	-0.5525618	0.1199828	-4.6053	4.118e-06 *
secedu	0.4362571	0.0652884	6.6820	2.357e-11 *
unemp	-0.0539721	0.0150676	-3.5820	0.000341 *
triptime	0.4329511	0.5354103	0.8086	0.418726
densapts	-0.0030280	0.0012422	-2.4375	0.014789 *
metro	-29.0637370	4.2593199	-6.8236	8.881e-12 *
mp_spill	-7.8723454	3.9306249	-2.0028	0.045196 *
λ	0.508751	0.058675	8.6707	< 2.2e-16 *
R ²	0.982			

Note: * $p < 0.05$

2.2 Census tract-level model (metro proximity)

The results of a census tract-level model that evaluates the effect of station proximity are presented in the Table 2. All variables reflecting metro station proximity were found to be significant and inversely related to the number of car trips, with decreasing influence on the dependent variable as distance increased. Being located within and up to 400-meter buffer from a metro station was associated with an average decrease of 43.5 in the number of car trips per 1000 inhabitants. This estimate would fall by about 10 car trips with each additional 400 meters until a 1200-meter limit from the station. Surprisingly, the estimates for census tracts located within 800 – 1600-meter buffer

are very similar (minus -23.5 and -21 trips, respectively). Finally, after 1600 meters the effect decreases quite sharply, falling to just -7 trips on average for the 1600 – 2000-meter buffer.

Besides distance, average daily train frequency on the closest station was significant, with each additional train being associated with an average decrease of -1.3 car trips per 1000 inhabitants.

Table 2: Census tract-level model estimation results

Variable	Estimate	Std.Error	t-value	p-value
lumix	-0.0060545	0.0043813	-1.3819	0.167001
lumix_lag	-0.0063069	0.0108327	-0.5822	0.560428
up13	0.3093290	0.0290926	10.6326	< 2.2e-16 *
plus65	-0.3318237	0.0247506	-13.4067	< 2.2e-16 *
secedu	0.5612359	0.0197284	28.4481	< 2.2e-16 *
unemp	0.1043231	0.0264995	3.9368	8.258e-05 *
triptime	0.7549798	0.2400946	3.1445	0.001664 *
densapts	-0.0036099	0.0011209	-3.2206	0.001279 *
densapts_lag	-0.0061943	0.0025846	-2.3966	0.016546 *
buff400	-43.5577626	2.5414572	-17.1389	< 2.2e-16 *
buff800	-31.5323893	2.4287144	-12.9832	< 2.2e-16 *
buff1200	-23.5689135	2.6972147	-8.7382	< 2.2e-16 *
buff1600	-21.0865937	2.9316362	-7.1928	6.349e-13 *
buff2000	-7.2745314	3.3007183	-2.2039	0.027530 *
trainhour	-1.2969749	0.1348745	-9.6162	< 2.2e-16 *
ρ	0.139225	0.028517	4.8822	0.000001049 *
R^2	0.963			

Note: * $p < 0.05$

2.3 Census tract-level model (TOD/TAD/PR proximity)

The results of the census tract-level model that accounts for different station types are presented in Table 3. The influence of TOD stations was considerably stronger than other station types, especially for census tracts located within 400-800 meters and 1200-2000 meters from a station. TOD's influence decreased with distance: from -52.2 car trips in the immediate station area (up to 400 meters) to -21.4 car trips for census tracts located within 1200 – 2000-meter distance, falling relatively steady on average by 10 units every 400 meters. TAD stations on the other hand showed very little variation in estimates for census tracts located within 400 – 1200-meter distance. Park&Ride was the only station type that did not show significant negative influence on the number of car trips in the immediate station areas (up to 400 meters from a station) and in the farthest areas (within 1200 – 2000 meters). The first issue probably occurred because for P&R immediate station areas were mainly scarcely populated census tracts with scattered private housing, where apparently station proximity did not change much people's habits. In the second case, maybe driving 1200 – 2000 meters to park a car (especially in the morning or evening peak hours) was already too far for the drivers, who in this case would prefer to drive directly to their destination. This could greatly

compromise P&R influence for the remote census tracts since P&R stations hardly had any other access options than car.

Table 3: Section-level model estimation results including station types

Variable	Estimate	Std.Error	t-value	p-value
lumix	-0.0060133	0.0043498	-1.3824	0.1668366
lumix_lag	-0.0054947	0.0107059	-0.5132	0.6077849
up13	0.3126143	0.0289178	10.8104	< 2.2e-16 *
plus65	-0.3347217	0.0246742	-13.5656	< 2.2e-16 *
secedu	0.5620046	0.0195939	28.6826	< 2.2e-16 *
unemp	0.1147700	0.0264694	4.3359	1.451e-05 *
triptime	1.0666852	0.2426099	4.3967	1.099e-05 *
densapts	-0.0029314	0.0011160	-2.6266	0.0086245 *
densapts_lag	-0.0053684	0.0025587	-2.0981	0.0358969 *
trainhour	-0.9431023	0.1427718	-6.6057	3.957e-11 *
TOD400	-52.2004192	2.9364575	-17.7767	< 2.2e-16 *
TAD400	-34.2333897	4.1281296	-8.2927	< 2.2e-16 *
PR400	-11.5242647	7.6726168	-1.5020	0.1330973
TOD800	-41.0154809	2.9967309	-13.6867	< 2.2e-16 *
TAD800	-20.9152552	3.6032867	-5.8045	6.456e-09 *
PR800	-21.1654416	5.6201780	-3.7660	0.0001659 *
TOD1200	-28.3715738	3.4836948	-8.1441	3.821e-16 *
TAD1200	-20.5843794	4.0690499	-5.0588	4.220e-07 *
PR1200	-15.9433585	5.7134844	-2.7905	0.0052630 *
TOD2000	-21.4575764	3.1243976	-6.8677	6.522e-12 *
TAD2000	-9.9354572	4.4929458	-2.2113	0.0270119 *
PR2000	-8.5700027	5.3097598	-1.6140	0.1065254
ρ	0.12392	0.02872	4.315	0.00001596 *
R^2	0.963			

Note: * $p < 0.05$

Acknowledgments

The research described in this paper has been carried in the framework of the doctoral thesis of Anna Ibraeva (PD/BD/135417/2017), funded by FCT – Fundação para a Ciência e Tecnologia.

References

- Anselin, L., Bera, A. K., Florax, R., Yoon, M. J. (1996). ‘Simple diagnostic tests for spatial dependence’, *Regional Science and Urban Economics*, 26, 77 – 104, DOI: 10.1016/0166-0462(95)02111-6.
- Bardaka, E., Hersey, J. (2019). ‘Comparing the travel behavior of affordable and market-rate housing residents in the transit-rich neighborhoods of Denver, CO’, *Travel Behaviour and Society*, 15, 74 – 87, DOI: 10.1016/j.tbs.2019.01.001.
- Brown, B. B., Werner, C.M. (2008). ‘Before and After a New Light Rail Stop: Resident Attitudes, Travel Behavior, and Obesity’, *Journal of the American Planning Association*, 75(1), 5-12, DOI: 10.1080/01944360802458013.

- Cao, X., Mokhtarian, P. L., Handy, S. L. (2007). 'Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach', *Transportation*, 34, 535 – 556, DOI: 10.1007/s11116-007-9132-x.
- Cervero, R. (2007). 'Transit-oriented development's ridership bonus: A product of self-selection and public policies', *Environment and Planning A*, 39(9), 2068-2085, DOI: 10.1068/a38377.
- Cervero, R., Day, J. (2008). 'Suburbanization and transit-oriented development in China', *Transport Policy*, 15(5), 315-323, DOI: 10.1016/j.tranpol.2008.12.011.
- Cervero, R., Gorham, R. (1995). 'Commuting in transit versus automobile neighborhoods', *Journal of the American Planning Association*, 61(2), 210-225, DOI: 10.1080/01944369508975634.
- Cervero, R., Kockelman, K. (1997). 'Travel demand and the 3Ds: Density, diversity, and design', *Transportation Research Part D: Transport and Environment*, 2(3), 199-219.
- Cervero, R., Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods, *Transport Policy*, 3 (3), 127 – 141, DOI: 10.1016/0967-070X(96)00016-9.
- Chatman, D. G. (2013). Does TOD need the T? On the importance of factors other than rail access, *Journal of the American Planning Association*, 79(1), 17-31, DOI: 10.1080/01944363.2013.791008.
- Croissant, Y., Millo, G. (2019). Panel Data Econometrics with R, *John Wiley & Sons Ltd*, Hoboken, NJ, USA.
- Griffiths, B., Curtis, C. (2017). Effectiveness of Transit Oriented Development in Reducing Car Use: Case Study of Subiaco, Western Australia, *Urban Policy and Research*, 35(4), 391-408, DOI: 10.1080/08111146.2017.1311855.
- Handy, S. L., Cao, X., Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California, *Transportation Research Part D: Transport and Environment*, 10(6), 427-444, DOI: 10.1016/j.trd.2005.05.002.
- Ibraeva, A., Correia, G. H. A., Silva, C., Antunes, A. P. (2020). Transit-oriented development: A review of research achievements and challenges, *Transportation Research Part A: Policy and Practice*, 132, 110-130, DOI: 10.1016/j.tra.2019.10.018.
- Nasri, A., Zhang, L. (2014). The analysis of transit-oriented development (TOD) in Washington, D.C. and Baltimore metropolitan areas, *Transport Policy*, 32, 172-179, DOI: 10.1016/j.tranpol.2013.12.009.
- Olaru, D., Curtis, C. (2015). Designing TOD precincts: accessibility and travel patterns, *European Journal of Transport and Infrastructure Research*, 15(1), 6 – 26.
- Van Acker, V., Witlox, F., Van Wee, B. (2007). The Effects of the Land Use System on Travel Behavior: A Structural Equation Modeling Approach, *Transportation Planning and Technology*, 30(4), 331-353, DOI: 10.1080/03081060701461675

Van de Coevering, P., Maat, K., Kroesen, M., Van Wee, B. (2016). Causal effects of built environment characteristics on travel behavior: a longitudinal approach, *The European Journal of Transport and Infrastructure Research*, 16(4), 674-697, DOI: 10.18757/ejtir.2016.16.4.3165.

Modelling urban population density of employment centres: Israel case study

David Burg^{1,2}

¹Tel Hai Academic College

Qiryat Shemona, 1220800, Israel

²Program for the Human Environment, Rockefeller University

New York, New York, 10065, USA

+972-50-7372797, biomodel.tac@gmail.com

Keywords: Population density, Employment centres, Dynamical modelling

Israel's population grew significantly during the 20th century and intensified in the early 21st century with predictions in the range of 12-25 million inhabitants by 2048 in this small geographic area by the Israeli National Economic Council (INEC). With the one of the highest densities in the world, it is clear that many challenges are posed to such as an economy and society in terms of energy, consumption, transportation, pollution, public health and more. As most people live near and commute to employment centers, 1-in-9 in Israel do so, it is important to study the development of these hubs of human and economic activity.

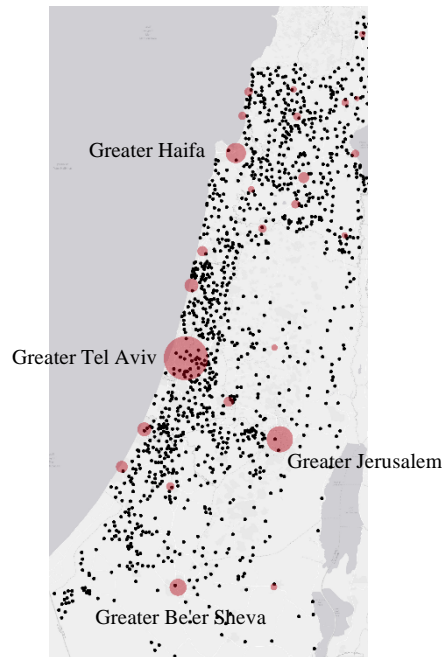
1. Urban Employment Centers

While administrative divisions of cities represents a statistical and scientific problem this work details the development of the population size and density on the scale of employment centres (Silva and Poelman 2016). Digitizing the map published to the public (CBS, 2013) recovered 28 centres with continuous built areas where 'Greater Tel Aviv', 'Greater Jerusalem', 'Greater Haifa' and 'Greater Be'er Sheva' stand out (Map 1). Extant data for population sizes between 1980-2020 was collated and aggregated across cities with consolidated bordering built areas. This was then fit to the multi-phase logistic model (Meyer and Ausubel, 1999):

$$P(t) = \sum_{i=1}^n \left(\frac{K_i}{1 + e^{-\frac{\ln(81)}{\Delta_i}(t-\tau_i)}} \right)$$

where P is the modelled population size, K_i is the upper asymptote and τ_i is the inflection point for n distinct growth curves. Δ_i is the "characteristic time" for the population in phase i to grow for 10-90%. Results show that all employment centers have developed in a manner consistent with two "growth

waves" (*i.e.*, $n=2$) which are now close to peak growth, or even showing a convergence to saturation ($d^2P/dt^2 < 0$). Figure 1 shows the bi-phasic trajectory of the four major employment centers.



Map 1: Israel's cities, municipalities, and communities (*black*) and major employment centres (*red*). The size of the employment centres is calculated as the consolidated population sizes.

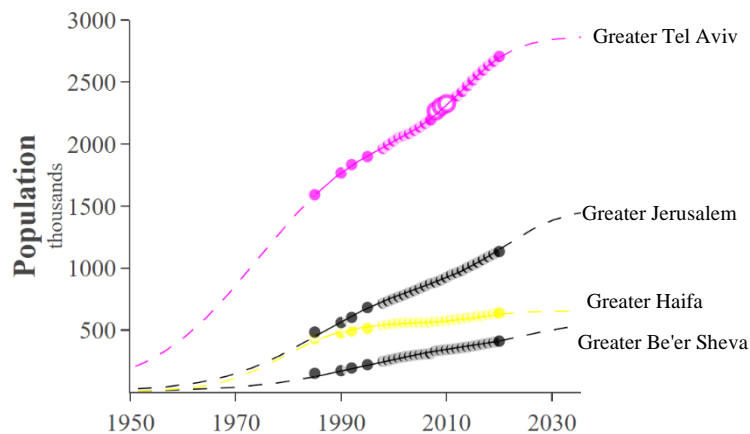


Figure 1: The population dynamics of the aggregated employment centers of Israel

2. Population Density

Israel has an average population density, customarily calculated as a function of the entire land mass. However, this description is inadequate as 45% of the land mass is inarable and over 90% of the population is urbanized. Here the urban built area (CBS 2020) is used to derive population density.

The 'Greater Tel Aviv' area is characterized by a population density of ~12 people/dunam, while in Jerusalem it is ~15 people/dunam, and in Haifa ~10 people/dunam. For comparison, New York City (five boroughs) has a density of 10.1 and Singapore with approximately 9 people/dunam built-up area but both with improved infrastructure relative to Israel.

3. Housing Building Rates

Further evidence of the intense development of the Greater Tel Aviv area is revealed in the trend of housing construction. This indicates the emphasis placed upon an area and the demand to be in proximity to it. During the last three decades Israel has created nearly 1 million (10^6) new homes, with the simultaneous necessary changes in land use. Concomitantly, the small Greater Tel Aviv area has received 50% of all new homes constructed (Figure 2). Due to compressed land resources, intense urban renewal programs and general demand for proximity to employment centers densification is a necessary requirement. However, continuation of this process, couple with the lack of land to implement transport infrastructure, will only increase economic, environmental and health issues, not alleviate them.

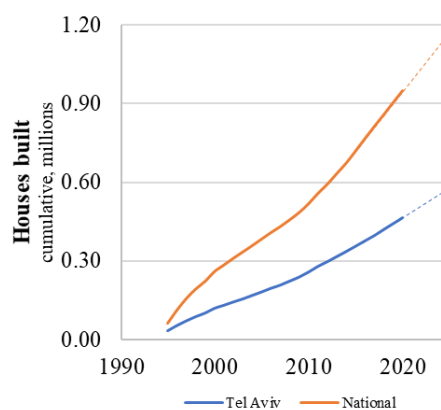


Figure 2: Housing construction trends for Israel (orange) and the Greater Tel Aviv area (blue). Linear projections are shown (dashed lines).

Population density should be re-evaluated in many contexts, as shown here. As of 2020 the 'Greater Tel Aviv' area consists of 2.7 million inhabitants, comprising 36% of Israel's Jewish population (Arab populations tend to associate in smaller communities). This is orders-of-magnitude larger than is common where the primal city is 5-10% of the nation's population. New York City and London with 6 and 13% of the national populace, respectively, and even there they face existential problems. Israel's

largest metropolitan area may be an important case study to research the impacts of policies to generate unsustainable densities with insufficient interventions to attempt to attenuate the resulting costs. Densification and overcrowding creates daily nuisances but this is expected to escalate into a national disaster for the environment, and even more impactful on public health in the coming decades. Economies-of-scale are not unbounded and costs are imposed when crowding impedes efficiency. Commuting and transportation to work hubs (Kaplan et al. 2020) are at historically low speeds (Ahmed and Stopher, 2017) and not projected to improve in the near future. The ramifications expected to face society will be discussed, as well as possible interventions. Despite governmental clearly observable "densifying policy" the current environmental and public health situation, as epitomised by COVID-19's high density dependence, may be a warning sign for decision-makers to re-examine the existing development policies and provide insights to create a sustainable trajectory for nature and human wellbeing.

Acknowledgments

The author discloses no funding for this study.

References

- Ahmed, A. and Stopher, P., 2014. Seventy minutes plus or minus 10—a review of travel time budget studies. *Transport Reviews*, 34(5), 607-625.
- CBS, 2013. Employed population aged 15 and over, by locality of residence and by employment centers. Retrieved from: CBS.gov.il.
- CBS, 2020. Population data. Retrieved from: CBS.gov.il
- Kaplan, N., Burg, D. and Omer, I, 2020. The spatial organization of accessibility and functional hierarchy: The case of Israel. *Computers, Environment and Urban Systems*, 80, 101429.
- Ismael, H.M. 2020. Urban form study: the sprawling city—review of methods of studying urban sprawl. *GeoJournal*, 1-12.
- Meyer, P.S. and Ausubel, J.H., 1999. Carrying capacity: a model with logistically varying limits. *Technological forecasting and social change*, 61(3), pp.209-214.
- Silva, F.B. and Poelman, H., 2016. Mapping population density in functional urban areas. JRC Technical Reports, European Commission. DOI:10.2791/06831.

Decision Making and Land Market Preferences: A Land Use Agent Based Model

Yahya Gamal¹; Nuno Pinto²

¹University of Manchester, SEED, Architecture Department & Cairo University, Faculty of Engineering,
Architecture Department

Arthur Lewis Building, Oxford Road, Manchester, M13 9PL, UK
+447533532763, yahya.gamal@manchester.ac.uk (correspondent author)

²University of Manchester, SEED, Planning and Environmental Management Department
Humanities Bridgeford Street, Oxford Road, Manchester M13 9PL, UK
United Kingdom+44 (0) 161 275 2729, nuno.pinto@manchester.ac.uk

Keywords: land market, procedural utility, agent based model

Land use change involves regulatory procedures that govern exchanging land as a commodity and dictating its land use. In that sense, the material benefit from land is only achievable through a set of procedural regulations. That is, land is the economic output and land markets are the procedural regulations governing the pricing and exchange of land.

In that sense, both outcomes (land) and procedures (land markets) should be considered during economic decision making (Sen, 1995). On the one hand, outcomes are explicitly described as an aggregation of the physical attributes of plots such as size and proximity to services. On the other hand, procedures are implicitly included in different forms within neoclassical economics. For instance, in calculating the cost of capital in a land market, procedures are considered in the form of risk premium values (Pratt and Grabowski, 2014). This includes the risks due to uncertainties in supply-demand trends within the land market as a whole due. It also includes unique risks for each different market regulation – these market regulations can coexist in the same context (Gurran and Bramley, 2017) and we label them as land sub-markets. To exemplify, an informal land sub-markets can entail high risk premium if the buyer views the regulatory procedures as uncertain. In other cases, procedures are included in the transaction price based on information availability (Botosan, 2006). The less the information is available within a land market for a buyer, the higher the risks and the perceived transaction cost.

Although the aforementioned approaches can be used to describe land market preferences, they abstract such preferences in monetary terms. This neglects the underlying psychological factors that motivate a buyer to prefer a specific market regulation. To address this limitation, we operationalise the concept of procedural utility to describe the psychological motivations in the context of land markets. We formulate a quantitative relation between procedural utility and the traditional (output) utility. Based on this, we observe market preferences in the case study of Cairo, Egypt, and we apply these observation in a land use change Agent Based Model (ABM) incorporating both output and procedural utility.

To clarify, traditional output utility is used to describe material benefit in economic decision making (Emil Kauder, 1965). Quantitative land use analyses, specifically simulation methods, use such utility as an aggregate representation of a set of land plot attributes – this aggregation is usually through a linear (e.g. Xie, Batty and Zhao, 2007) or a Cobb-Douglass formula (e.g. Filatova, Veen and Voinov, 2008). While such utility can include procedures in the form of capital risk or transaction prices, it inherits the aforementioned limitation of ignoring the psychological factors.

Procedural utility addresses such limitation properly. It is concerned with the value of an economic process rather than its output (Frey, 2008). It ties its achievement to the satisfaction of the three psychological innate needs: autonomy (being causal); relatedness (connected and a member of a larger group); competence (capable and

effective). These needs have been initially introduced within the Self-Determination Theory (SDT) (Ryan and Deci, 1982). More importantly, SDT ties the satisfaction of such needs to a set of observable motivations for being involved in a specific process (Ryan and Connell, 1989). In doing so, SDT introduces four motivational categories on a continuum scale starting from externalised motivations and ending with internalised ones: external motivation (achieving material benefit); introjected motivation (avoiding a sense of guilt); identified motivation (alignment with one's moral values); integrated motivation (enjoyment of the process) (Ryan and Connell, 1989). SDT states that each motivation is observable on a scale of 0 to 4 through a set of questionnaires that identify why an individual undergoes a specific process. The more internalised the motivations, the more satisfied the innate needs are (Ryan and Connell, 1989). Subsequently, this entails a higher achievement of procedural utility. In other words, motivations can be used as a quantitative indicator of procedural utility.

Accordingly, we introduce a mathematical representation to aggregate the four motivational categories into procedural utility. Further, we transform such procedural utility to align with the output utility through a factor (β). This factor corresponds to the procedural utility weight, and it is only observable through comparing land plots in different land sub-markets. Hence, we formulate a set of questionnaires including different cases of land plots and sub-markets to observe both output and procedural utility. We deploy such questionnaire in Greater Cairo for a sample of inhabitants living in formal housing units. We selected Cairo due to its rich context of different land market procedures including: private real estate (supply-demand competitive market); government market (government sets prices and first come first served buyers); social market (informal bargaining); and lottery market (random buyers allocation) (based on David Sims, 2010).

The initial findings indicate that each individual have different preferences towards the aforementioned four market procedures. Further, market preferences has a significant effect on individual land choices – the procedural utility weight β value has been as high as 0.8 for some participants. To identify the collected effect of such decisions on land use change, we formulated an ABM based on the mathematical formulation of procedural utility. We ran this model using inputs from the observation in Greater Cairo to test the effect of procedural utility on spatial segregation. The results show that market preferences lead to the formulation of spatial clusters of agents preferring the same market despite the random process of searching for land plots (Figure 1).

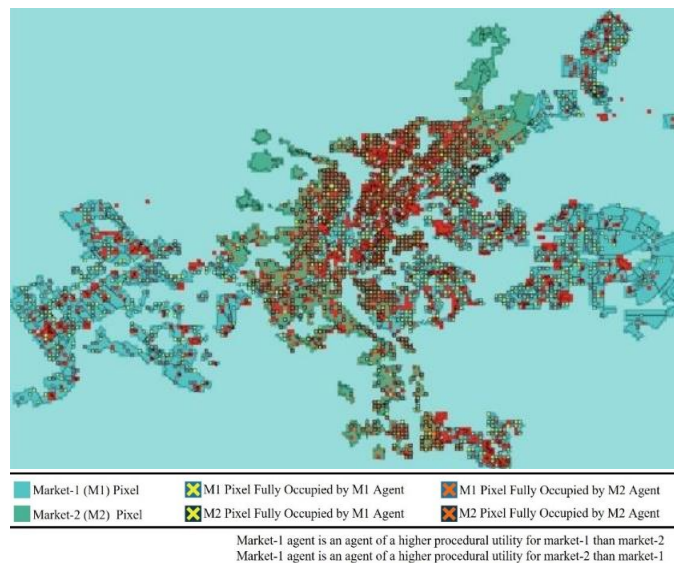


Figure 1, Sample ABM Run in Greater Cairo

In reflection, this paper showcases the relevance of market procedures on the demographic spatial distribution across urban contexts. This implies that land sub-market should be subdivided based on their regulatory procedures, rather than only on prices and immigration trends (Jones and Watkins, 2009). Further, this can have implications on planning decisions where market regulations of development projects can be manipulated to target specific groups based on their market preferences.

References

- Botosan, C. A. (2006) 'Disclosure and the cost of capital: What do we know?', *Accounting and Business Research*, 36(SPEC. ISS), pp. 31–40. doi: 10.1080/00014788.2006.9730042.
- David Sims (2010) *Understanding Cairo The Logic of a City Out of Control*. Cairo: American University in Cairo Press.
- Emil Kauder (1965) *A History of Marginal Utility Theory*. New Jersey: Princeton University Press.
- Filatova, T., Veen, A. Van Der and Voinov, A. (2008) 'An agent-based model for exploring land market mechanisms for coastal zone management', *Proc. iEMSs 4th Biennial Meeting - Int. Congress on Environmental Modelling and Software: Integrating Sciences and Information Technology for Environmental Assessment and Decision Making, iEMSs 2008*, 2, pp. 792–799.
- Frey, B. S. (2008) *Happiness A revolution in Economics*. Edited by H.-W. Sinn. MIT Press.
- Gurran, N. and Bramley, G. (2017) *Urban Planning and the Housing Market, Urban Planning and Renewal*. London: Palgrave Macmillan UK. doi: 10.1057/978-1-137-46403-3.
- Jones, C. and Watkins, C. (2009) *Housing Markets & Planning Policy, The Encyclopedia of Housing*. West Sussex: Wiley-Blackwell.
- Pratt, S. P. and Grabowski, R. J. (2014) *Cost of capital: applications and examples*. Fifth Edit. John Wiley & Sons.
- Ryan, R. M. and Connell, J. P. (1989) 'Perceived locus of causality and internalization: Examining reasons for acting in two domains', *Journal of Personality and Social Psychology*, 57(5), pp. 749–761. doi: <http://dx.doi.org/10.1037/0022-3514.57.5.749>.
- Ryan, R. M. and Deci, E. L. (1982) 'The Psychology of Self-Determination', (March 2018). doi: 10.2307/2067164.
- Sen, A. (1995) 'Rationality and Social Choice', *The American Economic Review*, 85(1), pp. 1–24. Available at: <https://www.jstor.org/stable/2117993>.
- Xie, Y., Batty, M. and Zhao, K. (2007) 'Simulating emergent urban form using agent-based modeling: Desakota in the Suzhou-Wuxian Region in China', *Annals of the Association of American Geographers*, 97(3), pp. 477–495. doi: 10.1111/j.1467-8306.2007.00559.x.

Special Session SS03_04.2

4 November 11:20 am-1:20 pm (GMT)

Exploration and validation of spatial simulation models

Proponents

Juste Raimbault, Denise Pumain, Eric Koomen and Chris Jacobs-Crisioni

Exploring methods for simulating urban negotiations

Aya BADAWEY¹; Nuno PINTO²; Richard KINGSTON^{1,2}

¹Planning and Environmental Management, University of Manchester
M13 9PL, UK

+44 (0) 7903087712, aya.badawy@manchester.ac.uk

² Planning and Environmental Management, University of Manchester
M13 9PL, UK

+44 (0) 161 275 2729, nuno.pinto@manchester.ac.uk

^{1,2} Planning and Environmental Management, University of Manchester
M13 9PL, UK

+44 (0) 161 275 1936, Richard.kingston@manchester.ac.uk

Keywords: urban negotiations, developer contributions, agent-based modelling, Finite-State-Machine (FSM), Belief-Desire-Intention (BDI)

1. Introduction

Researchers develop urban or geographical simulation models usually to understand how certain policies or decisions may affect the urban system and dynamics. However, they ignored asking a precedent question which is how stakeholders arrive at these decisions in the first place. Using models to answer this question can give better insights and lead to a more proactive approach to urban decision-making.

The urban system involves intersecting sectors, so several stakeholders play a role in the decision-making process. Because they have different interests, they usually negotiate to reach a common ground. The way they negotiate leads to the decisions and policies they take in the urban system. Thus, to answer the question posed at the beginning, urban negotiations should be simulated. On this basis, our research explores the different methods for simulating urban negotiations. Our aim was not only to identify the potential methods by reviewing the literature, but also to prove their applicability by developing models with them. We identified two agent-based methods that can be used to model negotiation: the Finite-State Machine (FSM) and the Belief-Desire-Intention (BDI).

2. Potentials identified in two agent-based methods

FSM designs the interaction between agents in the form of states and transitions. It is the same concept of cellular automata and has been widely used in models of land use change. However, it can be tailored differently to simulate negotiation. Instead of states representing different land use, they will represent the different types of messages that can be exchanged between negotiators e.g., propose, accept

proposal, refuse...etc. A negotiator can be in a state called “received proposal” where it evaluates the offer of the opponent. According to the evaluation, the transition can be to one of several states: make counter-proposal, accept, reject...etc. It is a simple way of simulating the interaction.

As for the BDI method, it is more advanced than FSM. It does not only simulate what agents do but also how they think before deciding what to do. As explained, conventional urban models usually simulate the effect of policies regardless how they have been issued. Therefore, they did not need to use the BDI. However, when simulating negotiation, BDI can be quite useful. The concept of BDI is that it creates a knowledge base for the agent where information is categorized into beliefs: what the agent knows, and desires: what the agent wants. When the agent is ready to implement a desire, it turns this desire to an intention, then execute it with one of the suitable plans in the plan library (Georgeff *et al.*, 1999). The agent’s knowledge base is continuously dynamic. Beliefs are updated with the new information that the agent gets from the environment and its peers which can trigger new desires. The architecture can help in simulating the reasoning of stakeholders while negotiating— how they perceive opponents, proposals, and the overall situation at each step then use a tactic to react.

3. Testing the applicability of methods in simulating negotiations

The literature, specifically the computational literature, used FSM and BDI for simulating negotiation in other domains. The aim is to prove their applicability for simulating urban negotiation which are more complex as negotiators explore new ideas to solve the problem and consider the urban system while negotiating.

3.1 The case of developer contributions

We applied the methods to simulate negotiations of developer contributions in England, particularly in projects of housing development. The process starts with developers who propose medium/large scale urban projects. They need to contribute to some public infrastructure— determined according to the context and need—to get their projects approved by the Local Planning Authority (LPA). However, these contributions are negotiable. Developers can negotiate with LPA the value of contributions, the method and schedule of provision...etc. (Healey, Purdue and Ennis, 1995) With housing projects, developers are often asked to allocate a percent of the project as affordable units as a form of contribution. Affordable units are not provided for ownership but for rent which means they will require a landlord. Therefore, developers usually collaborate with Registered Social Providers (RSL) to whom they can hand-out the units. They negotiate on the price of selling the units. In summary, developers negotiate with at least two stakeholders: the LPA and RSL.

The rationale for choosing this case is as follows. First, there have been various social and qualitative studies which describe the mechanisms of developer contributions and the involved stakeholders. Second, data is available for the values of the different issues that negotiators may discuss. The first two reasons will help in understanding the process, hence modelling it easily. The third reason is that negotiation here is not a one-off activity, but it occurs as a step while developing a new project. Thus, it will be suitable to test how to simulate negotiation within a bigger urban system.

3.2 Conceptual design of models

We simulated the process on the agent-based platform Gama. The platform provides several architectures that the modeller can choose from to design agents, among which are the FSM and BDI. To test their applicability, we developed two models for the case study. The negotiations in one are simulated with FSM architecture, and in the other with BDI. However, the activities other than negotiation e.g., preparing the project brief, completing the planning application, conducting a viability assessment for contributions...etc. are simulated similarly in both models using a task-based architecture.

As explained, each architecture designs agents with a different concept, and this affected how we identified the purpose of each model. With FSM, it is easy to represent all the possible interactions and add simple conditions for the transition between them. Therefore, we used the method to explore the possible outcomes of negotiation and the probability of each. It was done by randomizing the transition rules and running the simulation several times, hence allowing agents to interact with all the ways possible. With BDI, it is easy to represent the reasoning of agents which affects their choice of strategies and tactics. Therefore, we used the method to understand how changing the agenda and expectations of an agent can affect its reasoning and, in turn, the outcome. The models have different purposes but designing them gave us better insight about simulating urban negotiations.

4. Discussion

When comparing both methods for simulating urban negotiations, we concluded the following. FSM is easier to use, hence creates more interpretable models. This can make it a better choice for large scale models that involve many negotiations in an urban system. Some complexity can be added by integrating a spatial model that agents can use when evaluating proposals or simulating other forms of interaction like auction. However, we think that FSM capabilities will restrict negotiation to be an exchange of proposals between stakeholders who concede in their goals until one of them ends the discussion. It is not suitable for simulating how agents argue, restructure the problem by adding or changing issues to reach a solution, or at least analyse their opponents' behaviour and react

accordingly. Here, the BDI method is better. The BDI opens avenues for improving the model to simulate negotiations with more complexity. Most importantly, working with BDI highlighted its potentials to simulate reasoning in urban decision making and governance in general not only in negotiations. For example, answering a question like: what is the relation between the pattern of decision-makers' reasoning and the policies they enforce in the environment?

5. Conclusion

To sum up, FSM and BDI methods can both be used to simulate urban negotiations but from different perspectives. Selecting between them goes back to the purpose of the simulation and the required degree of complexity. Whatever the method, the main takeaway from the research is that simulation can help in understanding how stakeholders arrive at their decisions not only in exploring the impact of their decisions on the system.

References

- Georgeff, M. *et al.* (1999) 'The Belief-Desire-Intention Model of Agency', in Müller, J. P., Rao, A. S., and Singh, M. P. (eds) *Intelligent Agents V: Agents Theories, Architectures, and Languages*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–10.
- Healey, P., Purdue, M. and Ennis, F. (1995) *Negotiating development: rationales and practice for development obligations and planning gain*. London: E and FN Spon.

Validating a global urbanisation model

Pendula A. Ferdinand¹; Bo P. J. Andréé¹; Eric Koomen¹

¹ Vrije Universiteit Amsterdam, Faculty of Economics and Business Administration

Department of Spatial Economics/ Spatial Information Laboratory (SPINlab)

De Boelelaan 1105, 1081 HV Amsterdam, Netherlands

+31 20 5986095, p.a.ferdinand@vu.nl (correspondent author)

Keywords: land-use modelling, predictive modelling, logit model, urban development

Abstract

For this study, predictive modelling is used to validate a high-resolution model that simulates urban development. A logit model is used to predict the likelihood of a cell being classified as urban or non-urban based on historical data. Several essential choices had to be made during the validation process about the classification model and the model evaluation metrics. This presentation focuses on three important choices in performing a validation analysis: choosing between a static or dynamic analysis, selecting an appropriate performance indicator, and deciding between a linear or non-linear classification.

Introduction

Predictive modelling relates to “the process of uncovering relationships within data for predicting some desired outcome” (Kuhn and Johnson, 2013). This is an essential step in the validation of land-use change models. This presentation highlights how predictive modelling is used to validate a high-resolution model that simulates urban development. The 2UP model we apply in this case simulates changes in urban land use and population simultaneously (Huijstee *et al.*, 2018; Koomen *et al.*, 2019), but we focus on the former: how likely is a cell classified as urban or non-urban? Historical data can be used to train statistical models to make accurate out-of-sample predictions and simulate future urban growth. Historical data can consist of all existing cells classified as urban land-use (static analysis) or only of cells that became urban within a specific period (dynamic analysis). The core objective of predictive land-use models is to accurately predict land use classes without necessarily understanding the underlying processes. Nevertheless, expert knowledge is fundamental to develop an effective predictive model and select relevant data. This presentation focusses on three important choices in performing a validation analysis: choosing between a static or dynamic analysis, selecting an appropriate performance indicator, and deciding between a linear or non-linear classification.

Static versus dynamic validation

The choice between static and dynamic analysis depends on the research aim. For the validation of the 2UP model, both types of analyses were performed and compared. For the static analysis, a classification model was trained for each continent separately to predict the urban area in 2014, based on socio-economic, biophysical variables and policy predictors. For the dynamic analysis, a model was trained to only predict urban growth between 1990 and 2014 using the same predictors as for the static analysis. The advantage of static analysis is that all urban cells available can be used for model training and testing. Thus, the model has more urban cells to learn patterns between urban land use and the predictive variables. For the dynamic analysis, only urban cells for the specific period, e.g., 1990 and 2014, can be used. Results for the 2UP model showed that the model's predictive power for the static analysis was higher than for the dynamic analysis (Ferdinand, Andrée and Koomen, 2020). The disadvantage of static analysis is that the extent to which the obtained patterns can be extrapolated accurately into the future is limited. Especially initial drivers of agglomeration such as the proximity to natural phenomena (rivers or relatively flat terrain) that played an influential role in the emergence of the first urban clusters will be less important for future urban growth. If the aim is to predict future urban growth, a dynamic analysis, similar to studies by Braimoh and Onishi (2007), Batisani and Yarnal (2009) or Vermeiren *et al.* (2012), might be the better choice. A static analysis might be more useful if the research objective is to validate whether the model can reproduce existing land-use patterns. The static and dynamic validation results for the 2UP model have shown that the number of statistically significant predictors and the size of their effect vary between both analyses (Ferdinand, Andrée and Koomen, 2020). For example, for Europe, the travel time predictor became nearly twice as important for the shorter and more recent study period, and the terrain roughness index (TRI) became relatively less important. This is in line with studies by Vermeiren *et al.* (2012) and Cao *et al.* (2020).

Evaluating the performance of classification models

Assessing the performance of logistic regression models is not as straightforward as it is for regression models with continuous outcomes, where matrices like the RMSE and R-squared can be used (Kuhn and Johnson, 2013). To assess the performance of regression models with discrete outcomes, the actual and predicted classes are compared in a confusion matrix. Different evaluation metrics can be derived from the confusion matrix, such as the accuracy, kappa statistic or sensitivity and specificity. However, not all of these metrics are useful for evaluating the urbanization model due to the extreme class

imbalance between urban and non-urban classes. For all continents, the number of cells classified as urban only take up around one per cent of the data. Therefore, the model can have a nearly perfect accuracy by predicting all cells to be non-urban. To account for the imbalancedness, the final model for the validation of 2UP was selected based on a balanced accuracy result that gives equal weight to accuracy on both urban and non-urban validation samples. This performance metric thus ignores the imbalance in classes in a validation sample and only rewards model predictions that can differentiate between both urban and non-urban samples within both classes rather than simply reflecting the near-zero rate of occurrence of urban land use itself. More discussion on this validation metric, its policy interpretation in the context of model-based decision making, and its connection to a Weighted Maximum Likelihood criterion can be found in Andree *et al.* (2020).

Choosing the right model: linear versus non-linear classification models

Choosing the right modelling approach requires a trade-off between model complexity and performance. To validate the 2UP model, different machine learning approaches were compared to the same classification problem. The logit model was used as a benchmark model, as it is the most commonly used linear classification model for predicting a response variable with two categories. Besides the simple logit model, logistic regressions with an added penalty term have been explored to optimize the complexity and performance of the logit model. These so-called penalized logistic regression models place constraints on the coefficients, which prevents overfitting and reduces the number of independent variables without the need for the researcher to make a selection a priori. The three most widely used forms of penalized logistic regressions have been compared: ridge regression, lasso regression, and elastic net regression, a mix of a ridge and lasso regression. Additionally, a non-linear random forest model was compared to the linear approaches. The results showed no substantial improvements of the more complex approaches over the logit model (Ferdinand, 2020). This made the logit model the optimal choice as it is relatively straightforward while still achieving strong predictive performance.

References

- [1] Andree, B. P. J. *et al.* (2020) 'Predicting food crises'. World Bank, Washington, DC.
- [2] Batisani, N. and Yarnal, B. (2009) 'Urban expansion in Centre County, Pennsylvania: Spatial dynamics and landscape transformations', *Applied Geography*. Elsevier Ltd, 29(2), pp. 235–249. doi: 10.1016/j.apgeog.2008.08.007.
- [3] Cao, Y. *et al.* (2020) 'Urban spatial growth modeling using logistic regression and cellular automata: A case study of Hangzhou', *Ecological Indicators*. Elsevier, 113(December 2019), p. 106200. doi: 10.1016/j.ecolind.2020.106200.
- [4] Ferdinand, P. (2020) *Forecasting patterns of urban expansion; A statistical analysis of forces determining locational urban growth and their regional differences*. MSc thesis STREEM, Vrije Universiteit Amsterdam.
- [5] Ferdinand, P., Andrée, B. P. J. and Koomen, E. (2020) *Revised calibration of the 2UP model*.
- [6] Huijstee, J. *et al.* (2018) *Towards an Urban Preview: Modelling future urban growth with 2UP*. The Hague.
- [7] Koomen, E, van Huystee, J, van Bommel, B, Bouwman, A, Ligtoet, W & Andree, BPJ 2019, Localising global urban development; simulating local exposure to natural hazards in the global 2UP model. in G Caruso, P Gerber, K Jones, O Klein & C Perchoux (eds), *Programme and Book of abstracts: 21st European Colloquium on Theoretical and Quantitative Geography*. Luxembourg Institute of Socio-Economic Research (LISER), Esch-sur-Alzette/Belval, Luxembourg, pp. 126-126.
- [8] Kuhn, M. and Johnson, K. (2013) *Applied predictive modeling*. Springer.
- [9] Vermeiren, K. *et al.* (2012) 'Urban growth of Kampala, Uganda: Pattern analysis and scenario development', *Landscape and Urban Planning*. Elsevier B.V., 106(2), pp. 199–206. doi: 10.1016/j.landurbplan.2012.03.006.

Benchmarking road network growth models

Juste RAIMBAULT^{1,*}

¹Center for Advanced Spatial Analysis, University College London
Gower Street, London WC1E 6BT, United Kingdom

* j.raimbault@ucl.ac.uk

Keywords: Road network, Model benchmark, Feasible space

Processes underlying the growth of road networks are diverse and complementary, as for example with the combination of self-organisation and top-down planning (Barthelemy et al., 2013). Multiple generative models, more or less parsimonious and data-driven, have been introduced in the literature to reproduce existing networks and provide potential explanations on main processes driving their growth. Whereas each model includes plausible mechanisms and often yields reasonable empirical results, a systematic and quantitative comparison of such models remains to be explored. We propose in this contribution such a benchmark of road network growth models. We include in the comparison (i) a random null model; (ii) a random potential breakdown model (Raimbault, 2020); (iii) a deterministic potential breakdown model (Raimbault, 2019); (iv) a cost-benefit compromise model (Louf et al., 2013); (v) a biological network generation model (Raimbault, 2018); and (vi) a self-reinforcement model (Molinero and Hernando, 2020). We use the GHSL dataset for functional urban areas worldwide and OpenStreetMap to extract real networks and population distributions for the 1000 largest urban areas, and to compute corresponding values of diverse network measures (including betweenness and closeness centralities, accessibility, performance, diameter, density, average link length, average clustering coefficient). The models are integrated into the spatialdata scala library (Raimbault et al., 2020) and into the OpenMOLE software for model exploration and validation (Reuillon et al., 2013). We then run a diversity search algorithm, the Pattern Space Exploration algorithm (Cherel et al., 2015), for each model with their own free parameters and with the population distribution also as input parameter among the sampled areas. This algorithm is specifically tailored to provide feasible spaces of model outputs in relatively low dimensions. We thus proceed to a principal component analysis on real data points and project simulated values on the two first components, taken as objectives of the diversity algorithm. We obtain different shapes of feasible point clouds and corresponding effective degrees of freedom, some regions in the objective space reachable by a single model only, and a small number of urban areas which can not be approximated by the models. This quantitatively confirms the complementarity of diverse processes driving road network growth, and the need for a plurality of models to explain it.

References

- Barthelemy, M., Bordin, P., Berestycki, H., & Gribaudo, M. (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific reports*, 3(1), 1-8.
- Chérel, G., Cottineau, C., & Reuillon, R. (2015). Beyond corroboration: Strengthening model validation by looking for unexpected patterns. *PloS one*, 10(9), e0138212.
- Louf, R., Jensen, P., & Barthelemy, M. (2013). Emergence of hierarchy in cost-driven growth of spatial networks. *Proceedings of the National Academy of Sciences*, 110(22), 8824-8829.
- Molinero, C., & Hernando, A. (2020). A model for the generation of road networks. *arXiv preprint arXiv:2001.08180*.
- Raimbault, J. (2018). Des systèmes naturels aux systèmes urbains: génération de réseaux de transport optimaux par modèle slime-mould. In *BioMim Expo 2018*.
- Raimbault, J. (2019). Second-order control of complex systems with correlated synthetic data. *Complex Adaptive Systems Modeling*, 7(1), 1-19.
- Raimbault, J. (2020). Unveiling co-evolutionary patterns in systems of cities: a systematic exploration of the simpopnet model. In *Theories and Models of Urbanization* (pp. 261-278). Springer, Cham.
- Raimbault, J., Perret, J., & Reuillon, R. (2020). A scala library for spatial sensitivity analysis. GISRUUK 2020 Proceedings.
- Reuillon, R., Leclaire, M., & Rey-Coyrehourcq, S. (2013). OpenMOLE, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, 29(8), 1981-1990.

Spatial and temporal transferability of urban growth models

- A data-driven framework of multi-area, multi-period calibration, growth mode clustering, and scenario development

Jingyan YU¹; Alex HAGEN-ZANKER¹; Naratip SANTITISSADEEKORN²; Susan HUGHES¹

¹Department of Civil and Environmental Engineering, University of Surrey, Guildford, Surrey, GU2 7XH, UK
+44 (0)07767940299, jingyan.yu@surrey.ac.uk (correspondent author)

²Department of Mathematics, University of Surrey, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Keywords: Transferability; Calibration; Non-stationarity; Urban growth models; Cellular Automata land-use change models

1. Introduction

Urban growth and land use change models are widely used to investigate urban form and dynamics (White et al., 2015). The models rely on parameters that are typically obtained by the calibration of the models to observed changes over a past period (Straatman et al., 2004, Engelen and White, 2008, van Vliet et al., 2016). The temporal non-stationarity of land use change processes between the calibration period and prediction period (Blecic et al., 2015, Chen et al., 2016, Mas et al., 2018, Feng et al., 2019, Qian et al., 2020, Aguejdad, 2021) as well as the spatial heterogeneity between the calibrated area and the prediction area (Brunsdon et al., 1996, McDonald and Urban, 2006, Luo and Wei, 2009, Li et al., 2015, Mondal et al., 2015, Mirbagheri and Alimohammadi, 2017, Zhang et al., 2017, Wang et al., 2019, Meentemeyer et al., 2013, Ke et al., 2016, Feng et al., 2018, Firozjaei et al., 2019, Shu et al., 2020, Xia and Zhang, 2021) challenge the transferability of the model and calibrated parameters (Santé et al., 2010, Li et al., 2013, Lin and Li, 2016, Liu et al., 2017, Silva, 2004, Clarke et al., 2007, Clarke, 2018, Votsis and Haavisto, 2019).

2. Methodology

This study proposes and implements a data-driven framework of urban growth modelling to investigate the spatial and temporal transferability of an urban growth model (Yu et al., 2021), which consists of multi-area, multi-period calibration, growth mode clustering, and growth scenario development.

The framework is first applied to ten study areas across five western European countries. The urban growth model is calibrated for all areas and two calibration periods, both on emerging patterns, through

a simulation of the full history of urban growth from urban genesis (designated as “0”) to 2000 (Figure 2) and on patterns of change over a shorter period from 1975 to 2000 (Figure 1). Cluster analysis of calibrated parameters identifies groups of parameters that represent different urban growth modes (Figure 3). Extensive cross-application and validation reveal the transferability of individual study area parameters and the parameter clusters of alternative growth scenarios.

3. Results

Results show that parameters calibrated by observed land use change data of a study area are not always the most suitable to project future urban growth of the same area (Figure 4). In most of the ten studied areas, this is due to a transition from a compact urban growth mode over the calibration period to a more dispersed urban growth mode of studied areas in the projection period. The identification of parameter clusters that correspond to different urban growth modes and the application of these clusters for urban growth projection illuminates uncertainty beyond the variation around a mean, but instead in terms of qualitatively different yet realistic scenarios.

4. Conclusions

The transferability of land use change model and calibrated parameters is due to the inherent uncertain nature of land use change processes. Transferability of land use change models shall be discussed with reference to the different modes of urban growth: within each mode of urban growth, there is good spatial and temporal transferability; however, the transferability between areas and time periods is poor when they belong to different growth modes. The studied areas and periods contain substantial variations in growth mode. Identifying clusters of calibrated parameters that represent different urban growth modes provides insights on the transferability issue: it is clusters of different urban growth modes that are transferable spatially and(or) temporally.

Results also suggest a repudiation of the common practice of calibrating models to the particular history of a study area and using this calibration for extrapolation. Instead, it is recommended to reflect the uncertainty of the nature of future urban growth and develop multiple scenarios of future growth under qualitatively different urban growth regimes. Urban growth and land use change modelling are recommended to utilize the framework of multi-case calibration, clustering, and scenario development to address the uncertainty of urban growth and land use change processes, rather than providing one definite prediction.

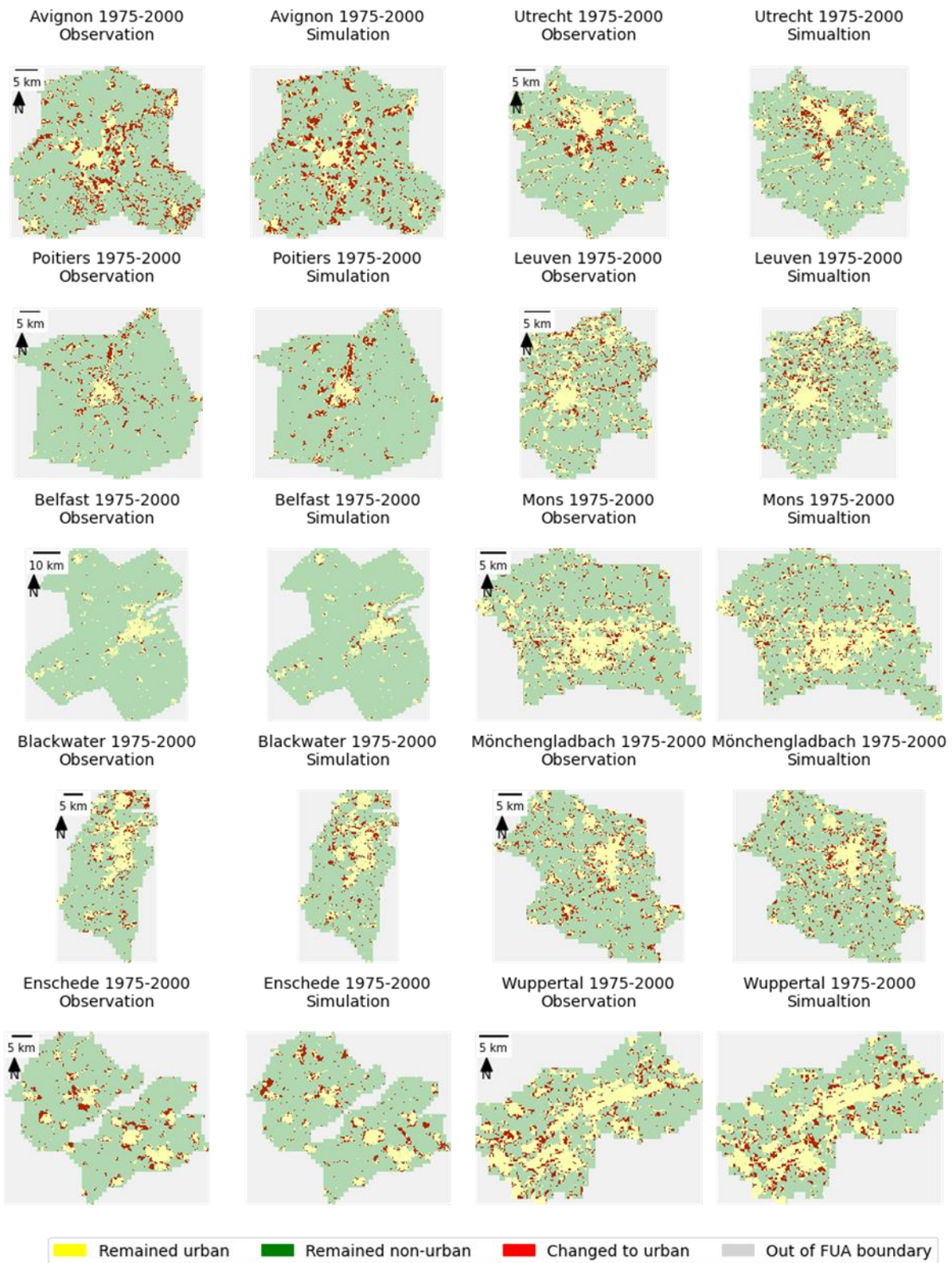


Figure 1: Short term 1975-2000 calibration result visualization

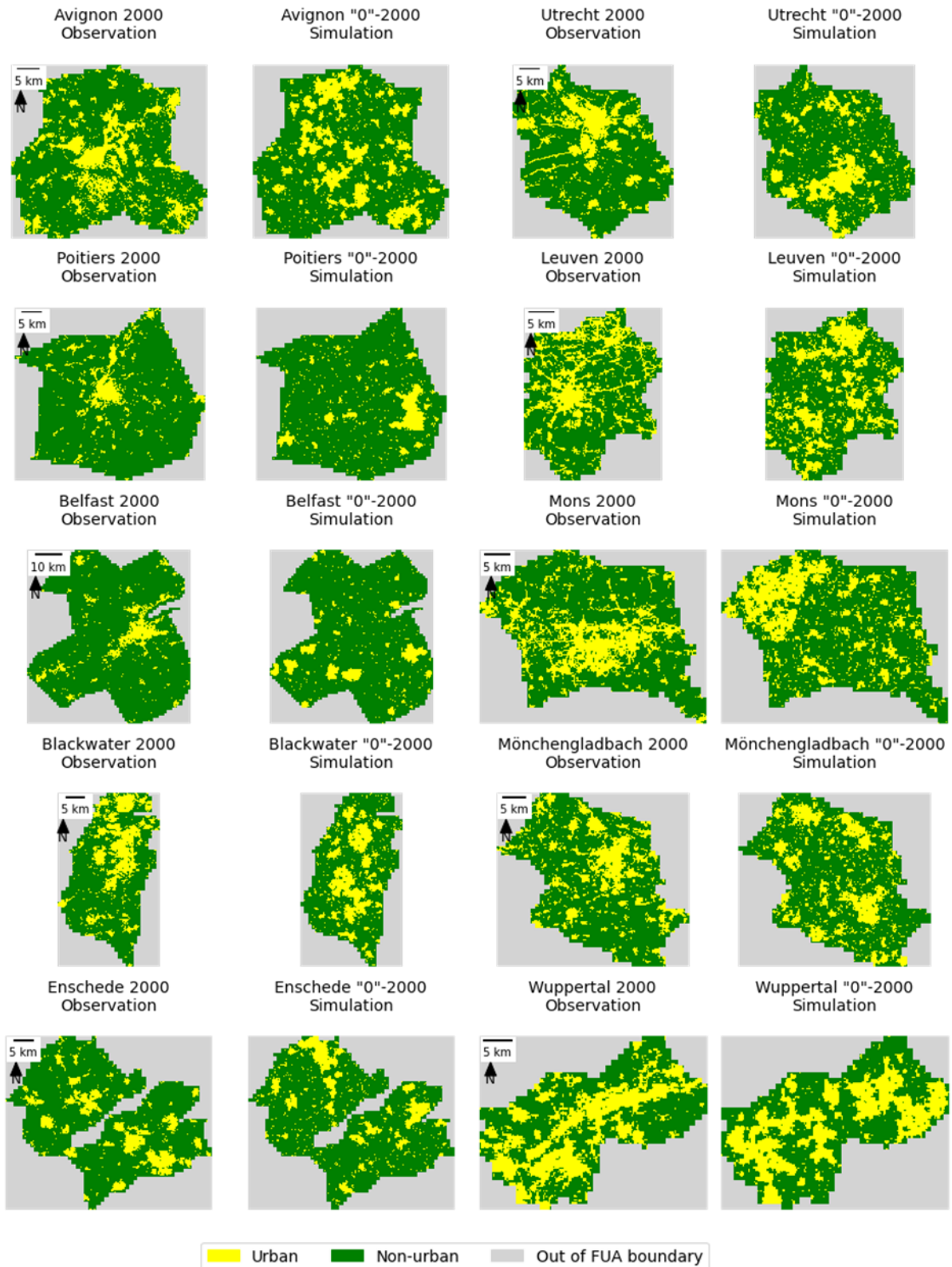


Figure 2: Long term “0”-2000 calibration result visualization

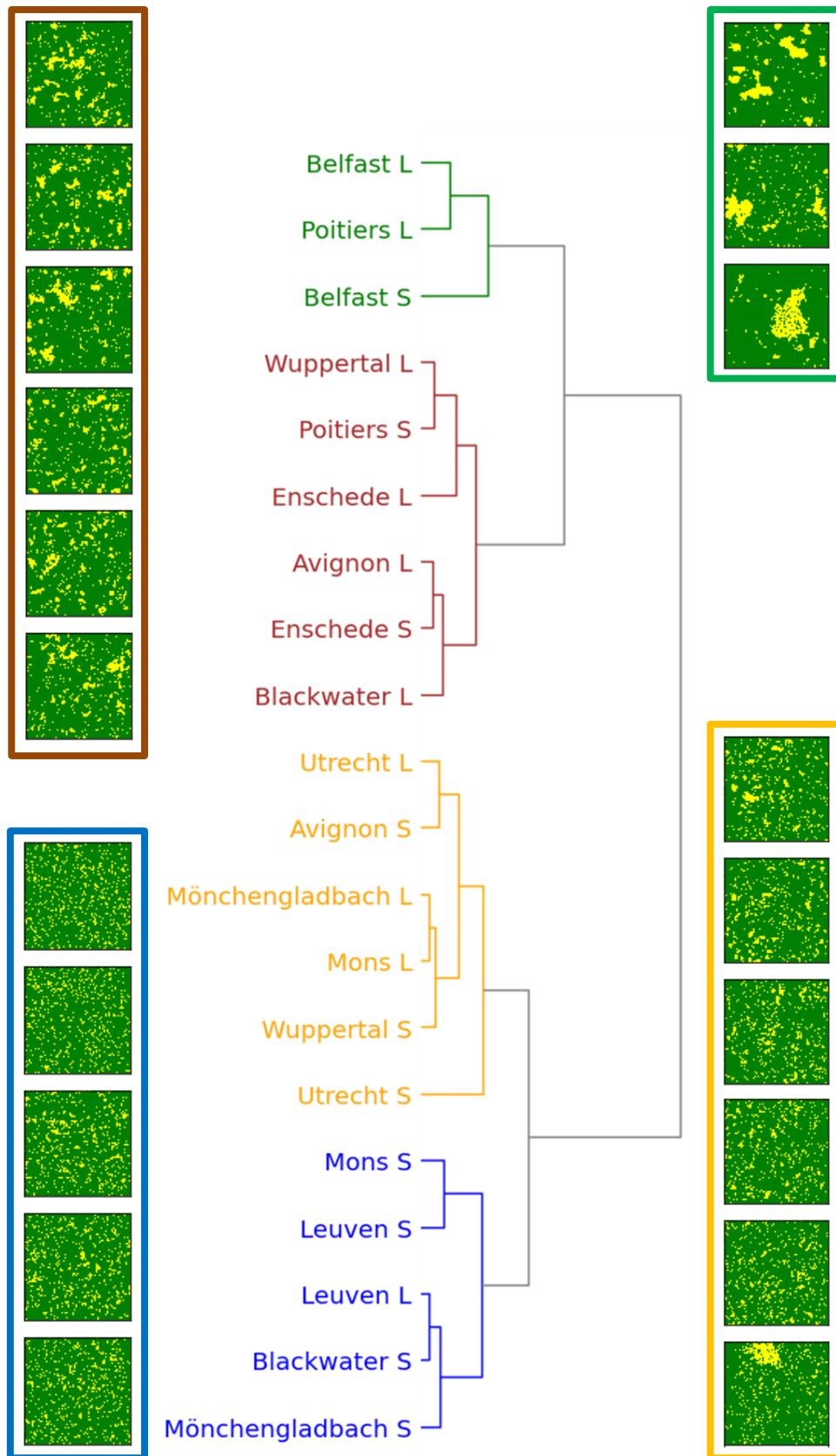


Figure 3: Clustering of the long term (L) and short term (S) calibrated parameters according to their simulated urban growth patterns into compact, medium compact, medium dispersed, dispersed

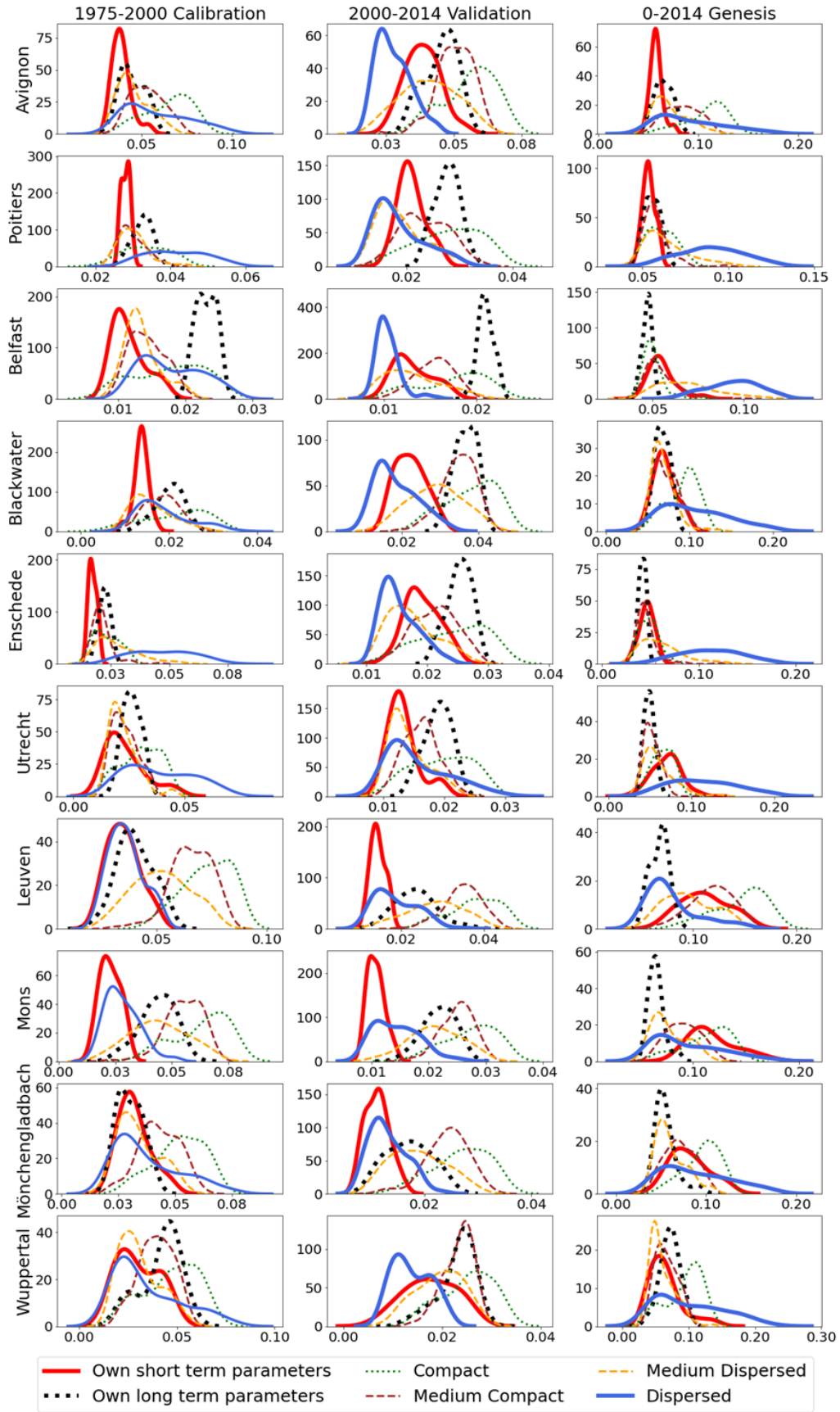


Figure 4: Validation measure - Calibrated parameter clusters' edge density in calibration period 1975-2000, short validation period 2000-2014, and long validation period "0"-2014

Acknowledgments

UKRI, NERC Landscape Decisions programme project number NE/T004150/1.

References

- AGUEJDAD, R. 2021. The Influence of the Calibration Interval on Simulating Non-Stationary Urban Growth Dynamic Using CA-Markov Model. *Remote Sensing*, 13, 468.
- BLECIC, I., CECCHINI, A. & TRUNFIO, G. A. 2015. How much past to see the future: a computational study in calibrating urban cellular automata. *International Journal of Geographical Information Science*, 29, 349-374.
- BRUNSDON, C., FOTHERINGHAM, A. S. & CHARLTON, M. E. 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, 28, 281-298.
- CHEN, Y., LI, X., LIU, X., AI, B. & LI, S. 2016. Capturing the varying effects of driving forces over time for the simulation of urban growth by using survival analysis and cellular automata. *Landscape and Urban Planning*, 152, 59-71.
- CLARKE, K., GAZULIS, N., DIETZEL, C. K. & GOLDSTEIN, N. 2007. A decade of SLEUTHing: Lessons learned from applications of a cellular automaton land use change model. *Classics from IJGIS. Twenty Years of the International Journal of Geographical Information Systems and Science*, 413-425.
- CLARKE, K. C. 2018. Land Use Change Modeling with SLEUTH: Improving Calibration with a Genetic Algorithm. In: CAMACHO OLMEDO, M. T., PAEGELOW, M., MAS, J.-F. & ESCOBAR, F. (eds.) *Geomatic Approaches for Modeling Land Change Scenarios*. Cham: Springer International Publishing.
- ENGELN, G. & WHITE, R. 2008. Validating and calibrating integrated cellular automata based models of land use change. *The dynamics of complex urban systems*. Springer.
- FENG, Y., WANG, J., TONG, X., LIU, Y., LEI, Z., GAO, C. & CHEN, S. 2018. The effect of observation scale on urban growth simulation using particle swarm optimization-based CA models. *Sustainability*, 10, 4002.
- FENG, Y., WANG, R., TONG, X. & SHAFIZADEH-MOGHADAM, H. 2019. How much can temporally stationary factors explain cellular automata-based simulations of past and future urban growth? *Computers, Environment and Urban Systems*, 76, 150-162.
- FIROZJAEI, M. K., SEDIGHI, A., ARGANY, M., JELOKHANI-NIARAKI, M. & ARSANJANI, J. J. 2019. A geographical direction-based approach for capturing the local variation of urban expansion in the application of CA-Markov model. *Cities*, 93, 120-135.

- KE, X., QI, L. & ZENG, C. 2016. A partitioned and asynchronous cellular automata model for urban growth simulation. *International Journal of Geographical Information Science*, 30, 637-659.
- LI, X., LIU, X. & GONG, P. 2015. Integrating ensemble-urban cellular automata model with an uncertainty map to improve the performance of a single model. *International Journal of Geographical Information Science*, 29, 762-785.
- LI, X., LIU, Y., LIU, X., CHEN, Y. & AI, B. 2013. Knowledge transfer and adaptation for land-use simulation with a logistic cellular automaton. *International Journal of Geographical Information Science*, 27, 1829-1848.
- LIN, J. & LI, X. 2016. Knowledge Transfer for Large-Scale Urban Growth Modeling Based on Formal Concept Analysis. *Transactions in GIS*, 20, 684-700.
- LIU, Y., HU, Y., LONG, S., LIU, L. & LIU, X. 2017. Analysis of the Effectiveness of Urban Land-Use-Change Models Based on the Measurement of Spatio-Temporal, Dynamic Urban Growth: A Cellular Automata Case Study. *Sustainability*, 9, 796.
- LUO, J. & WEI, Y. H. D. 2009. Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing. *Landscape and Urban Planning*, 91, 51-64.
- MAS, J. F., PAEGELOW, M. & CAMACHO OLMEDO, M. T. 2018. Lucc Modeling Approaches to Calibration. In: CAMACHO OLMEDO, M. T., PAEGELOW, M., MAS, J.-F. & ESCOBAR, F. (eds.) *Geomatic Approaches for Modeling Land Change Scenarios*. Cham: Springer International Publishing.
- MCDONALD, R. I. & URBAN, D. L. 2006. Spatially varying rules of landscape change: lessons from a case study. *Landscape and Urban Planning*, 74, 7-20.
- MEENTEMEYER, R. K., TANG, W., DORNING, M. A., VOGLER, J. B., CUNNIFFE, N. J. & SHOEMAKER, D. A. 2013. FUTURES: Multilevel Simulations of Emerging Urban–Rural Landscape Structure Using a Stochastic Patch-Growing Algorithm. *Annals of the Association of American Geographers*, 103, 785-807.
- MIRBAGHERI, B. & ALIMOHAMMADI, A. 2017. Improving urban cellular automata performance by integrating global and geographically weighted logistic regression models. *Transactions in GIS*, 21, 1280-1297.
- MONDAL, B., DAS, D. N. & DOLUI, G. 2015. Modeling spatial variation of explanatory factors of urban expansion of Kolkata: a geographically weighted regression approach. *Modeling Earth Systems and Environment*, 1, 29.
- QIAN, Y., XING, W., GUAN, X., YANG, T. & WU, H. 2020. Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation. *Science of The Total Environment*, 722, 137738.

- SANTÉ, I., GARCÍA, A. M., MIRANDA, D. & CRECENTE, R. 2010. Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and urban planning*, 96, 108-122.
- SHU, B., ZHU, S., QU, Y., ZHANG, H., LI, X. & CARSHJENS, G. J. 2020. Modelling multi-regional urban growth with multilevel logistic cellular automata. *Computers, Environment and Urban Systems*, 80, 101457.
- SILVA, E. A. 2004. The DNA of our regions: artificial intelligence in regional planning. *Futures*, 36, 1077-1094.
- STRAATMAN, B., WHITE, R. & ENGELEN, G. 2004. Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems*, 28, 149-170.
- VAN VLIET, J., BREGT, A., BROWN, D., VAN DELDEN, H., HECKBERT, S. & VERBURG, P. 2016. A review of current calibration and validation practices in land-change modeling. *Environmental Modelling and Software*, 82, 174-182.
- VOTSIS, A. & HAAVISTO, R. 2019. Urban DNA and Sustainable Cities: A Multi-City Comparison. *Frontiers in Environmental Science*, 7.
- WANG, H., STEPHENSON, S. R. & QU, S. 2019. Modeling spatially non-stationary land use/cover change in the lower Connecticut River Basin by combining geographically weighted logistic regression and the CA-Markov model. *International Journal of Geographical Information Science*, 33, 1313-1334.
- WHITE, R., ENGELEN, G. & ULJEE, I. 2015. *Modeling cities and regions as complex systems: From theory to planning applications*, MIT Press.
- XIA, C. & ZHANG, B. 2021. Exploring the effects of partitioned transition rules upon urban growth simulation in a megacity region: a comparative study of cellular automata-based models in the Greater Wuhan Area. *GIScience & Remote Sensing*, 1-24.
- YU, J., HAGEN-ZANKER, A., SANTITISSADEEKORN, N. & HUGHES, S. 2021. Calibration of cellular automata urban growth models from urban genesis onwards - a novel application of Markov chain Monte Carlo approximate Bayesian computation. *Computers, Environment and Urban Systems*, 90, 101689.
- ZHANG, L., WEI, Y. D. & MENG, R. 2017. Spatiotemporal Dynamics and Spatial Determinants of Urban Growth in Suzhou, China. *Sustainability*, 9, 393.

Special Session SS02.1

4 November 9:00-11:00 am (GMT)

The multitude of spatiotemporal scales in urban systems

Proponents

Janka Lengyel, Seraphim Alvanides, Stephane Roux and Patrice Abry

On the modeling of the interdependences of spatial scales in urban systems

J. Friedrich^{1,*}, S. Albanides², J. Lengyel³

¹ Institute of Physics and For Wind, University of Oldenburg
Küpkersweg , 26129 Oldenburg, Germany
+49 441 798-5023, jan.friedrich@uni-oldenburg.de
* Corresponding author

² GESIS-Leibniz Institute for the Social Sciences
Unter Sachsenhausen 6-8, 50667 Köln, Germany
+49 221 47694-423, serafeim.alvanides@gesis.org

³ Univ. Lyon, ENS de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique
46 allée d'Italie, 69342 Lyon, France
+33 6 26 71 23 16, janka.lengyel@ens-lyon.fr

Keywords: Multiscale framework, Strongly correlated systems, Urban analysis and modelling

Submitted to Special Session SS02

We propose a novel approach to the analysis and modeling of urban structures which is based upon concepts and methods from strongly interacting systems in physics. To this end, we perform a multi-scale analysis of urban structures on a large dataset of rent price values in the Ruhr area, Germany. We argue that, due to their many interacting degrees of freedom, urban systems exhibit similar features as other strongly correlated systems, e.g., turbulent flows or stock market indices, notably the occurrence of very strong fluctuations at small scale separations. It is shown explicitly how this analogy can be used to model such multiscale urban features and how common fractal/self-similar models have to be modified in order to account for this peculiar behavior.

1. The notion of scales in urban systems

Spatial scales and their relevance for urban systems have been addressed both from the analysis objective and as means to gain further insights into urban characteristics [Manson, 2007]. Whereas the latter focuses for instance on the inter- and intra-city spatial organization [Liu et al., 2018], the former is oftentimes used to reveal universal features of the urban form, such as urban scaling laws [Betten-court, 2013, Cottineau et al., 2017, Lemoy and Caruso, 2020] and fractal analysis [Batty and Longley, 1994, Tannier and Pumain, 2005]. Nonetheless, the fundamental question of how to reveal and quantify the interdependence of processes in between various spatial and temporal scales is less often addressed. Using the example of a rent price field in the Ruhr area in Germany depicted in Fig. 1(a), we will first demonstrate that strong intra-scale interactions lead to the emergence of non-self-similar (or multifractal) behavior of rent price fluctuations at small scales and then establish an analogy to the phenomenon of small-scale intermittency in turbulent flows [Friedrich, 2021]. This analogy can be used in order to model strong interdependencies in between scales by surrogate models of the urban form.

2. Surrogate models of urban forms

On the basis of the proposed analogy and the statistical analysis outlined in Sec. 1, we can now utilize

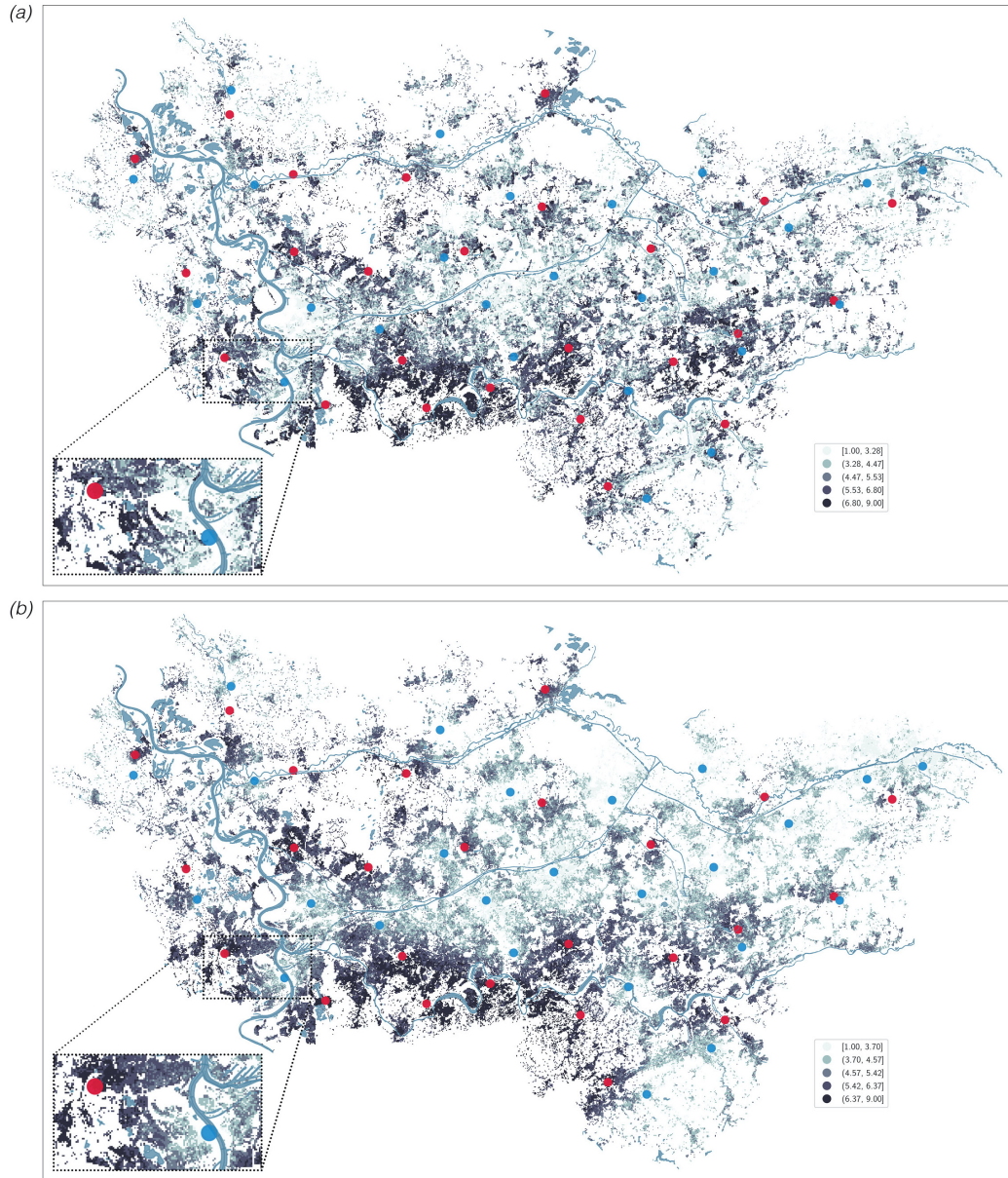


Figure 1: (a) Original rent price field in the Ruhr area, Germany (2016): Average value of mean rent price indices on a 100 x 100 meters grid. The field consists of 440572 points. Mean = 4.969. (b) Surrogate model of the rent price field in 2016. The field is reconstructed from 50 sampling points and possesses the same mean as in (a). *Red dots*: Hot-spots. *Blue dots*: Cold-spots. *Insets*: Zoom into close vicinity of two hot- and cold-spots.

a framework that has been devised in the context of surrogate turbulent fields. The framework combines a stochastic interpolation of a given set of prescribed points (e.g., the hot- and cold-spots of the rent price field in Fig. 1) by a so-called multipoint fractional Brownian bridge [Friedrich et al., 2020] and the knowledge of a recently developed explicit construction of a joint statistics of a multifractal field [Friedrich et al., 2021]. The latter allows for the introduction of strong fluctuations at small scales and thus accounts for the empirically observed non-self-similarity. A typical example of the synthesis of such a surrogate field is depicted in Fig. 1(b) where model parameters have been estimated from the statistical analysis of the original rent price field in Fig. 1(a). We further give an outlook on the applicability of our modelling approach to data-reconstructing and data-remodelling efforts for other urban fields (e.g., land use, temperature, or aerosol concentration).

References

- [Batty and Longley, 1994] Batty, M. and Longley, P. A. (1994). *Fractal cities: a geometry of form and function*. Academic press.
- [Bettencourt, 2013] Bettencourt, L. M. (2013). The origins of scaling in cities. *Science*, 340(6139):1438–1441.
- [Cottineau et al., 2017] Cottineau, C., Hatna, E., Arcaute, E., and Batty, M. (2017). Diverse cities or the systematic paradox of urban scaling laws. *Computers, environment and urban systems*, 63:80–94.
- [Friedrich, 2021] Friedrich, J. (2021). *Non-perturbative Methods in Statistical Descriptions of Turbulence*. Springer.
- [Friedrich et al., 2020] Friedrich, J., Gallon, S., Pumir, A., and Grauer, R. (2020). Stochastic interpolation of sparsely sampled time series via multipoint fractional brownian bridges. *Physical Review Letters*, 125(17):170602.
- [Friedrich et al., 2021] Friedrich, J., Peinke, J., Pumir, A., and Grauer, R. (2021). Explicit construction of joint multipoint statistics in complex systems. *Journal of Physics: Complexity*.
- [Lemoy and Caruso, 2020] Lemoy, R. and Caruso, G. (2020). Evidence for the homothetic scaling of urban forms. *Environment and Planning B: Urban Analytics and City Science*, 47(5):870–888.
- [Liu et al., 2018] Liu, X., Derudder, B., and Wang, M. (2018). Polycentric urban development in China: A multi-scale analysis. *Environment and Planning B: Urban Analytics and City Science*, 45(5):953–972.
- [Manson, 2007] Manson, S. M. (2007). Challenges in evaluating models of geographic complexity. *Environment and Planning B: Planning and Design*, 34(2):245–260.
- [Tannier and Pumain, 2005] Tannier, C. and Pumain, D. (2005). Fractals in urban geography: a theoretical outline and an empirical example. *Cybergeo: European Journal of Geography*.

Structural fragility of cities to airborne releases

Sofia FELLINI^{1,2}; Pietro SALIZZONI¹; Luca RIDOLFI²

¹Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, Université de Lyon, Ecole Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon I
36 Av. Guy de Collongue, 69134 Ecully, France

²Department of Environmental, Land and Infrastructure Engineering
Corso Duca degli Abruzzi 24, 10129 Turin, Italy
sofia.fellini@hotmail.it (correspondent author)

Keywords: Complex network theory, Urban pattern, Urban fluid mechanics

Cities are particularly vulnerable to air pollution due to the presence of many potential sources and a high population density. Urban air pollution is mainly related to vehicular traffic and the heating of buildings. However, urban areas are also exposed to accidental releases such as those related to gas leaks, industrial plants or the transport of hazardous material. Evidently, this has also a link to city security (Coaffee et al. 2008), since toxic substances could be maliciously dispersed in the urban atmosphere for terroristic purposes.

In this framework, public authorities are urged to adopt operational tools to rapidly assess the transport of pollutants within the urban canopy and their impact on citizens health.

To this aim, computational fluid dynamics and simplified models based on empirical parametrizations have been widely adopted in the last decades (Blocken 2015, Tominaga & Stathopoulos, 2016). These models suggested that flow patterns within the urban canopy are strongly driven by the layout of buildings and the way street canyons cross each other. In this sense, the city structure plays an active role in conveying pollutants, especially for the transport dynamics at the pedestrian level. Consequently, the same pollutant source may have a different impact if released in cities with different urban plans or in different places within the same urban district.

In recent years, the link between urban morphology and air pollution has been investigated by focusing on key geometrical properties as the urban shape (Fan et al. 2019), the packing density of buildings (Buccolieri et al. 2015, Peng et al. 2019) and geometric characteristics of the street canyons (Miao 2020). However, the way street connectivity and layout can affect urban ventilation is still an uncharted territory.

In this work, we investigate these aspects with the aim to shed light on one of the most challenging questions in the field of urban safety and planning: what makes a city or an urban area vulnerable to

localized airborne releases? How much does topology, in which the history of a city is written, affect ventilation in the streets and thus the vulnerability of the city in case of a toxic point source?

To answer these questions we adopt the approach introduced by Fellini et al. (2019), i.e. a complex network perspective to study dispersion processes in cities. The urban canopy is modelled as a weighted and directed complex network: the streets and the street intersections are the links and the nodes of the network. The geometry of the street canyon and the wind field inside it determine the direction and the weight of each link. Within this approach, dispersion in the urban atmosphere is modelled as a spreading process on a network and local vulnerability is easily computed (in a variety of meteorological conditions) by means of a centrality metric (Fellini et al. 2020) that associates to each node in the network a value based on its spreading potential, and thus on the extent of the contaminated area when a release takes place in the node.

Adopting this innovative model, we compare the vulnerability of four districts in Lyon, Paris (France), Firenze (Italy) and New York (US) (Figure. 1). These urban areas were chosen as emblematic of different topologies, given by different historical urban layering.



Figure 1: Vulnerability maps for (a) Lyon, (b) Firenze, (c) Paris, and (d) New York for a single wind direction. Panels a1, b1, c1 and d1 show the urban pattern for the different cities.

Simulation results evidence the different resilience of cities to gas propagation. Moreover, vulnerability is highly heterogeneous within a single urban area and sensitive to wind direction.

The reasons for the different fragility of cities (and their patterning) to gas propagation are embedded in the centrality metric adopted to compute urban vulnerability. The key factors for vulnerability can then be analytically recognized in the metric definition. Here, we decompose the analytical model in its fundamental bricks and we isolate the variables that drive pollutant spreading in the streets. In this way, we are able to link the vulnerability of a city to its tangible, urban characteristics and to capture the dominant role of urban topology.

Acknowledgments

This study was supported by the Italo-French/French-Italian University via the Vinci Program, and by the Région Auvergne Rhône-Alpes via the SCUSI Project.

We would also like to express our gratitude to Lionel Soulhac and the equipe AIR for giving access to data.

References

- Coaffee, J., Moore, C., Fletcher, D., & Bosher, L. (2008). Resilient design for community safety and terror-resistant cities. In *Proceedings of the Institution of Civil Engineers-Municipal Engineer*, vol. 161, 103–110.
- Blocken, B. (2015). Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. *Building and Environment*. 91, 219–245 .
- Tominaga, Y. & Stathopoulos, T. (2016). Ten questions concerning modeling of near-field pollutant dispersion in the built environment. *Building and Environment*. 105, 390–402.
- Fan, Y., Wang, Q., Yin, S. & Li, Y. (2019). Effect of city shape on urban wind patterns and convective heat transfer in calm and stable background conditions. *Building and Environment*. 162, 106288.
- Buccolieri, R., Salizzoni, P., Soulhac, L., Garbero, V. & Di Sabatino, S. (2015). The breathability of compact cities. *Urban Climate* 13, 73–93.
- Peng, Y., Gao, Z., Buccolieri, R. & Ding, W. (2019). An investigation of the quantitative correlation between urban morphology parameters and outdoor ventilation efficiency indices. *Atmosphere* 10, 33.
- Miao, C., Yu, S., Hu, Y., Bu, R., Qi, L., He, X., & Chen, W. (2020). How the morphology of urban street canyons affects suspended particulate matter concentration at the pedestrian level: An in-situ investigation. *Sustainable Cities and Society* 55, 102042.

Fellini, S., Salizzoni, P., Soulhac, L. & Ridolfi, L. (2019). Propagation of toxic substances in the urban atmosphere: A complex network perspective. *Atmospheric Environment*. 198, 291–301.

Fellini, S., Salizzoni, P. & Ridolfi, L. (2020). Centrality metric for the vulnerability of urban networks to toxic releases. *Physical Review E* 101, 032312.

Defining roughness and intermittency within metropolitan regions

J. Lengyel¹, S. Roux², F. Sémécurbe³, S. Jaffard⁴, P. Abry⁵

¹ Univ. Lyon, ENS de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique
46 allée d'Italie, 69342 Lyon, France
+33 4 72 72 80 00, janka.lengyel@ens-lyon.fr
* Corresponding author

² Univ. Lyon, ENS de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique
46 allée d'Italie, 69342 Lyon, France
+33 4 72 72 83 78, stephane.roux@ens-lyon.fr

³ THÉMA, Université de Franche-Comté
32 rue Mégevand, F-25030 Besancon, France
+33 3 81 66 54 01, francois.semecurbe@univ-fcomte.fr

⁴ Univ. Paris-Est - Créteil Val-de-Marne, LAMA
61 avenue du Général de Gaulle, F-94010 Créteil, France
+33 (1) 45 17 65 84, stephane.jaffard@u-pec.fr

⁵ Univ. Lyon, ENS de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique
46 allée d'Italie, 69342 Lyon, France
+33 4 72 72 84 93, patrice.abry@ens-lyon.fr

Keywords: Multiscale urban analysis, Roughness, Intermittency

The main aim of this contribution is to introduce a novel method that is able to describe and analyse roughness and intermittency in smaller-scale urban structures, i.e. within municipal or metropolitan boundaries. We show how the direct *application of wavelet theory to urban multifractal analysis and modeling* is beneficial with regard to our latter goal. In more detail, the spatially continuous scanning of building distributions in the three largest French conurbations - over their entire territories and at length scales ranging from parcel to neighbourhood level - will allow us to derive and compare globally and locally characteristic scaling exponents. Depending on the multiscale quantity under analysis, the resulting exponents reveal qualitatively distinct structural properties. The viability of our findings will be further verified on four exemplary typologies of multiscale behavior in urban structures, which may provide a compelling alternative to how we define scale, size and distributional variability within urban landscapes.

1. The box-counting and wavelet coefficient in urban systems

We start by describing the two kinds of multiscale quantities, the box-counting and wavelet coefficient, that will build the basis for all our computed quantities. More precisely, we demonstrate that as a result of its inherent nature, the rather commonly used box-counting coefficient [Chen and Zhou, 2003, Lu and Tang, 2004, Lagarias and Prastacos, 2020] is oftentimes inapt for differentiating between neighbourhood-level scaling characteristics. Therefore, we introduce an additional multiscale quantity in geographical analysis the so-called wavelet coefficient. Even though wavelet theory has already been explored in the geospatial signal processing domain, e.g. for remote sensing and corresponding

feature-extraction [Couloigner and Ranchin, 1998, Huang and Zhang, 2012, Raja et al., 2013], it has been scarcely adopted to gain further fundamental insights on urban structures. Wavelets, in general, are shown to be highly advantageous for the multiresolution analysis of non-stationary signals as their scaling characteristics are believed to be more local in nature [Lee and Yamamoto, 1994]. Wavelet analysis were notably shown to provide computationally efficient multiscale representations, robust to trends and non-stationarities, enabling to assess scale-free dynamics in signal and images [Arneodo et al., 2000, Decoster et al., 2000, Wendt et al., 2009]. The here applied wavelet transfer for morphological analysis will be shown to be useful for describing the concepts of roughness and intermittency as understood in classical multifractal analysis using the metropolitan level as an example as well as for differentiating between neighbourhood-level scaling characteristics.

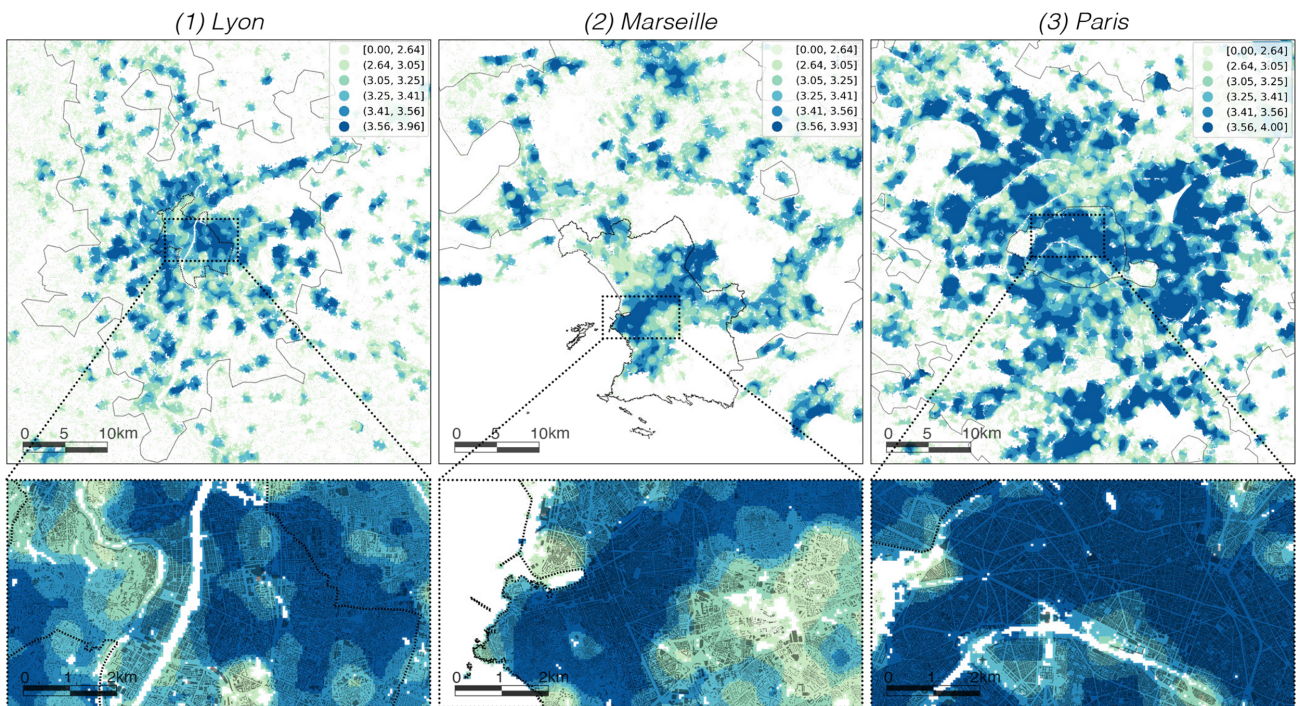


Figure 1: **Local scaling exponent ζ_2^c of the second-order using normalized sand-box coefficients $\tilde{N}_e(r)$, the exponent is computed for every grid point $g(x_g, y_g)$, $\mathbf{T} = 650\text{m}$, $\mathbf{r} = 50\text{-}400$ meters (using 25-meter gaps), $\mathbf{p} = 50$ meters, **Data:** Building centroids $e(x_e, y_e)$. **Top:** Results for the entire observed territory of 2500 square-kilometers **Bottom:** Inset showing results for the city centers. (1) Lyon, mean $R^2 = 0.987$ (2) Marseille, mean $R^2 = 0.99$ (3) Paris, mean $R^2 = 0.996$**

2. Geographically weighted multiscale analysis

In the second part of our talk, we introduce the Geographically Weighted Multiscale Analysis (GWMSA) which is partly based on previous work by [Sémécurbe et al., 2019]. GWMSA is a local, spatially non-stationary way of conducting geographical analysis and its goal is to scan how urban scaling and intermittency measures locally fluctuate across the observed area and what this might reveal about the spatial organization of underlying settlement structures. Our observations emphasize that one must introduce a careful distinction between the notions of *global and local* density, diversity, compactness and complexity and introduce differing multiscale parameters ($N_e(r)$, $D_e(r)$) and methodologies (classical multifractal analysis and GWMSA) to describe them. Finally, we combine our local GWMSA

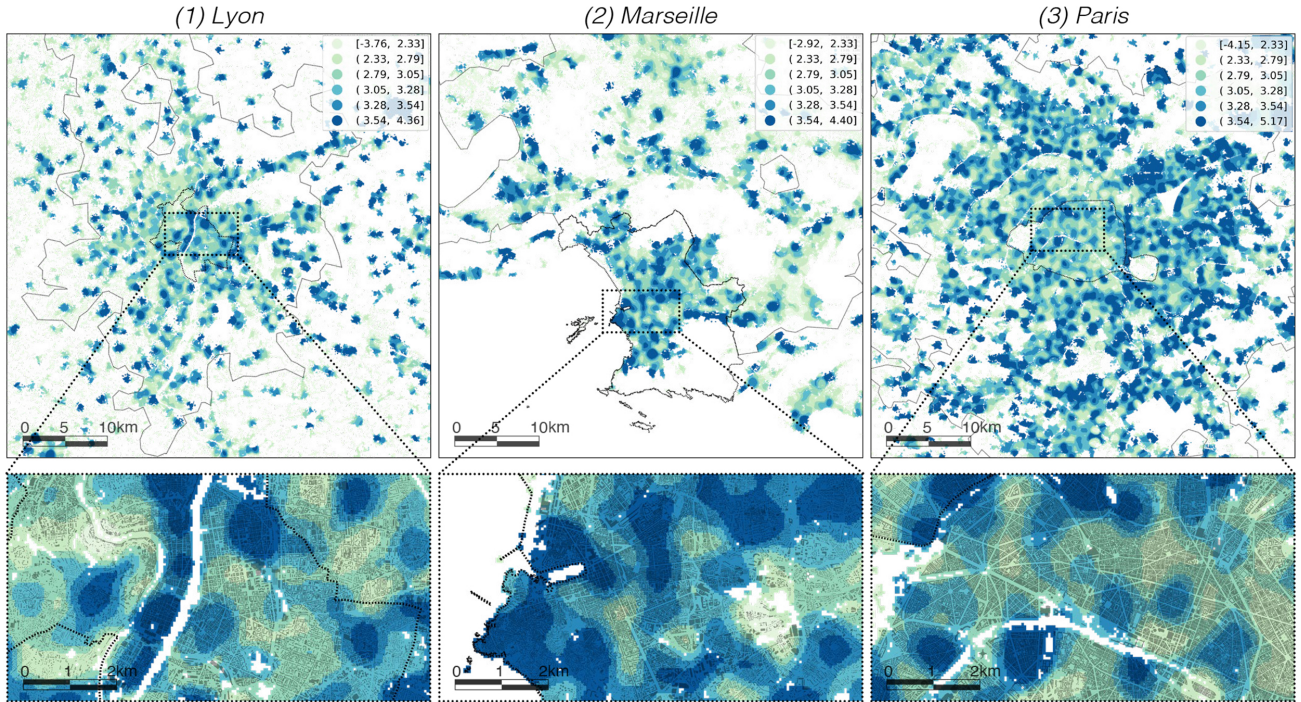


Figure 2: **Local scaling exponent ζ_2^d - roughness exponent - of the second-order structure function using normalized wavelet coefficients $\tilde{D}_e(r)$** , exponents are computed for every grid point $g(x_g, y_g)$, $\mathbf{T} = 650\text{m}$, $\mathbf{r} = 50\text{-}400$ meters (using 25-meter gaps), $\mathbf{p} = 50$ meters **Top**: Results for the entire observed territory of 2500 square-kilometers **Bottom**: Inset showing results for the city centers (1) Lyon, mean $R^2 = 0.947$ (2) Marseille, mean $R^2 = 0.961$ (3) Paris, mean $R^2 = 0.978$

results for the typologisation of multiscale behavior in urban structures, which allows us to make a distinction between two kinds of predominant scale-dependent behaviour: The expanding and contracting cities.

References

- [Arneodo et al., 2000] Arneodo, A., Decoster, N., and Roux, S. . (2000). A wavelet-based method for multifractal image analysis. i. methodology and test applications on isotropic and anisotropic random rough surfaces. *The European Physical Journal B-Condensed Matter and Complex Systems*, 15(3):567–600.
- [Chen and Zhou, 2003] Chen, Y. and Zhou, Y. (2003). The rank-size rule and fractal hierarchies of cities: mathematical models and empirical analyses. *Environment and Planning B: Planning and Design*, 30(6):799–818.
- [Couloigner and Ranchin, 1998] Couloigner, I. and Ranchin, T. (1998). Extraction of urban network from high-spatial-resolution imagery using multiresolution analysis and wavelet transform. In *Wavelet Applications in Signal and Imaging Processing VI*, volume 3458, pages 103–112. International Society for Optics and Photonics.
- [Decoster et al., 2000] Decoster, N., Roux, S., and Arneodo, A. (2000). A wavelet-based method for multifractal image analysis. ii. Applications to synthetic multifractal rough surfaces. *The European Physical Journal B-Condensed Matter and Complex Systems*, 15(4):739–764.
- [Huang and Zhang, 2012] Huang, X. and Zhang, L. (2012). A multiscale urban complexity index based on 3d wavelet transform for spectral-spatial feature extraction and classification: an evaluation on the 8-channel WorldView-2 imagery. *International Journal of Remote Sensing*, 33(8):2641–2656.
- [Lagarias and Prastacos, 2020] Lagarias, A. and Prastacos, P. (2020). Comparing the urban form of south european cities using fractal dimensions. *Environment and Planning B: Urban Analytics and City Science*, 47(7):1149–1166.
- [Lee and Yamamoto, 1994] Lee, D. T. and Yamamoto, A. (1994). Wavelet analysis: theory and applications. *Hewlett Packard journal*, 45:44–44.

- [Lu and Tang, 2004] Lu, Y. and Tang, J. (2004). Fractal dimension of a transportation network and its relationship with urban growth: a study of the Dallas-Fort Worth area. *Environment and Planning B: Planning and Design*, 31(6):895–911.
- [Raja et al., 2013] Raja, R. A., Anand, V., Kumar, A. S., Maithani, S., and Kumar, V. A. (2013). Wavelet based post classification change detection technique for urban growth monitoring. *Journal of the Indian Society of Remote Sensing*, 41(1):35–43.
- [Sémécurbe et al., 2019] Sémécurbe, F., Tannier, C., and Roux, S. G. (2019). Applying two fractal methods to characterise the local and global deviations from scale invariance of built patterns throughout mainland france. *Journal of Geographical Systems*, 21(2):271–293.
- [Wendt et al., 2009] Wendt, H., Roux, S. G., Jaffard, S., and Abry, P. (2009). Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Processing*, 89(6):1100–1114.

Measuring the Polycentricity based on urban and intercity transportation networks in Greater Bay Area: a cross-scale method in the context of Node-Place model

Haocheng Sun¹; Michael Batty²; Chen Zhong^{1,2}

¹University College London, Centre for Advanced Spatial Analysis (CASA)

Gower Street, London WC1E 6BT, United Kingdom

ucfnhsu@ucl.ac.uk (correspondent author)

² University College London, Centre for Advanced Spatial Analysis (CASA)

90 Tottenham Court Road, London W1T 4TJ, United Kingdom

44-(0)7768 423 656, m.batty@ucl.ac.uk

^{1,2} University College London, Centre for Advanced Spatial Analysis (CASA)

Gower Street, London WC1E 6BT, United Kingdom

c.zhong@ucl.ac.uk

Keywords: GBA, Spatial Structure, Morphological Polycentricity, Functional Polycentricity, Node-Place Model

For a long time, the study of urban spatial structure, especially the relationship between cities, has been regarded as an important research field in urban planning and regional science. The study of urban spatial structure can be traced back to the beginning of the 20th century, which mainly originated from the theory of urban location (Alonso, 1964). As time enters the 21st century, the concept of monocentricity has gradually given way to polycentricity (Green, 2007; Meijers and Burger, 2010; Batty, 2016). There are various signs that people have perceived that polycentricity is happening in the city, but in fact this concept is vague (Burger and Meijers, 2011a; Brezzi and Veneri, 2014; Giffinger and Suitner, 2014). This ambiguity is mainly reflected in the cognition of the centre and spatial structure.

The Guangdong-Hong Kong-Macao Greater Bay Area (GBA) plan was promulgated by the State Council of China in 2015 (Walker and Schafran, 2015). As one of China's three Mega-City Clusters, GBA is considered an important innovation platform for national implementation of resource allocation, coordination, and division of labour (Hui et al., 2020). It is foreseeable that the spatial structure of GBA will gradually move closer to polycentricity in the future. Therefore, exploring the current spatial structure of GBA will provide a strong basis for a series of urban construction measures in the future. This is the primary motivation of this research from the perspective of practical application.

Therefore, this research will use the Node-Place model proposed by Bertolini in 1999 as an entry point to fill the above-mentioned research gaps in this field. This research proposes an extended model defined as Node-Place-Settlement-Container+Mobility (NPSC+M) to evaluate the different dimensions of urban station areas. Among them, the ‘NPSC’ dimensions are used to evaluate the morphological spatial structure, and the ‘M’ dimension is used to evaluate the functional polycentricity. We provide a multi-dimensional and comprehensive multi-index polycentricity evaluation framework based on this extended model, and further analyse the connections of morphological polycentricity and functionality polycentricity.

1. Visualize and analyse the spatial structure of GBA

In order to quantitatively evaluate the morphological polycentricity and functional polycentricity of GBA and its major cities, this paper processes the collected data to construct 19 indicators in five dimensions of NPSC+M. On this basis, the indicators are weighted by the CRITIC method and integrated into the index of each dimension. Finally, through Global Moran’s I, we put forward a quantitative evaluation result of the Centralized-Dispersed degree for each city in GBA, and visualized their morphological and functional spatial structure through the LISA spatial autocorrelation method to evaluate their polycentricity. We found that currently only Shenzhen in GBA meets the definition of morphological polycentricity and functional polycentricity (Figure 1 and 2). Guangzhou is currently in the process of transforming from monocentricity to polycentricity. Compared with the former two, Foshan and Dongguan have weaker urban competitiveness. This results in them still being monocentric or even dispersed. At the regional level, GBA already possesses morphological polycentricity (although only Guangzhou and Shenzhen), but the flow of functions tends to gather in Shenzhen.

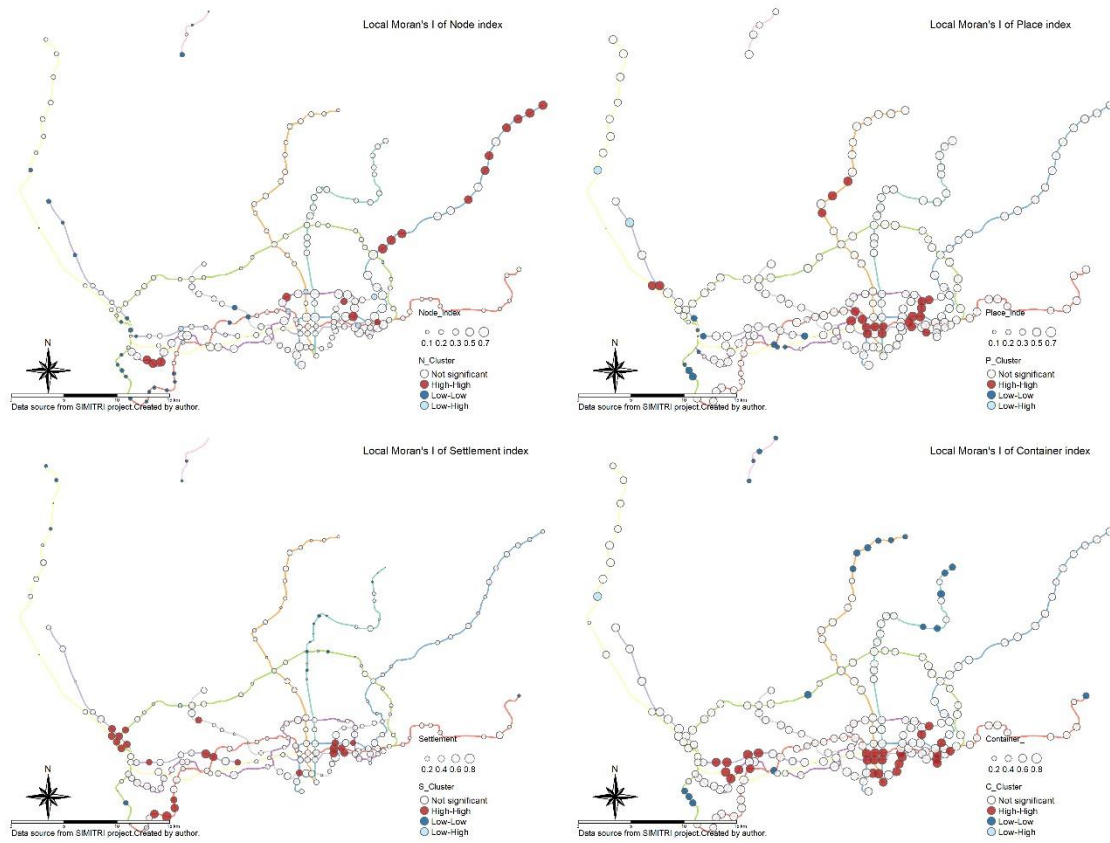


Figure 1: LISA maps of morphological spatial results in Shenzhen, based on the k2 matrix; 1) Node dimensional; 2) Place dimensional; 3) Settlement dimensional; 4) Container dimensional morphological spatial structure

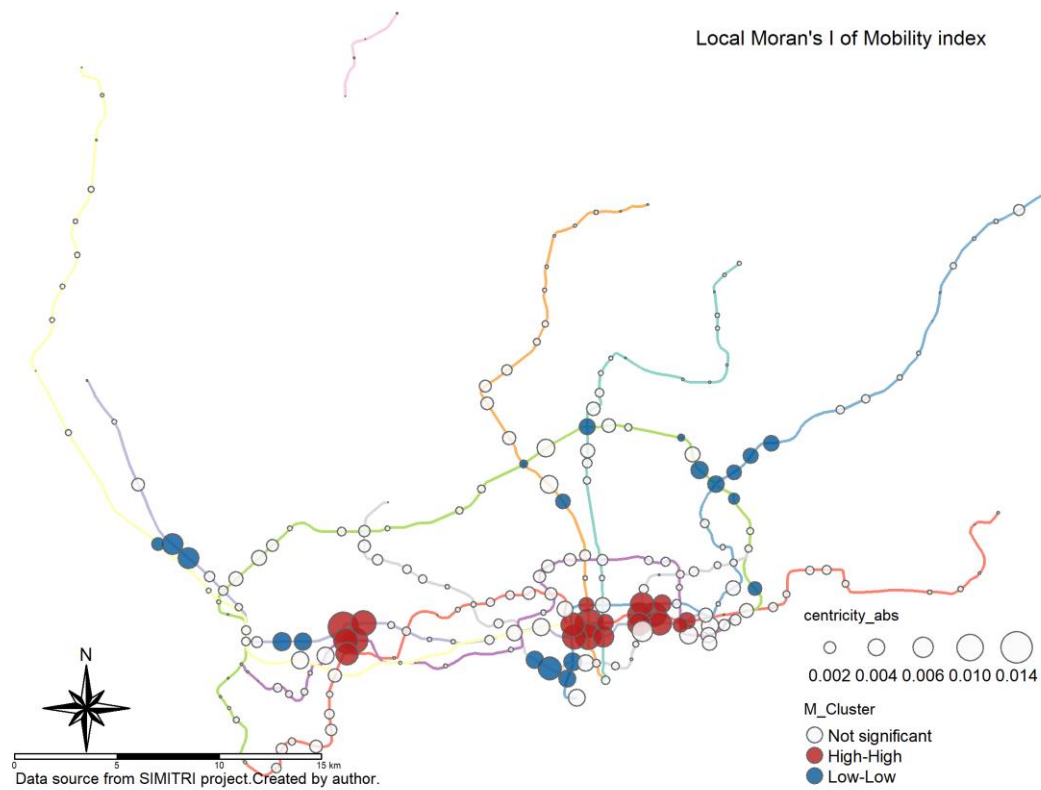
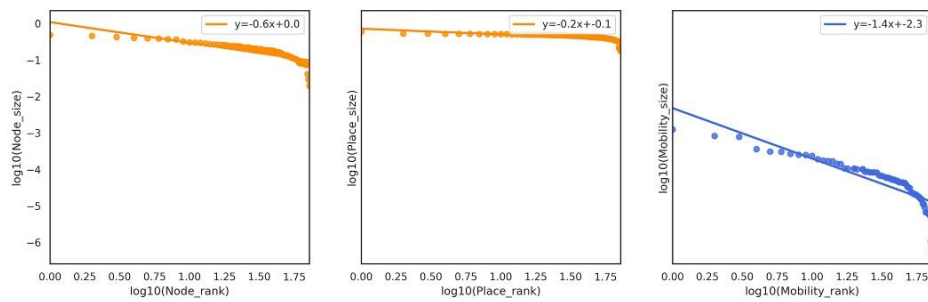
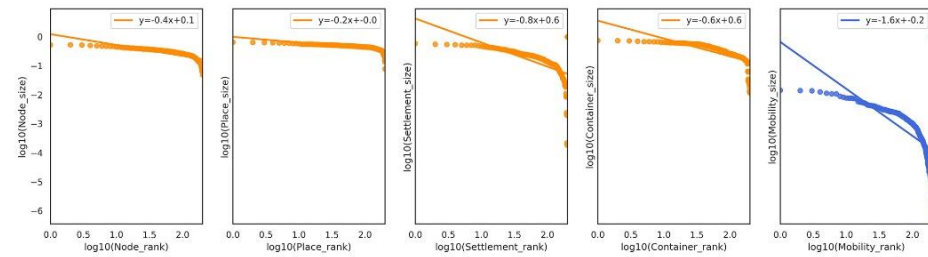


Figure 2: LISA map of Mobility dimensional results in Shenzhen, based on the k2 matrix

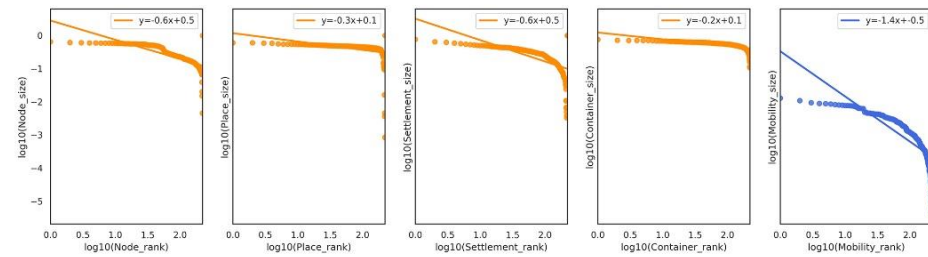
1) Rank-size distribution of Node, Place and Mobility dimensions in GBA



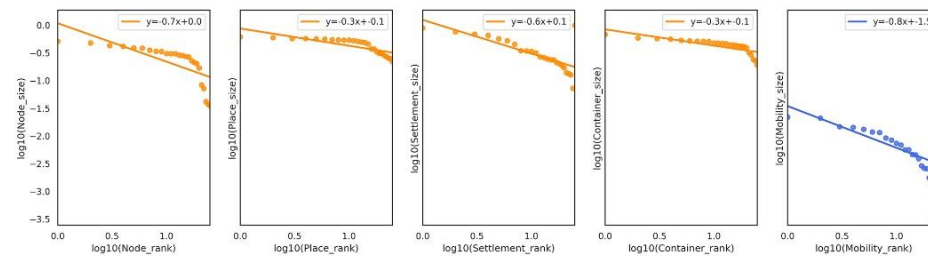
2) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Guangzhou



3) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Shenzhen



4) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Foshan



5) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Dongguan

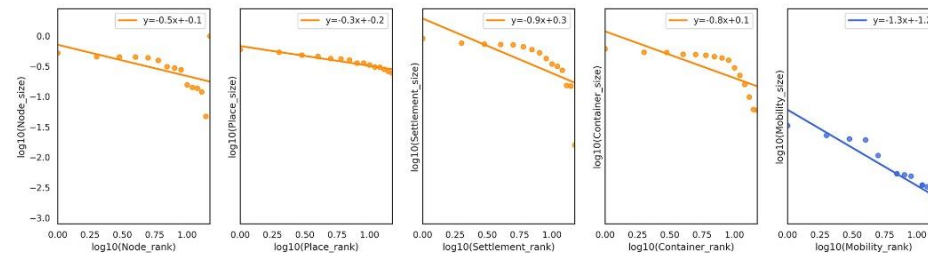
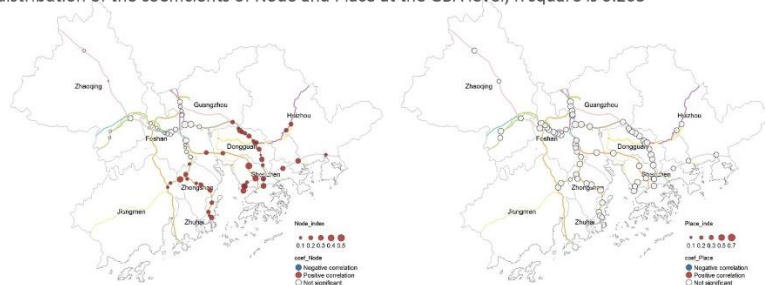
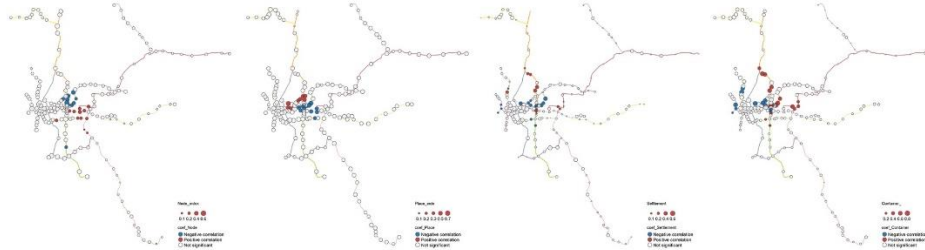


Figure 3: Rank-size distribution of 1) GBA, 2) Guangzhou, 3) Shenzhen, 4) Foshan, 5) Dongguan in NPSC+M dimensions;

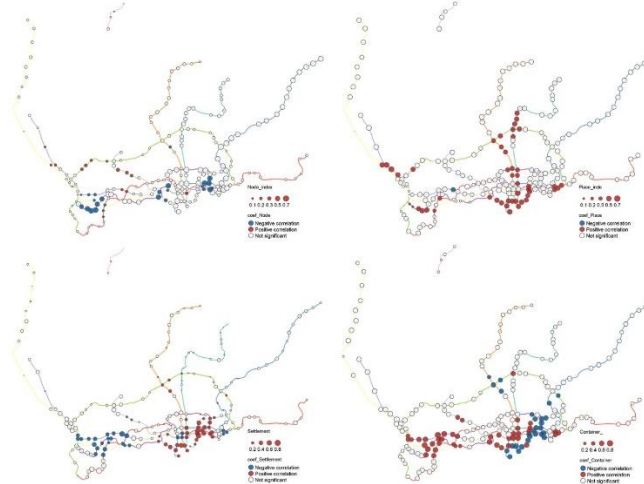
1) The distribution of the coefficients of Node and Place at the GBA level; R square is 0.263



2) The distribution of the coefficients of Node, Place, Settlement and Container at the Guangzhou level; R square is 0.918



3) The distribution of the coefficients of Node, Place, Settlement and Container at the Shenzhen level; R square is 0.924



4) The distribution of the coefficients of Node, Place, Settlement and Container at the Foshan level; R square is 0.467



5) The distribution of the coefficients of Node, Place, Settlement and Container at the Dongguan level; R square is 0.596

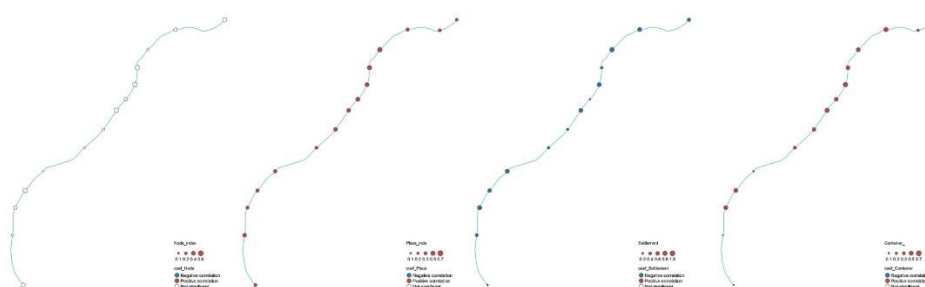


Figure 4: The coefficients of the GWR model are distributed in 1) GBA, 2) Guangzhou, 3) Shenzhen, 4) Foshan, 5) Dongguan. Node, Place, Settlement, Container are used as independent variables, and Mobility is used as dependent variable.;

2. Comparison of morphological polycentricity and functional polycentricity in GBA

On this basis, we further studied the relationship and mismatch between morphological polycentricity and functional polycentricity in GBA. For the former, we chose the Rank-size distribution method which has been confirmed by Meijers (2011) to be feasible. For visualizing the mismatch of morphological polycentricity and functional polycentricity, we tried to use geographically weighted regression (GWR) and confirmed the contribution of this method in this field. we made some conclusions from the results (Figure 3 and 4):

- 1) The development of physical space of GBA is more balanced than the development of functional connection, which is completely contrary to the conclusions obtained in the United Kingdom and the United States (Burger *et al.*, 2011; Arribas-Bel and Sanz-Gracia, 2014).
- 2) Morphological polycentricity and functional polycentricity have been shown to be interdependent. This finding also verifies the conclusion of Meijers (2011);
- 3) For regional sub-central cities (Foshan, Dongguan), Horizontal land use is the key to stimulating functional flow. For regional central cities (Guangzhou, Shenzhen), the flow of functions has a closer relationship with the urban spatial form.
- 4) Focusing on mismatches, the urban centres of Shenzhen and Guangzhou show that morphological elements are not enough to support high- strength functional flow. The main problem in Foshan is that the residential space dominates the vertical spatial form of the city. There is a serious imbalance between the distribution of communities in Dongguan and the flow of population.

In general, the concept of polycentricity is still vague (Möck and Küpper, 2019). Nevertheless, we have boldly attempted to combine the Node-Place model with the polycentricity evaluation and respond to the ambiguity with a fixed spatial unit (station) and multi-scale evaluation method. However, our current method still has many limitations, such as the ambiguity caused by the choice of analysis space unit. We encourage other scholars to look at polycentricity from the perspective of ‘evaluation’ rather than ‘identify’. In this way, not only will polycentricity not become a far-fetched and meaningless concept, but it will also help the city's cognition and reality transformation.

Acknowledgments

The first author would like to express his gratitude to his supervisor, the second author, for his continuous teaching and help. This article benefits from SIMETRI project.

References

- Alonso, W. (1964). 'Location and Land Use'. *Location and Land Use*. Harvard University Press. doi: 10.4159/HARVARD.9780674730854/HTML.
- Arribas-Bel, D. and Sanz-Gracia, F. (2014). 'The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010'. <http://dx.doi.org/10.1080/02723638.2014.940693>. *Routledge*, 35 (7), pp. 980–997. doi: 10.1080/02723638.2014.940693.
- Batty, M. (2016). 'Polynucleated Urban Landscapes': <http://dx.doi.org/10.1080/00420980120035268>. Sage PublicationsSage UK: London, England, 38 (4), pp. 635–655. doi: 10.1080/00420980120035268.
- Brezzi, M. and Veneri, P. (2014). 'Assessing Polycentric Urban Systems in the OECD: Country, Regional and Metropolitan Perspectives'. <http://dx.doi.org/10.1080/09654313.2014.905005>. *Routledge*, 23 (6), pp. 1128–1145. doi: 10.1080/09654313.2014.905005.
- Burger, M. J., de Goei, B., van der Laan, L. and Huisman, F. J. M. (2011). 'Heterogeneous development of metropolitan spatial structure: Evidence from commuting patterns in English and Welsh city-regions, 1981–2001'. *Cities*. Pergamon, 28 (2), pp. 160–170. doi: 10.1016/J.CITIES.2010.11.006.
- Burger, M. and Meijers, E. (2011a). 'Form Follows Function? Linking Morphological and Functional Polycentricity': <http://dx.doi.org/10.1177/0042098011407095>. SAGE PublicationsSage UK: London, England, 49 (5), pp. 1127–1149. doi: 10.1177/0042098011407095.
- Burger, M. and Meijers, E. (2011b). 'Form Follows Function? Linking Morphological and Functional Polycentricity': <http://dx.doi.org/10.1177/0042098011407095>. SAGE PublicationsSage UK: London, England, 49 (5), pp. 1127–1149. doi: 10.1177/0042098011407095.
- Giffinger, R. and Suitner, J. (2014). 'Polycentric Metropolitan Development: From Structural Assessment to Processual Dimensions'. <http://dx.doi.org/10.1080/09654313.2014.905007>. *Routledge*, 23 (6), pp. 1169–1186. doi: 10.1080/09654313.2014.905007.
- Green, N. (2007). 'Functional Polycentricity: A Formal Definition in Terms of Social Network Analysis'. *Urban Studies*. *Routledge*, 44 (11), pp. 2077–2103. doi: 10.1080/00420980701518941.
- Meijers, E. J. and Burger, M. J. (2010). 'Spatial Structure and Productivity in US Metropolitan Areas': <http://dx.doi.org/10.1068/a42151>. SAGE PublicationsSage UK: London, England, 42 (6), pp. 1383–1402. doi: 10.1068/A42151.
- Möck, M. and Küpper, P. (2019). 'Polycentricity at its boundaries: consistent or ambiguous?' <https://doi.org/10.1080/09654313.2019.1666802>. *Routledge*, 28 (4), pp. 830–849. doi: 10.1080/09654313.2019.1666802.

Multiplicative random cascade models of urban structures

T.F. STEPINSKI¹; M. SAEEDIMOGHADDAM²

¹Dept. of Geography and GIS, University of Cincinnati

215 Braunstein Hall, Cincinnati, OH 45221, USA

+15135563583, stepintz@uc.edu (correspondent author)

Dept. of Land, Air and Water Resources, University of California

Davis, CA, 95616, USA

msaeedi@ucdavis.edu

Keywords: modeling urban structure, multifractal pattern, multiplicative cascade, street density

1. Introduction

Cities have spatially complex structures that are best described in terms of multifractal geometry (Benguigui and Czamanski, 2004). The multifractal character of the urban form indicates the presence of a hidden formation process that operates at different spatial scales resulting in the observed structure. Urban spatial models (USMs) use simple, heuristic sets of rules in order to simulate the observed pattern of urban structure, in particular its scaling property. If a USM produces patterns with scaling properties similar to those observed in a real urban structure and visually resembles an actual layout of the modeled city, its heuristic rules may provide an insight into the character of an actual hidden process. USMs based on different heuristics have been proposed in the literature. The feasibility of these models listed above was assessed on the basis of visual comparison of simulated patterns to observed urban structures and comparison of radial profiles of simulated and actual population densities, but not on fractal properties (except for Murcio et al., 2015), or on multifractal spectra.

Here, we investigate the plausibility of using the Multiplicative Random Cascade (MRC) (Stanley and Meakin, 1988) to model urban spatial structure. The MRC has some desirable properties that other USMs lack. (1) It is a priori known to produce a multifractal pattern. (2) It is expressed in an analytic form which makes it possible to find the best-fit model. Using street intersection points pattern (SIPP) data as a proxy of actual urban structure we investigated the following. (1) Similarity of observed Renyi's spectrum to the best-fit spectrum of MRC pattern. (2) Similarity of MRC-

modeled probability distribution function (PDF) of local SIPP densities to PDFs of local densities in actual urban structures. (3) Layout of MRC model versus actual layout.

name (D_0^{emp})	MRC multinomial model			MRC binomial model			
	s_0 (D_0)	w_1, w_2, \dots, w_{s_0}	RMSE	s_0	s_1	w_1	RMSE
Knoxville (1.37)	7 (1.4)	0.316,0.279,0.091,0.091,0.091,0.091,0.04	0.029	7	2	0.3133	0.048
Philadelphia (1.37)	7 (1.4)	0.276,0.245,0.107,0.107,0.107,0.107,0.051	0.025	7	3	0.2529	0.048
Oklahoma City (1.58)	9 (1.58)	0.254,0.207,0.197,0.079,0.077,0.068,0.056,0.04,0.022	0.016	9	3	0.2377	0.034
Orlando (1.55)	9 (1.58)	0.223,0.193,0.186,0.166,0.053,0.049,0.049,0.049,0.033	0.019	9	3	0.2124	0.026
Portland (1.42)	7 (1.4)	0.207,0.190,0.171,0.156,0.148,0.064,0.064	0.007	7	4	0.1999	0.028
Phoenix (1.54)	9 (1.58)	0.207,0.173,0.172,0.171,0.066,0.064,0.062,0.06,0.025	0.012	9	3	0.1979	0.029

Figure 1: Table showing best-fit values of MCR models parameters to account for SIPP's pattern in six US cities.

2. MRC model

The MRC model is an iterative algorithm for placing Ω points in a unit square. At each iteration step points are randomly redistributed into s equal-size square subdivisions of the parent square using redistribution probabilities $\{w_1, \dots, w_s\}$, where $\sum w_i=1$. This process is repeated for each subdivision for n rounds of iteration. The end result is the probability distribution function (PDF) $\{p_i\}_\varepsilon = \{p_1, \dots, p_n\}_\varepsilon$ where n is the number of subdivisions (boxes) at iterative step ε . Placing points according to this PDF results in the multifractal pattern whose appearance depends on values of w_i . We compare PDFs, Renyi's spectra D_q , and layouts of MRC and SIPP (observations) patterns. D_q describes scaling properties of the pattern; q is the moment order, D_{qj} describes scaling of the part of the pattern consisting of N_{qj} boxes with the highest point density. D_0 is a fractal dimension. D_q for the MRC has a closed analytic form; we refer to this form as the MRC multinomial model. We also consider a simplified analytic form to which we refer as the binomial MRC model. In order to account for variations of SIPP patterns we consider MRC with $s=16$ and regulate the fractal dimension of the MRC pattern by changing the number (s_0) of nonempty boxes. For details see Saeedimoghaddam and Stepinski (2020).

2. Results

We fitted MRC model-generated D_q curves to the empirical D_q^{emp} curve derived from SIPP from six cities in the US (United States Census Bureau, 2017). D_q^{emp} is calculated using the box-counting method. The results are shown in Figure1. The table in this figure shows the best-fit parameters of MRC models, multinomial model offers a better fit (smaller value of RMSE) but binomial model is easier to interpret. SIPP-derived curves (D_q^{emp}) and best-fit MRC generated curves D_q^{emp}

are shown in Figure 2. Figure 2 and the values of RMSE in the table indicate a good fit - the MRC model can account for multifractality of urban forms.

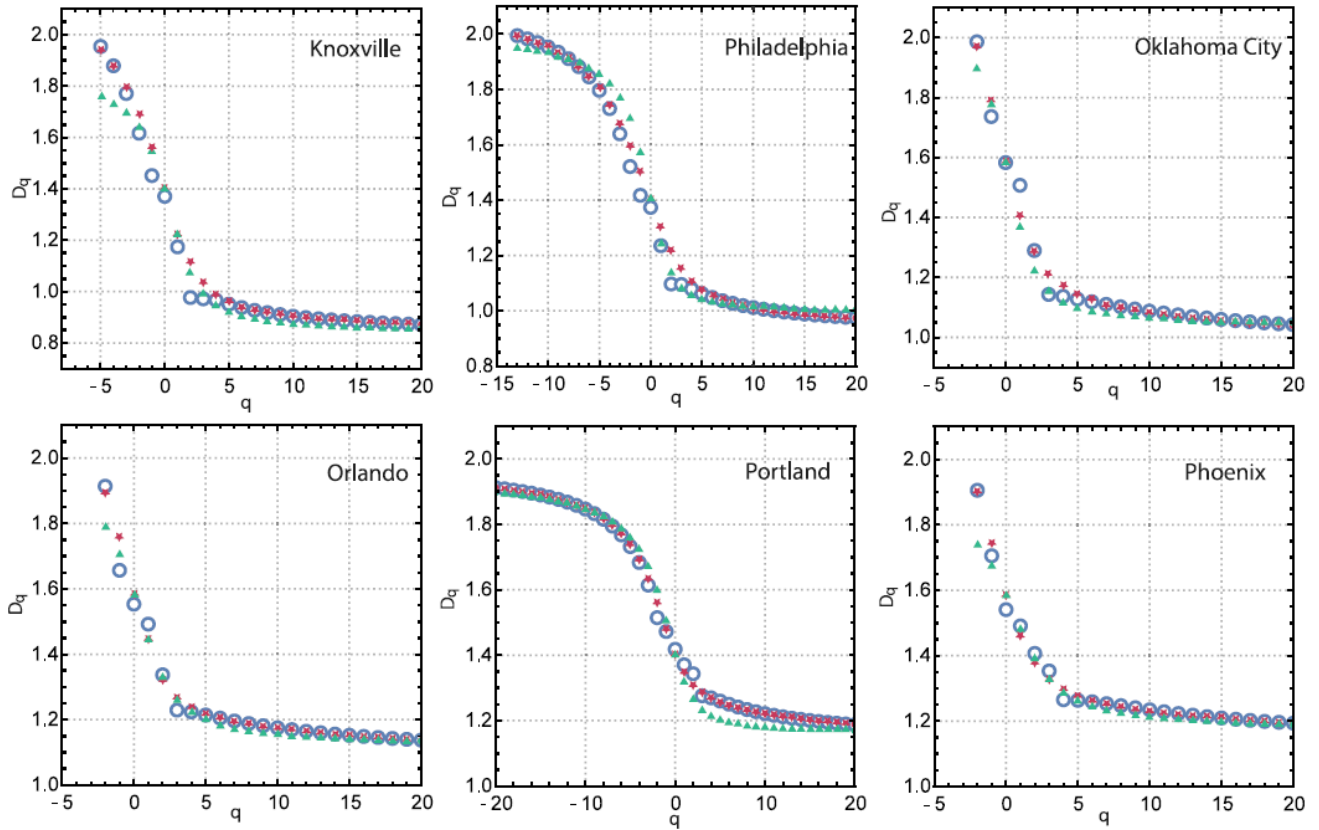


Figure2: Comparison of empirical Renyi's spectra of SIPPs in six MSAs with their best-fit MRC models. Blue circles represent empirical spectra, red stars represent multinomial models, and green triangles represent binomial models.

Interpretation of binomial model is straightforward For example, the model with $s_0=9$, $s_1=3$, $w_1=0.2$ has the following interpretation. Because SIPP is a proxy for a density of urban development, this model encapsulates the situation when 7/16 or 44% of the city's area is undeveloped, 3/16 or 19% has a higher level of development, and 6/16 or 37% has a lower level of development (1/3 of the development of the higher level). This redistribution of development also holds for all subsequent subdivisions of the city. According to our results, Oklahoma City, Orland, and Phoenix have such urban structures. The remaining three cities have different, more disaggregated urban structures.

Another test of the model is to see whether a distribution of intersection density generated by the model is comparable to the actual distribution of SIPP. The results of such a comparison are shown in Figure 3. Analysis of this figure reveals the following. (1) Modeled distributions can be closely

fitted to the log-normal distribution.. (2) Observed density distributions are not log-normal, although in most cases (Portland and Phoenix are exceptions) they are close to being log-normal. (3) Observed and modeled density distributions are different, although, in the cases of Phoenix and Philadelphia the difference is not profound. In all cases the model predicts few boxes with densities higher than observed; these values are unrealistic, they are present because the model does not have an upper threshold. In four out of six cases (Philadelphia and Portland are exceptions) model predicts more low-density cells than are actually present. It also predicts the existence of cells with unrealistically low densities (smaller than 1) although their predicted numbers are low.

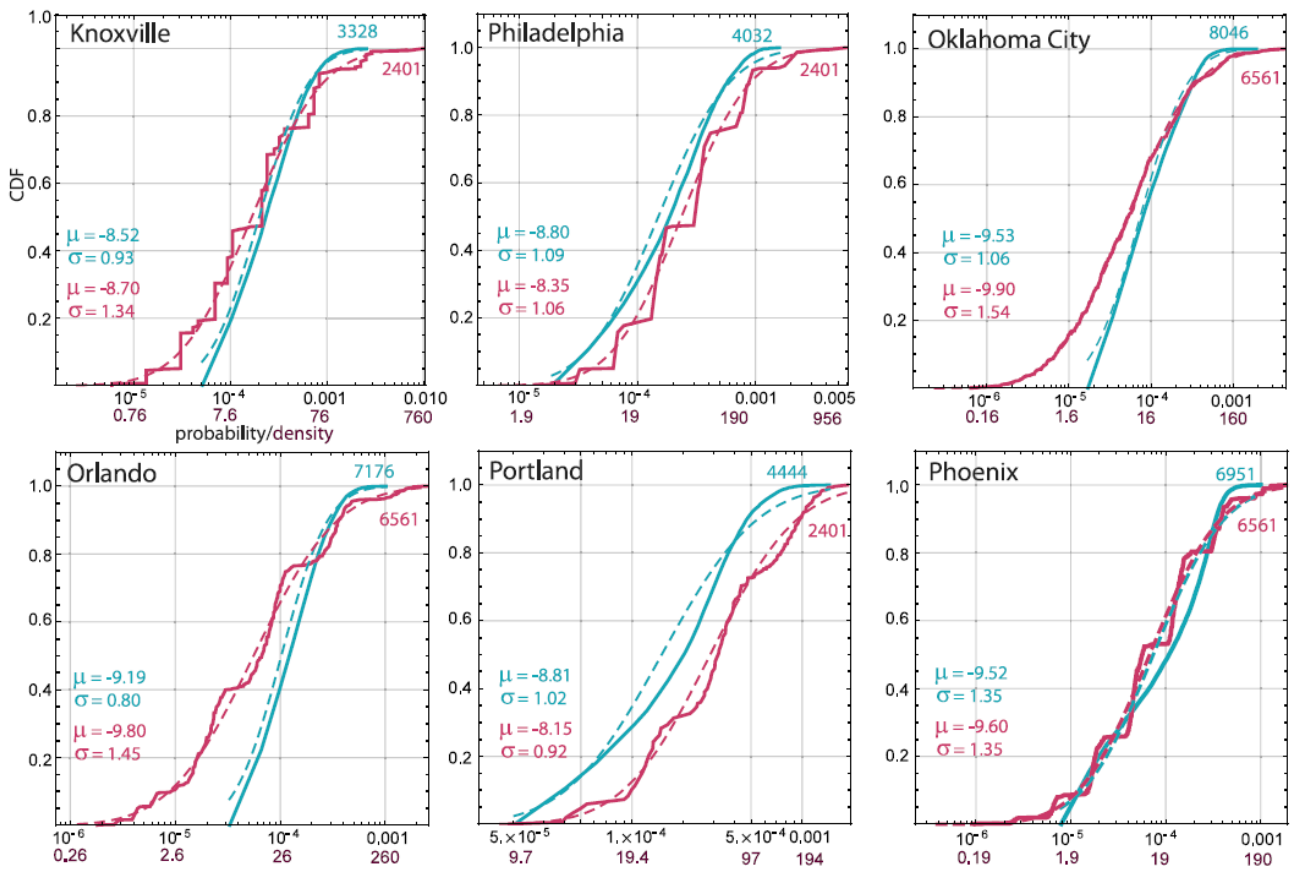


Figure 3: Cumulative Distribution Functions (CDFs) of observed intersections densities (green solid lines) and modeled intersections densities (red solid lines) for grid/lattice characterized by $\epsilon = 1/256$. The number of nonempty boxes is shown next to each CDF curve. Also shown (by dashed lines) are best-fitted CDFs of log-normal distributions. Parameters of best-fitted log-normal distributions are listed. Values of probabilities (black) and densities (brown) are displayed on the logarithmic scale. The unit of density is $(\# \text{ of intersections})/\text{km}^2$

The layout of the MRC pattern depends on the values of redistribution coefficients and a rule on which coefficient applies to which daughter cells in a redistribution process. Renyi's spectrum does not depend on the allocation rule, but a pattern's layout does. In the default MRC model, redistribution coefficients are assigned randomly to daughter boxes which result in a layout not similar to

those displayed by urban structures. Figure 4 shows layouts in Oklahoma City stemming from modifications of default MRC.

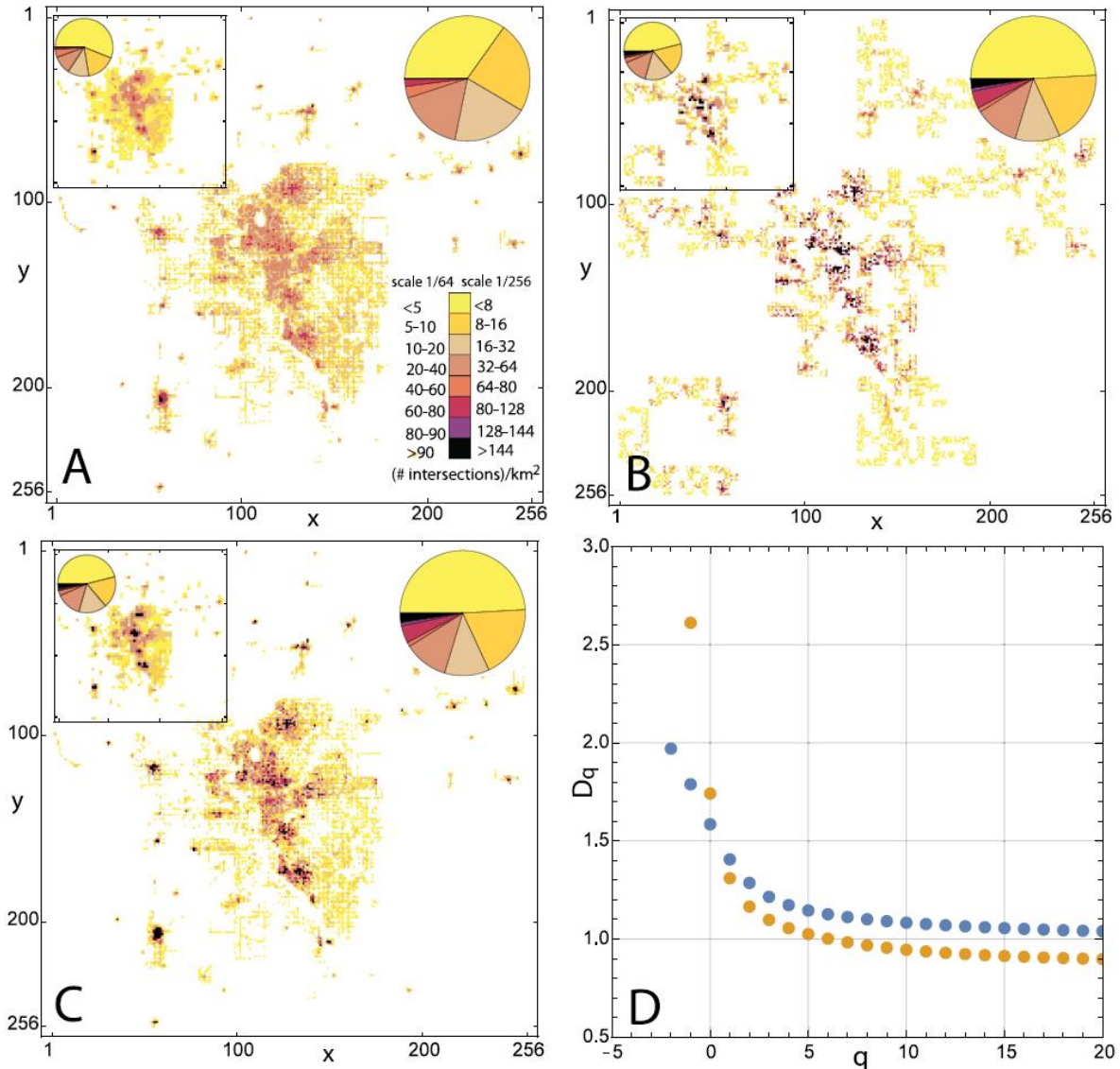


Figure 4: Layouts of intersection densities in Oklahoma City. (A) Observed layout. (B) The guided spatial allocations layout (C) Simulated layout. Pie diagrams show shares of boxes/cells of different densities as listed in the legend. Main panels are for $\epsilon = 1/256$, insets are for $\epsilon = 1/64$. (D) Renyi's spectrum of the multinomial MRC model (blue) and the simulated model (orange).

In this figure panel (A) shows the observed layout. Panel B shows the layout of the MRC-modeled densities with allocation guided by observation. This layout resembles observations only in the most rudimentary features. Panel C shows a pattern constructed by wrapping modeled values of intersection densities over the observed layout. This process does not preserve redistribution

coefficients of the MRC model, but its spectrum is quite similar to the spectrum of the MRC model (Figure 4D).

4. Conclusion

The MRC model of urban structure can reproduce multifractal spectra of actual urban structures as given by SIPP. To the best of our knowledge, this is the first heuristic model to do so. However, the density distribution given by the MRC departs from actual density distributions. The model has a log-normal distribution and it overpredicts the number of boxes with low and very high densities. This is because the model does not have density thresholds on both, low and high end of the distribution. Without additional information, the model's layout (not shown) does not resample the observed layout, but the model modified by providing configuration information produces a layout similar to the observed layout.

References

- Benguigui, L., Czamanski, D. (2004). Simulation analysis of the fractality of cities, *Geogr. Anal.* 36 (1) 69–84.
- Murcio, R., Masucci, A.P., Arcaute, E., Batty, M. (2015). Multifractal to monofractal evolution of the London street network, *Phys. Rev. E* 92 (6), 062130.
- Saeedimoghaddam, M., Stepinski, T.F. (2020). *Physica A* 569, 125767.
- Stanley, H.E., Meakin, P. (1988). Multifractal phenomena in physics and chemistry, *Nature* 335,405–409.
- United States Census Bureau. (2017)
(<https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2017&layergroup=Roads>)

Parallel Session PSC3

4 November 11:20 am-1:20 pm (GMT)

Mobility and Distance

Congestion and Parking Prices in a Shared Automated Vehicle Future: A Case Study of Jerusalem

Golan Ben-Dor¹; Ido Klein¹; Aleksey Ogulenko¹; Eran Ben-Elia²; Itzhak Benenson¹

Porter School of the Environment and Earth Sciences, Tel Aviv University

PO Box 39040 Ramat Aviv Tel Aviv, 69978, Israel

+972522937463, golanben@mail.tau.ac.il (corresponding author)

² Department of Geography and Environmental Development, Ben-Gurion University

PO Box 653 Beer Sheva, 84105 Israel

Keywords: Shared-Automated-Vehicles, Congestion Pricing, Parking Pricing; Agent-Based Simulation, MATSim

1. Introduction

Public Transport (PT) has entered an era where automated vehicles can potentially turn Demand Responsive Transit (DRT) into the leading transportation mode (Litman, 2017). Our goal is to assess this future regarding Shared Automated Vehicle (SAVs) service in the Jerusalem Metropolitan Area (JMA). According to a recent publication (Schaller, 2021), most of the ride-hailing users in the US were former PT users, and AVs are unlikely to change auto users' resistance to shared services. A possible solution to enforce car users to move to shared services is to impose congestion and/or parking prices (AbuLibdeh, 2017). This paper evaluates the consequences of a new PT-focused transport policy in the JMA that includes parking and congestion pricing and introduces a new SAV transportation mode as a possible game-changer of the existing equilibrium between PT and private car use. We investigated the effects of the SAV fleet operation parameters: size and location of the service area, fleet size, and service priorities. Our overall goal is to propose a balanced set of carrot-and-stick measures that could enforce sustainable modal shifts that can reduce congestion during the planned development of the JMA PT network between 2020 and 2040.

2. The Jerusalem Metropolitan Area MATSim model

We employ an open-source MATSim modeling framework (Horni et al., 2016), a leading agent-based multimodal transportation simulation system. MATSim accepts urban infrastructure data as a set of GIS layers and transportation demand in the form of individual agent's plans of their daily activities. MATSim simulates urban traffic focusing on individual travel choices and behaviour: agents adapt to the existing level-of-service conditions through reinforced learning and self-correction while the entire transportation system converges to a steady User Equilibrium (UE). We

apply the MATSim downscaling procedure (one agent represents several) and adjust model parameters to correctly represent the full-scale dynamics (Ben-Dor et al., 2020).

The JMA MATSim application (JMATSIm, Figure 1) is investigated with 30% of the metropolitan population (434K agents) representing three major JMA social groups: Secular Jews, Ultra-Orthodox Jews, and East Jerusalem Arabs. 96K additional agents traveling to/from JMA to the rest of the country represent external traffic. The JMA road network is defined by 8.5K links and 3.5K junctions, covering 1800 km². In JMATSIm, we modified the mobility module to account for turn penalties by applying the Kirby-Potts expansions (Kirby & Potts, 1969).

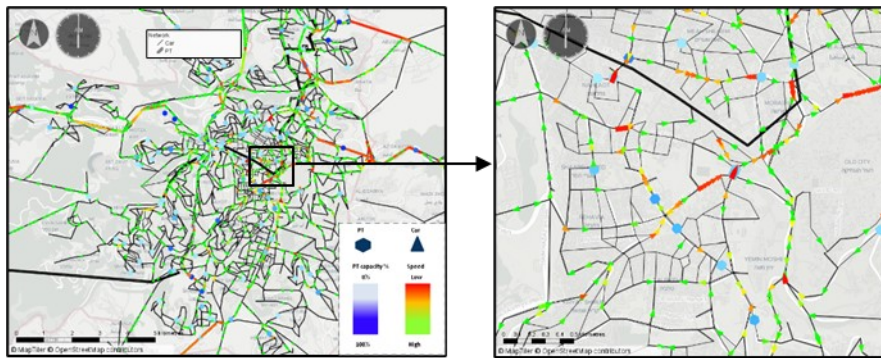


Figure 1: Snapshot of the JMATSIm at 06:00.

The JMATSIm transportation is multimodal, with private cars, bikes, buses, heavy trains, light rail, and walking. Agents' mode-dependent utilities are established by adjusting the Berlin scenario values (Ziemke et al., 2018). The model is calibrated by applying the Cadyts algorithm (Flötteröd, 2009) against 1K traffic counts collected during 2012-2020 and data on public transport use. The resulting match is very good, with $R^2 = 0.84$ between the car counts, and a similar fit is obtained for the PT data, e.g., model and smartcard-based light rail boardings (Figure 2).

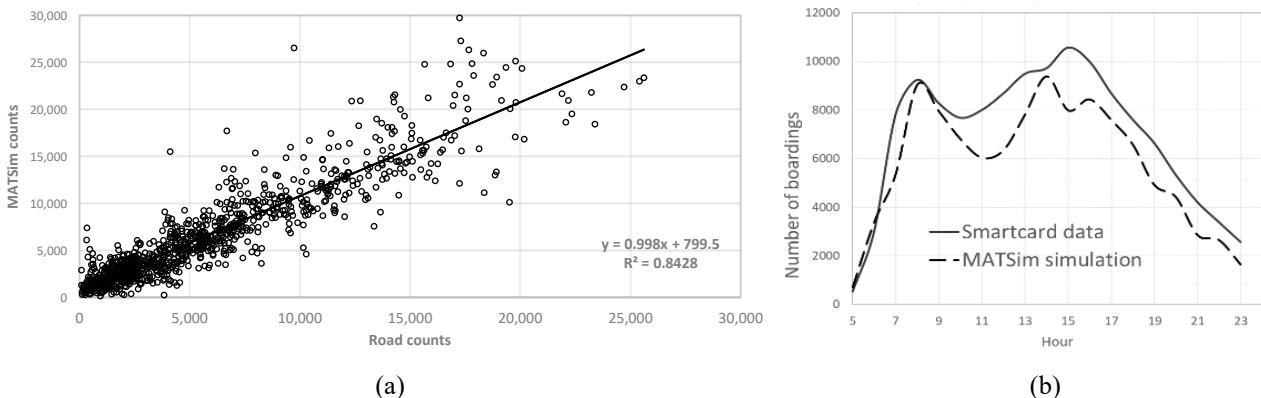


Figure 2: Calibration of the JMA model against (a) car traffic counts; (b) light train boardings smartcard data.

The calibrated version of the model properly reproduces JMA modal split (Table 1).

Table 1. Reproduction of the modal split

Mode	PT	Car	Ride	Walk	Bike
JMA Travel survey	21.28%	32.40%	8.00%	37.71%	0.60%
JMATSim	21.99%	33.09%	8.00%	35.83%	1.09%

3. Scenarios and results

Our scenarios compare the JMA congestion and modal split traffic as nowadays when there is no congestion price and parking prices established for the visitors of the core city area, to the system's state simulated with the JMATSIm for various congestion and parking pricing schemes, and SAV fleets of various sizes (250-1000). Importantly, we limit the SAV services area to the dense central core of $\approx 5 \text{ km}^2$ of Jerusalem (Figure 3, red). The price of the SAV trip is set equal to that of the PT, and SAV fleet dispatching is managed by a high-performance DRT algorithm (Bischoff et al., 2017; Zwick et al., 2020). We investigate the effect of (1) a fixed congestion charge for entering the center of Jerusalem with a private car and (2) time-based parking pricing there, having the overall goal to preserve travelers' flows to and from the city center. To determine the flows for given congestion prices, parking prices, and SAV fleet, we run the JMATSIm simulation and consider the persistent flows and modal split in the model's UE. We characterize road links as congested when the Volume-to-Capacity (V/C) ratio for this link is above 0.8 (Elefteriadou, 2016).

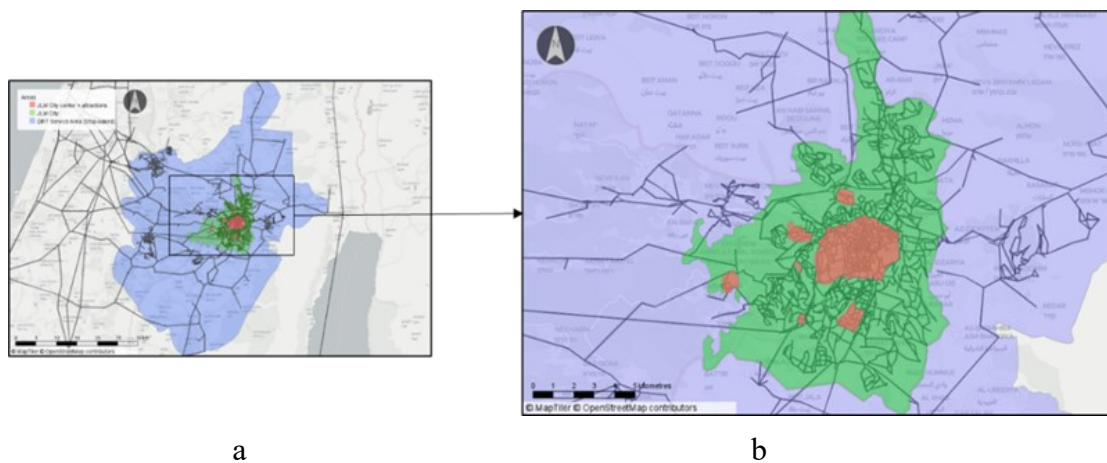


Figure 3: JMA (violet), Jerusalem city (Green) and City center and attraction areas (red), distant (a) and close (b) zoom.

Figure 4 presents the dependence of the number of private cars entering the center of Jerusalem and the percentage of links for which the Volume-to-Capacity ratio $V/C > 0.8$ on the congestion price for the scenarios when neither parking price nor SAV service is activated.

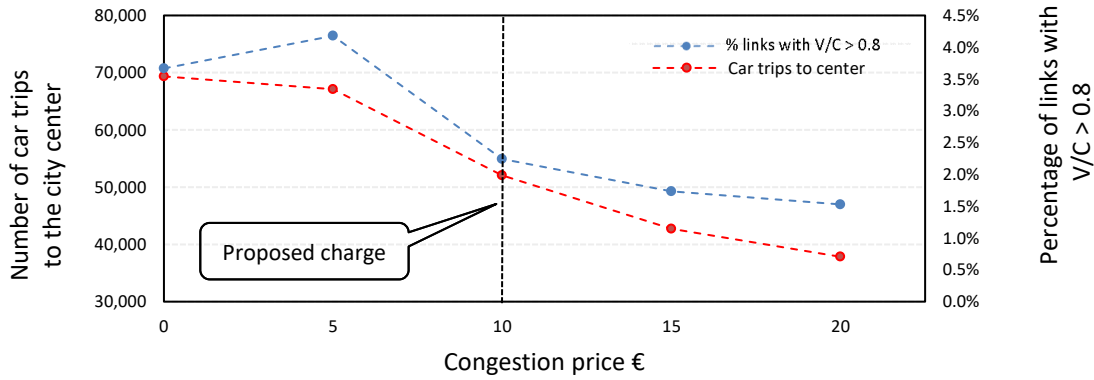


Figure 4. Dependence on the number of cars entering Jerusalem city center and the fraction of the links in the city center congested in the morning on the congestion price, in scenarios with parking prices and SAV service deactivated.

We thus propose a practically acceptable daily congestion price at a level of $\approx 10\text{€}$. This congestion price reduces the number of cars entering the city center by 15% and congestion by 50%.

Figure 5 shows the shift in a modal split among the travelers arriving at the city center induced by congestion or parking price in the case of the 250 SAV fleet. The left side of each Sankey diagram shows a modal split for the scenario with no SAVs, while the right side shows a modal split for the case of 250 SAVs serving core-periphery trips. In the case of zero congestion/parking price, the users of SAV are the former PT users almost exclusively (Figure 5a). In case of parking or congestion price, PT usage is preserved, and a significant share of travelers who traveled to a city center with private cars move to the SAV service (Figure 5b,c). Qualitatively, 5€ hourly parking price has a similar impact to a one-time 10€ congestion pricing, while the latter attracts more car users.



Figure 5: Sankey diagrams of modal split change only for the city center with 250 SAVs (*core-periphery SAV service*); (a) 250 SAVs, 0€; (b) 250 SAVs, 5€/h parking price and (c) 250 SAVs, 10€ congestion price.

To conclude, the future of the SAV services in cities like Jerusalem is defined by the city's policy in regards to the congestion/parking prices. Until these prices are imposed, the SAV vehicles will only attract PT users and add to the existing congestion. With the adequately established congestion/parking prices, SAV becomes a reasonable alternative for car users. We investigated the consequences of this understanding with the scenarios of the JMA transportation system development (additional light rail and metro lines) towards the year 2040. The full results of the study will be presented at the conference.

Acknowledgments

This research was supported by the Jerusalem Transportation Master Plan Unit and partially supported by Research Grant No. 3-15741 of the Ministry of Science & Technology, Israel, and The Ministry of Science & Technology of the People's Republic of China.

References

- Cecil, P. A. (1920). *The Economics of Welfare*. New Brunswick & Londres: Transaction Publishers, 4.
- AbuLibdeh, A. (2017). Traffic congestion pricing: Methodologies and equity implications. *Urban Transp. Syst*, 203-227.
- Ben-Dor, G., Ben-Elia, E., & Benenson, I. (2020). Population Downscaling in Multi-Agent Transportation Simulations: A Review and Case Study. *Simulation Modelling Practice and Theory*, 102233.
- Bischoff, J., Maciejewski M., & Nagel K. (2017) City-wide shared taxis: A simulation study in Berlin. *Intelligent Transportation Systems, ITSC 20*, 275-280.
- Flötteröd, G. (2009) Cadyts a free calibration tool for dynamic traffic simulations. Paper presented at the 9th Swiss Transport Research Conference.
- Elefteriadou, L. A. (2016). *The highway capacity manual 6th edition: A guide for multimodal mobility analysis*. *Itc journal*, 86(4).
- Horni, A., Nagel K., & Axhausen, W. K. (2016). *The multi-agent transport simulation MATSim*, Ubiquity Press London
- Kirby, R., & Potts, R. (1969) The Minimum Route Problem for Networks with Turn Penalties and Prohibitions. *Transp. Res.*, 3, 397–408.
- Litman, T. (2017) *Autonomous vehicle implementation predictions*, Victoria Transport Policy Institute Victoria, Canada.
- Mavrotas, G. (2009) Effective implementation of the ϵ -constraint method in Multi-Objective Mathematical Programming problems, *Applied Mathematics and Computation*, Volume 213, Issue 2, 455-465. doi:10.1016/j.amc.2009.03.037.
- Schaller, B. (2021). Can sharing a ride make for less traffic? Evidence from Uber and Lyft and implications for cities. *Transport policy*, 102, 1-10.

Which Regions Benefit from Rail Accessibility? Germany 1990-2030

Submitted to ECTQG2021

SS05: Geographical distances: measuring, mapping, theorizing

Fabian Wenner¹; Johannes Moser²; Alain L'Hostis^{1,2}

¹Chair of Urban Development, TU Munich

Arcisstr. 21, 80333 Munich, Germany

+49 (89) 289 - 22142, f.wenner@tum.de

²Chair of Urban Development, TU Munich

Arcisstr. 21, 80333 Munich, Germany

+49 (89) 289 - 22389, johannes.moser@tum.de

^{1,2}FLVMT, Université Gustave Eiffel, IFSITAR, Ecole des Ponts

F-77454 Marne-la-Vallée, France

+33 (0)3 20 43 84 98, alain.lhostis@univ-eiffel.fr

Keywords: High-Speed Rail, Visualisation, Accessibility, Time-Space, Transportation

The relation of time to space is ever changing. Recent innovations in transportation technologies such as High-Speed Rail (HSR) instantiate abrupt alterations on this relation. Within confined space constructs, the advent of HSR can cause distortions in the overall system of relational proximities. Usually, a HSR network is extended gradually with partial links that renders it a difficult task to meaningfully visualise the overall change in the time-space fabric. Therefore, better understanding of methods to operationalise the visualisation of effects of new transportation infrastructure on the time-space dimension is needed. This paper seeks to examine these methods through their strengths and weaknesses using the undertaken and planned German HSR network expansion from 1990 to 2030. The methods employed are choropleth, anamorphosis, spring and shriveling maps. Choropleth maps are the most widespread cartographic maps on a 2-dimensional scale visualising data often on administrative spatial units such as countries etc., using colours for differentiation. Anamorphic maps alter the visual representation (of administrative spatial units) by distorting their actual shapes with respect to the size of the variable under study. Spring maps, invented by Plassard and Routhier (1987) as well as Tobler (1987), show transport networks with fixed locations as nodes and edges (connections between nodes) in the form of springs that are the more sinuous the longer the (travel-time) distance is. Recent developments of the model have been made by Buchin et al. (2014). Shriveling maps, conceptualised by Mathis (1990) and implemented by L'Hostis (1996, 2009), also keep locations fixed

at their geographical position and draw connections with length proportional to travel time by using the third dimension.

We focus on the assessment and comparison of these four methods along the four properties of geographical time-space as identified by L’Hostis and Abdou (2021):

- 1) The acceleration causing the shrinking of geographical time-space with the introduction of faster transportation means over historical time.
- 2) The use of transportation networks to perform movements in geographical space.
- 3) The co-existence of transport modes characterised by different speeds.
- 4) The spatial inversion Tobler (1961) and Bunge (1962), caused by extreme forms of detour, where a journey starts in a direction opposite to the end destination; this phenomenon happens at all geographical scales, with for instance expressways, rail transport and aerial transport.

In Table 1, the feasibility of producing the respective maps with respect to the properties is depicted.

	1 Acceleration	2 Transport networks	3 Co-existence of modes	4 Spatial inversion
Choropleth map	Possible	Possible	Possible	No
Anamorphosis	Possible	No	No	No
Spring map	Possible	Yes	Difficult	No
Shriving	Possible	Yes	Yes	Yes

Table 1: How different types of representation convey the four geographical time-space properties

As regards the spatial inversion property, it becomes evident why shriving maps, which are the only 3-dimensional model class out of these four map types, are superior to the others.

In Wenner and Thierstein (2021) and Wenner and Moser (2020), the spatial distribution of regional rail accessibility in Germany between 1990 and 2030 is analysed. Deployed measures are potential accessibility of the population, closeness centrality, betweenness centrality and daily accessibility, which are all displayed as choropleth maps. They also produce time-space analyses in the form of anamorphosis maps. These studies show that there has been a ‘shrinkage’ of time-space relations along the axes of HSR. This means in particular the polarisation of space with relations between agglomerations becoming better connected. At the same time, with the abandonment of many regional and local connections, the periphery becomes disintegrated with the loss of long-distance connections. These results are in light with Spiekermann and Wegener (1994), who predict a shrinking of space across regions in overall but warn of precisely this kind of polarisation. However, these studies have not yet harnessed the strengths of spring and shriving maps. A new synthesis of these methods and

a more advanced version of the underlying data will help to understand the actual development of the regional rail accessibility and the change in interregional travel times from 1990 to 2030 in Germany as well as demonstrate the weaknesses and strengths of the various mapping methods.

To gain insights as regards the anamorphosis maps, consider figures 1-5. This series of figures (1-4) shows how German time-space has evolved from 1990 to 2020 with the development of the rail network and schedules. Should all projects of the federal ministry of transport be implemented as planned, the time-space map of 2030 can be depicted as in figure 5. It becomes evident how especially between 1990 and 2000 time-space on the former inner-German border shrunk with an enormous improvement of the rail network in the previous German Democratic Republic by integrating it into the West German network. Between 2000 and 2010, the introduction of HSR between Köln and Frankfurt am Main led to a shrinkage on this relation. In the decade 2010-2020, HSR was implemented between Nürnberg and Leipzig, which brings these regions closer together. Finally, in 2030, the overall time-space will be further compressed as there are many small-scale improvements causing travel time reductions across the whole country.

What the German long distance transport system with rail and air networks looks like in 2020 using a shriveling map is presented in figure 6. In order to reveal the shape of the three dimensional structure, two angles of the view of the same time-space, Germany 2020, are proposed at the same time-space scale. The only straight line, following the geodesic, is the direct air connection between Hamburg and Munich, at 450 kph, in blue the fastest connection in the 2020 German time-space. The slope of cones is determined by the ratio between the fastest speed and the average road speed at 80 kph.

The graphical message of the representation is that high speed rail is very efficient in time-space, performing similarly to air connections, that appear much slower than the maximum aircraft speed due to the short kilometer distances between cities in the German space.



Figure 1: Anamorphosis map of Germany in 1990



Figure 2: Anamorphosis map of Germany in 2000



Figure 3: Anamorphosis map of Germany in 2010



Figure 4: Anamorphosis map of Germany in 2020



Figure 5: Anamorphosis map of Germany in 2030

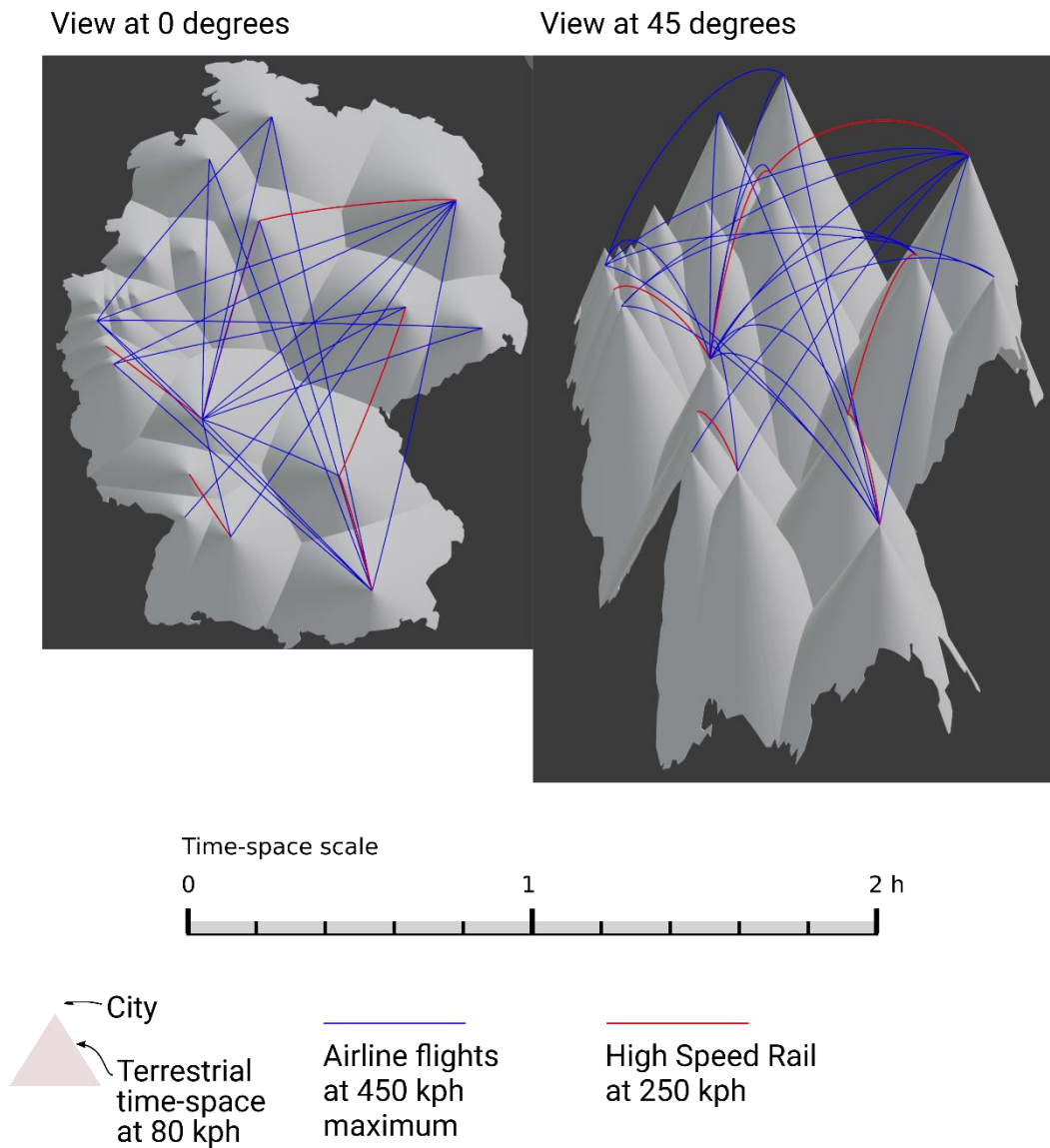


Figure 6: Shriveling map of Germany in 2020

Acknowledgments

There is no specific funding involved for this project.

References

Buchin, K., van Goethem, A., Hoffmann, M., van Kreveld, M., Speckmann, B., 2014. Travel-Time Maps: Linear Cartograms with Fixed Vertex Locations, in: Duckham, M., Pebesma, E., Stewart, K., Frank, A.U. (Eds.), *Geographic Information Science*, Springer International Publishing, Cham. pp. 18–33.

Bunge, W., 1962. *Theoretical Geography*. seconde éd. augmentée 1966 ed., Gleerup, Lund.

L'Hostis, A., 1996. Transports et Aménagement du territoire: Cartographie par images de synthèse d'une métrique réseau. *Mappemonde*, 37–43.

L'Hostis, A., 2009. The shrivelled USA: Representing time-space in the context of metropolitanization and the development of high-speed transport. *Journal of Transport Geography* 17, 433–439.

L'Hostis, A., Abdou, F., 2021. What Is the Shape of Geographical Time-Space? A Three-Dimensional Model Made of Curves and Cones. *ISPRS International Journal of Geo-Information* 10, 340

Mathis, P., 1990. Espace et graphe, le p-graphe t-modal 1-planaire, in: Table ronde ASRDLF "Distance et analyse spatiale", ASRDLF, Chamonix. p. 10

Plassard, F., Routhier, J.L., 1987. *Sémiologie Graphique et Évaluation*. A.R.T.U.R., Lyon.

Spiekermann, K., Wegener, M., 1994. The Shrinking continent: New time-space maps of Europe. *Environment and planning B.: planning and design* 21, 653–673

Tobler, W.R., 1961. *Map Transformation of Geographic Space*. Geography. University of Washington, Geography. Washington.

Tobler, W.R., 1997. Visualizing the impact of transportation on spatial relations, in: Western Regional Science Association Meeting, Western Regional Science Association, Hawaii. p. 7.

Wenner, F., & Moser, J. (2020). Which Regions Benefit from Rail Accessibility? Germany 1990-2030.

Wenner, F., & Thierstein, A. (2021). Rail Accessibility in Germany: Changing Regional Disparities between 1990 and 2020. *Raumforschung und Raumordnung. Spatial Research and Planning*, 79(2), 95-115.

Economic and Environmental Impacts of Dedicated Freight Corridors in India

Satoru KUMAGAI¹, Kenmei TSUBOTA^{1,2*}, Toshitaka GOKAN¹,

Kazunobu HAYAKAWA¹, Ikumo ISONO³, and Souknilanh KEOLA¹

¹ Institute of Developing Economies-Japan External Trade Organization.

3-2-2 Wakaba, Mihama-ku, Chiba-shi, Chiba, 261-8545, Japan

² Faculty of Global and Regional Studies, Toyo University

5 Chome-28-20 Hakusan, Bunkyo City, Tokyo 112-8606,

corresponding author: tsubota@toyo.jp

³ Economic Research Institute for ASEAN and East Asia

Senayan Senayan II 6F, Jalan Asia Afrika No.8 Gelora Bung Karno, Senayan Jakarta Pusat 10270, Indonesia

Keywords: Dedicated Freight Corridors, railways, India, reduction in CO2 emissions, Modal shift, Synergy effects,

1. Introduction

Dedicated Freight Corridors (DFCs) in India have considerable potentials to transform existing logistics and the location of economic activities. DFCs are composed

of two lines of freight trains connecting Delhi and other inland cities to Mumbai, Kolkata, international ports, and beyond to foreign markets. This transformation can most probably overtake the existing logistics networks in India, which is mainly dependent on road transportation for freight and passengers. This current unpopularity of railways is its slow speed and uncertain time schedules due to old technologies.¹ The instalment of DFCs can make the modernization of the technology and the separation of freight trains from passenger trains. These can allow DFCs to be punctual, have high average speed, and have larger capacities. Furthermore, DFCs can fulfil another important and expected outcome: reduction of Greenhouse Gas Emissions (GHG), which is an acute need to prevent global warming. India is ranked third for GHG emissions and sixth for per capita emissions in 2020.² Among the sectors, the transportation sector is regarded as one of the largest emitters. In

* Corresponding author; tsubota@toyo.jp

¹ See, for example, Bajaj (2010) and Lasseter and Kotoky (2015).

² See for Figure 2.2. of UNEP (2020) Emissions Gap Report 2020.

India, the transportation sector is ranked fourth, following Energy, Agriculture, Manufacturing & construction.³

This paper evaluates the economic and environmental impacts of Dedicated Freight Corridors in India. By employing Spatial Computable General Equilibrium Model called Geographical Simulation Model (IDE-GSM), we estimate the spatial economic impacts by industries. As our model is incorporated with modal choice among four transport modes for different manufacturing sectors, we can also explore the possible modal shifts from road to rail, which resulted in the low carbon emissions.

DFCs directly contribute to the improvements of transportation, which can change optimal routes, transportation time, and modes of transportation among regions. Given these changes, DFCs can even change the optimal location of industries in the future. On economic aspects, with better accessibility within and beyond India, we show that there are large economic benefits around the regions DFCs project has implemented as well as the hinterland regions, to which may be indirectly connected. These changes can accelerate economic growth of India.

On environmental aspects, DFCs can make modal shifts happened from tracks to railways, which can largely reduce CO₂ emissions. Of course, the induced economic impacts can make the freight traffic larger than before, which would offset the reduced CO₂ emissions. From our SCGE analysis, as the reduction of CO₂ emissions by the modal shifts is large, we show that DFCs can reduce the net amount of CO₂ emissions.

By showing the above impacts, our study shows so called Wider Economic Benefits (WEB), which include direct benefits as well as induced benefits and environmental impacts.⁴ The rest of our paper is organized as follows. Section 2 gives a brief overview of the DFCs. In section 3, we show our models, data, assumptions, the mechanisms and the scenarios. Then, showing the baseline scenario, and each scenario, we present the results on economic impacts in section 4 and on the environmental impacts in section 5. Some discussions shall follow in Section 6.

³ The data is as of 2016, obtained from Climate Data Explorer.

⁴ See ADB, UKAID, JICA, and WBG (2018) for WEB in South Asia, Wangsness et al (2017) on cross-country comparison of transport appraisal, and Roberts et al (2020) for meta-analysis of papers and reports on transport corridors and WEB.

2. Overview: India needs better intranational connectivity.

As the Indian Railways are one of the oldest and largest railways in Asia, the operations are poorly maintained, the average speed is as low as 23km/h, and hours of delays are frequent.⁵ Thus, road transportation by typical Indian trucks is the most popular.⁶ With the constant economic growth, transportation demands are also on the rise.

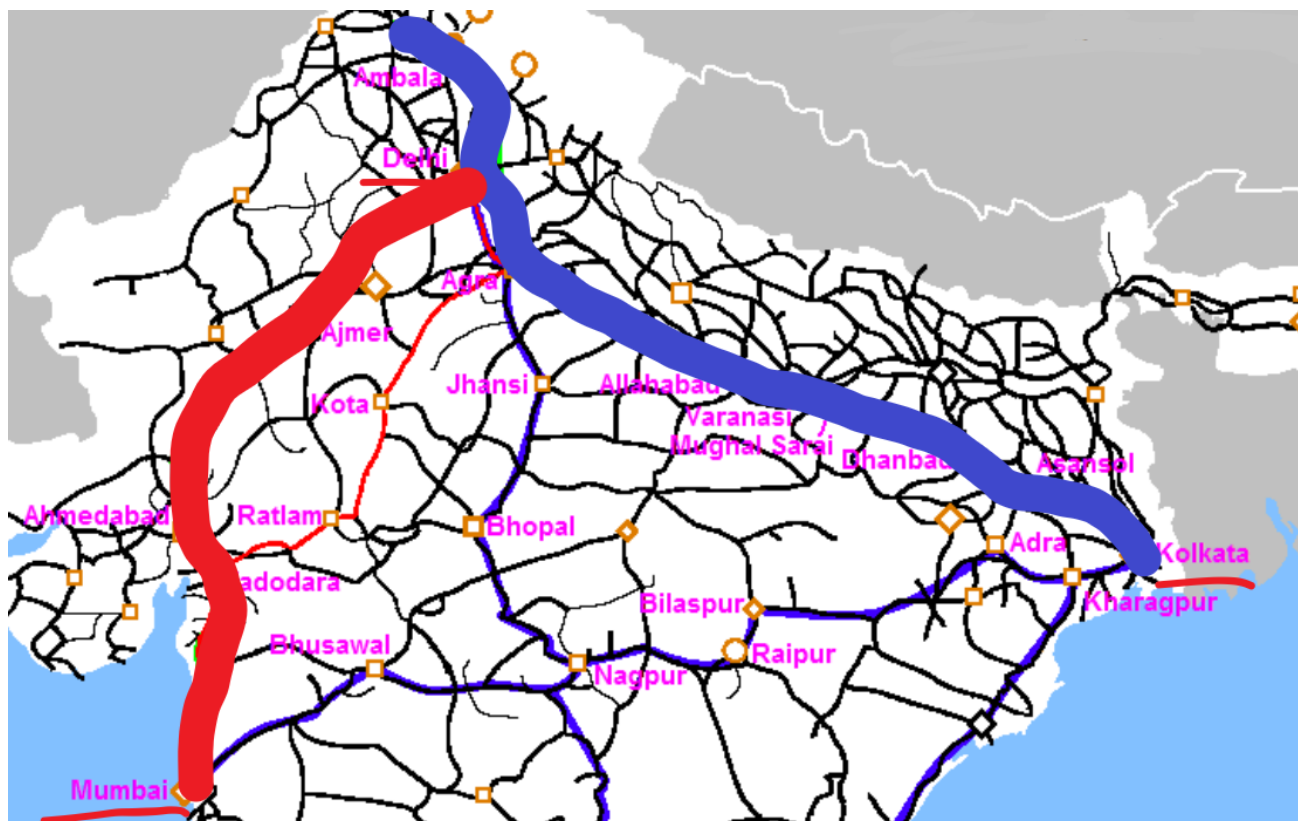


Figure 1: Project map of Dedicated Freight Corridors

Source: Authors' modification based on the map in Wikimedia

https://upload.wikimedia.org/wikipedia/commons/7/76/Indian_Railways_DFC.png

3. The model

3.1. The model

⁵ See, for example, Bajaj (2010) and Lasseter and Kotoky (2015).

⁶ According to Table 1.1. of Government of India (2014), shares of freight transportations were rail: truck: others = 30:61:9 in ton and 36:50:24 in net ton-km as of 2007/08.

We built an economic geography model of the Spatial General Equilibrium Model called Geographical Simulation Model, whose features are summarized by **1) agglomeration economies, 2) detailed intra-national regions, 3) multi-modal choice, and 4) multi-sectors.**

First, agglomeration, urbanization and industrial clustering are the phenomena widely observed in Asia. Our model has incorporated these agglomeration economies at the core of our model. Specifically, our model is built on Puga (1999), employing multisector and national general equilibrium of New Economic Geography using costly agricultural goods. The reason for taking this modelling strategy is to fit our analysis to the context of developing countries where most labour force may be in agriculture.⁷ We modified the specification of dispersion forces in Puga (1999) by setting land as inputs for agriculture and assuming that the rents from agricultural lands are distributed to households in the same region. The agricultural sector uses labour and land as its inputs under constant returns to scale technology. With this setting, we can reflect different endowments in agricultural productions. We assume that agricultural land rents accrue to households in the same region.⁸

3.2. Data

For our analysis, the most crucial variables are 1) population, 2) gross regional domestic product (GRDP), 3) industrial composition, and 4) the area size of arable land. We have compiled these variables into our geospatial data from various sources. The dataset covers 169 countries/economies and 3,262 sub-national regions, where most of the developing countries are at sub-national level as they are our focus.

3.3. The mechanisms

⁷ Typical models of spatial economics assume agricultural goods to be homogeneous and agricultural transportation to be costless. These simple specifications have benefits for theoretical modelling. However, within these assumptions, we cannot analyze the benefits of transport improvements to agricultural sector. Accordingly, we maintain the assumption of homogeneity in agricultural goods but dropped costless transportation. This is the modeling strategy of Puga (1999) in costly agricultural goods.

⁸ There are some different modeling strategies such as exogenous amenities, housing prices and others. However, it is one of the reasons why we don't take these strategies because it is impossible to collect such data for the world. Instead, we collect the agricultural arable land size for each country and region some from national agricultural census and others from satellite data of land use. Note that the input requirement of land is assumed to be only for agriculture.

The alternative scenarios contain specific policy measures at a specific year. With the alternative scenarios, we execute simulation and compare the outcome variables such as GRDP and GRDP per capita against the ones from the baseline scenario. The differences between the two scenarios show the impacts of the specific policy measures listed in the alternative scenario. When the GRDP of a region in a scenario is higher or lower than that under the baseline scenario, we regard the surplus or deficit as the economic impacts stemming out of the difference in policy measures (Figure 3). As a caveat, for example, when a region had a negative economic impact from a project at a certain point of time, it does not mean that the region will be worse off compared with the current situation (c.f 2020) but rather, it indicates a relatively slower growth rate compared with the baseline scenario. Negative values in the results simply imply relatively slower growth compared with the baseline scenario, which doesn't necessarily mean there is negative economic growth.

3.4. Scenarios

We conducted three scenarios

- 1) East Dedicated Freight Corridor (Kolkata to Ludhiana via Delhi)
- 2) West Dedicated Freight Corridor (Mumbai to Delhi via Ahmedabad)
- 3) Both Dedicated Freight Corridors

We set the average Speed of DFC as 85 km/h. The year of completion is set as the end of 2022.⁹ Also, we set the transshipment time at the station as 3 hours, which is half of the baseline.

4. Results

4.1. Economic impacts

The results of both DFCs show that spatially wider regions can positively benefited from the projects. The appearance of the spatial distribution of the impacts are very similar to both. More specifically, the map looks reflecting the combination of two previous maps. In Table 1, we can find that largest contribution comes from Textile and services, followed by

⁹ "Indian Railways DFC project needs daily monitoring to ensure completion by June 2022, say Piyush Goyal" (the then minister of railways) from *Financial Express*, January 12, 2021. With possible further delays, we set the timing of completion as the end of 2022.

agriculture, E&E and food processing. With Eastern DFC, the increase in real GDP per capita shall be 0.57%. Similarly, with Western DFC, the increase shall be 0.32%. Then the scenario with both DFCs shows that the increase shall be 0.96%. Interestingly, the impacts with both DFCs together are larger than the sum of each DFC by 0.06% point, which we call the synergy effects of two different infrastructure projects.

Scenario	Agriculture	Automotive	E&E	Textile	Food Processing	Other Manu.	Services	Mining	Real GDP per capita
Eastern DFC	0.27%	0.18%	0.28%	0.52%	0.22%	-0.06%	0.76%	-0.02%	0.57%
Western DFC	0.21%	0.26%	0.19%	0.65%	0.25%	0.11%	0.38%	-0.02%	0.32%
Both DFCs	0.53%	0.46%	0.50%	1.22%	0.49%	0.09%	1.21%	-0.04%	0.96%
East + West	0.48%	0.43%	0.48%	1.17%	0.47%	0.05%	1.13%	-0.03%	0.89%
Both/(E+W)	110%	105%	106%	104%	105%	189%	107%	-108%	108%

Note) The numbers show the percentage change of real value added or real GDP compared to the results from the baseline scenario. The simulation results were compared in 2042.

Source: Authors' calculation with IDE-GSM

Table 1: Comparison of the economic impacts against the baseline scenario

4.2. Environmental impacts

For further analysis on environmental impacts, we take three steps: freight flow estimation, CO2 reductions from modal shifts, and net effects in CO2 emissions. First two steps are on transportation sector and the third is the rest of the economies with the results of the first two. In environmental sense, this analysis is a kind of two-sector model of transportation and rest of the economies. After obtaining the prospects of CO2 emissions by each, we combine the result to have the net effects.

Scenario	Reduction of CO2 emissions in transport sector	Increase in CO2 emissions from induced economic activities ^{a)}	CO2 emissions in India ^{b, c)}		Total Reduction of CO2 emissions in India
			Share of transport sector	Share of all other sectors	
Baseline	0%	0%	11.82%	88.18%	0%
Eastern DFC	-9.72%	0.45%	10.75%	89.25%	-0.75%
Western DFC	-13.80%	0.26%	10.33%	89.67%	-1.40%
Both DFCs	-24.21%	0.76%	9.16%	90.84%	-2.19%
East + West	-23.52%	0.71%	9.24%	90.76%	-2.15%
Both/(E+W)	102.95%	107.60%	-	-	101.59%

Note)

a) The estimates are 80% of numbers in Table 1.

b) Estimated sectoral CO2 emissions in India for the baseline are obtained from climatewatchdata.org. We take the average of 2008-2018.

c) On the contribution of CO2 emissions by sector, we divide the economy into two parts: transport sector and all other sectors.

Source: Authors' calculation with IDE-GSM

Table 2: Comparison of the economic impacts against the baseline scenario

4.2.1. Net effects in CO2 emissions

On the economy wide CO2 emissions, by benchmarking the current shares in CO2 emissions, we estimate the scenarios. The current shares of CO2 emissions are taken by the average of 2008-2018.¹⁰ The share of transport sector is 11.82% and the rest is 88.18% as in first row of third column in Table 2.

For these contributions of sectoral CO2 emissions, we subtract the impacts of modal shift from the share of transport sector and add the impacts of induced economic activities to the share of all other sectors. Then, we obtain the revised composition of CO2 emissions. The comparison to the baseline case in total CO2 emissions are shown in the last column in Table 3.

From the results, we can see that the largest reduction in total CO2 emissions can be achieved by the scenario for both DFCs. Though the gap is as small as 0.04%, we can still find the synergy effects in CO2 emissions. In terms of reduction in CO2 emissions, Western DFC is larger than Eastern DFC, because there are more manufacturing industries in Western part of India.

5. Conclusion

In this paper, we have considered two DFCs from Delhi to Mumbai and Delhi to Kolkata and explored the magnitudes of Wider Economic Benefits (WEB) in terms of spatial economic impacts and environmental impacts. We have confirmed that DFCs in India have huge potentials to transform the economy geography in India as well as reduce CO2 emissions.

On spatial economic impacts, we find the complementary relation of the two DFCs by revealing the synergy effects, where both DFCs together can create larger economic impacts than the sum of each DFC. Both DFCs shall increase 0.96% of real GDP per capita of which the synergy effects are 0.07%. From the spatial perspectives, the regions near to DFCs have strongly positive impacts, while the Southern India shall have negative impacts (slower growth) as they are far from DFCs.¹¹ This is because Southern regions are yet to be directly connected. Probably, the completion of additionally planned DFCs in South and Southeast

¹⁰ The data source is climatewatchdata.org. The use of this numbers has an assumption that baseline scenario maintains the sectoral structure of CO2 emissions.

¹¹ Note that negative impacts do not mean the negative growth. As this is the comparison to the baseline, our literal interpretation is slower growth than the baseline result.

can flip these spatial results. However, note that while there were the largest clusters of industries are in Western DFC and the highly populated regions are along with Eastern DFC, other parts are relatively weaker in these economic characteristics. Thus, the possible impacts may be smaller than the first two DFCs.

On the environmental impacts, the impacts of DFCs have two opposite directions: modal shift to railways (reducing CO₂) and induced economic activities in production and transportation by cheaper transportation costs and faster transportation. We find that the reduction of CO₂ emission from modal shift is larger. Furthermore, we also find the synergy effects in CO₂ emissions, where reduction of CO₂ emissions is larger when both DFCs are competed together.

References

- Asian Development Bank, UKAID, JICA, and World Bank Group (2018) *The WEB of Transport Corridors in South Asia*, Washington, DC: World Bank.
- Kumagai, Satoru, Kazunobu Hayakawa, Ikumo Isono, Souknilanh Keola and Kenmei Tsubota (2013) "Geographical simulation analysis for logistics enhancement in Asia," *Economic Modelling*, 34: 145-153.
- Lall, Somik V. and Mathilde Lebrand (2020) Who wins, who loses? Understanding the spatially differentiated effects of the belt and road initiative, *Journal of Development Economics*, 146, 102496.
- Puga, Diego (1999) The rise and fall of regional inequality, *European Economic Review*, 43(2), 303-334.
- Puga, Diego and Anthony J. Venables (1996) The Spread of Industry: Spatial Agglomeration in Economic Development, *Journal of the Japanese and International Economies*, 10(4), pp440-464.
- Redding, Stephen J. and Esteban Rossi-Hansberg (2017) Quantitative Spatial Economics, *The Annual Review of Economics*, 9, 21-58.

Introducing "Contextual Metrics" as a mathematical definition for a comprehensive approach of geographical distances

Alain L'Hostis¹

¹LVMT, Univ Gustave Eiffel, IFSTTAR, Ecole des Ponts, F-77454 Marne-la-Vallée,
France

September 10, 2021

+33 (0)3 20 43 84 98, alain.lhostis@univ-eiffel.fr

Keywords: Geographical distance, Contextual metrics, Travel, Motility, Break

1 Abstract

Our goal is to propose and discuss a mathematical framework for the description of geographical distance in a comprehensive way. In our view, geographical distance always refer to potential or realized movement between places, and these displacements obey the least effort rule. While this optimization of effort is well known to imply the Triangle inequality in many situations, breaks in movement generate a paradox: effort optimization, taking into account the need to rest, results in apparent violations of the Triangle Inequality. In order to solve this issue, we introduce contextual metrics that consider space but also any contextual information relevant to travel, such as resources used for moving. Our approach permits to build a *subjective space* where distances are affected by the characteristics of the individual on the move. Contextual metrics frame the optimization problem in a space enriched by the context that the traveller has to take into account, making apparent that the violation of the Triangle Inequality in case of break was only an artefact of a model lacking crucial information. The range of geographical situations that can be modelled with this framework underline the level of generalization that can be expected from this approach.

The purpose of this paper is to expose the mathematical definition of contextual metrics following Kloeckner, L'Hostis, and Richard (2020), and explicit further the possible applications of the model in geography. The paper will heavily borrow from this published article by the same authors.

2 Introduction

We define geographical distances as the difficulty in relating two locations in geographical space. This definition implicitly conveys an idea of movement, actual or potential, between the two locations. Whether any potential movement will actually take place or not is guided by the expected ratio of benefit over cost; and the way in which the movement will occur (choice of path, of transport mean) will be guided by attempting minimization of the cost (Zipf 1949; Warntz 1967; Wu

and Miller 2001; Geertman and Eck 1995). Therefore, in order to understand movements in a geographical space it is necessary to consider more than the physical lengths, and by *distance* we will mean a disutility associated to movement. Moreover, in order to gain deeper understanding of movement, one has to take into account every element that influences the incurred disutility, as they may affect whether and how the movement will take place.

3 What is the problem?

We start from the mathematical definition of metrics¹ that corresponds to the notion of geographical distance. The mathematical metric relates two elements of a given space.

Definition 3.1 (Metrics). Let X be a set, which shall represent the geographical space; elements of X are called *locations*.

A *metric* on X is a function

$$d : (X)^2 \rightarrow [0, +\infty]$$

satisfying the following axioms:

- i.* (Identity) for all $p \in X$, $d(p, p) = 0$,
- ii.* (Separation In Space) for all $p, q \in X$, if $d(p, q) = 0$ then $p = q$,
- iii.* (Triangle Inequality) for all $p, q, r \in X$,

$$d(p, q) \leq d(p, r) + d(r, q),$$

The problem stems from how the notion of *break* has been considered incorrectly as an impediment to the universal validity of the triangular inequality (L'Hostis 2016; L'Hostis 2017; L'Hostis 2020). The *breaks* in itineraries, that are necessary to relaunch movement, do not entail the idea of sub-optimality in distances that the triangle inequality violation suggest. *Breaks* – quite counter-intuitively – contribute to form optimal paths, and hence should not generate violations of the triangle inequality. Losing triangle inequality means losing distance as an optimal measurement and entails difficult issues for the representation of space.

Our initial motivation is theoretical and concerns the Triangle Inequality. It has often been considered irrelevant to certain geographical distances, but we show that the triangular inequality holds in one form or another universally, in a very wide framework, since it follows logically from the minimization process. We close a critic of the critics of the Triangular Inequality opened in L'Hostis (2016), by showing that the last remaining case where it seemed not to hold is in fact a case where distances are ill-defined, precisely because they depend on a *context*.

4 Illustrating the problem: the motel case

Assume, as on figure 1, two cities A, B are connected by a single road with a motel M in between; to travel by car between A and M , or between M and B takes 8 hours. Assume further that one is not able or allowed to drive for more than 8 hours in a row, and that after 8 hours of driving an 8 hours rest is needed.

¹We shall use the word *metric* for the functions considered as a whole, and the word *distance* for individual measurements.

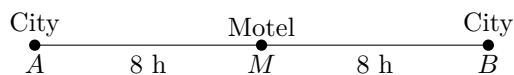


Figure 1: Two cities, a midway motel and the length of two paths AM and MB

It has been claimed that this is a counter-example to the triangular inequality (L'Hostis 2016): the distance (in travel time) from A to B would be 24 hours (twice 8 hours of driving and 8 hours of rest in between), larger than the sum of the distance from A to M and the distance from M to B (both 8 hours). The problem here actually does not lie in the triangular inequality, but in the fact that distances (travel times) are not well-defined out of *context*: for example the distance from A to M can be either 16h or 8h, depending on the necessity to consider the break time or not. Both measurement can make sense in a geographical context. For instance the time duration of long car trips provided by online routing services² usually do not include rest time, which may be avoided if two drivers take turns along the journey. Going back to our example, the travel time between any two points in fact depends on whether the traveler is rested or tired: there is a *context* to be taken into account.

5 What do we propose?

We propose to model the motel situation by a set of 6 couples, or *states*, where the context is written for cleaner notation in exponent of the location:

$$\{A^r, A^t, M^r, M^t, B^r, B^t\} \simeq \{A, M, B\} \times \{r, t\},$$

encoding the six possible combinations of a location and a context (r for rested, t for tired) of the traveler, as in figure 2. Each of the following actions take 8 hours: rest while staying in place, travel from a location to a neighboring one; traveling is only possible when rested and causes tiredness. Now the minimal length of a path between two states does satisfy the triangle inequality.

By enriching the space we see the initial model, where the triangle inequality seemed not to hold, as the mere shadow of a more accurate representation of the situation at hand. The apparent violation of the triangle inequality is revealed as an undeterminacy caused by important information being left out. Modeling much more general situations where non-spatial information, i.e. a context, influence the length of travel is the goal of the framework we are about to develop.

Definition 5.1 (Contextual metrics). Let X be a set, which shall represent the geographical space; elements of X are called *locations*. Let C be a set called the set of *contexts*. A pair in $X \times C$ corresponding to location p and context c will be called a *state* and denoted by p^c ;³ it compounds the information on the location p and the context c .

A *contextual metric* on X (with context set C) is a function

$$d : (X \times C)^2 \rightarrow [0, +\infty]$$

satisfying the following axioms:

²For instance: <https://maps.openrouteservice.org/>

³The usual mathematical notation would be (p,c) , but a lighter notation seemed better and confusion with taking power is unlikely.

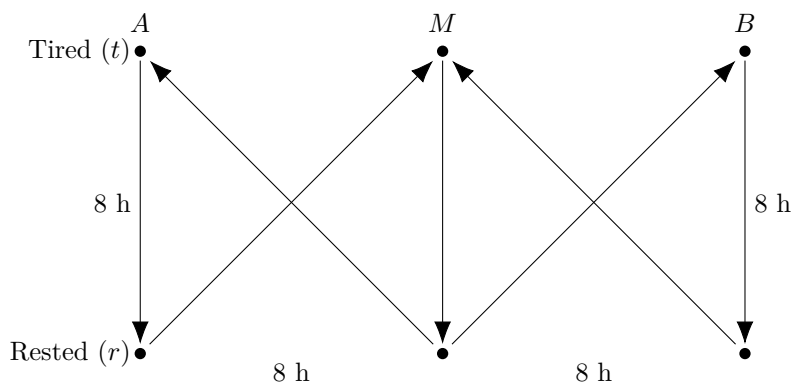


Figure 2: Two cities, a midway motel, and two contexts. Each arrow takes 8 hours.

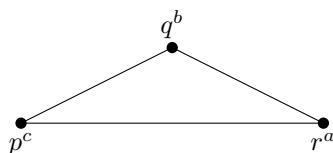


Figure 3: A triangle in the state space: each vertex is a state, i.e. a pair (location^{context}), and each edge is a path.

- i.* (Identity) for all $p^c \in X \times C$, $d(p^c, p^c) = 0$,
- ii.* (Separation In Space) for all $p, q \in X$, if $\inf_{c, b \in C} d(p^c, q^b) = 0$ then $p = q$,
- iii.* (Triangle Inequality) for all $p^c, q^b, r^a \in X \times C$,

$$d(p^c, q^b) \leq d(p^c, r^a) + d(r^a, q^b),$$

The number $d(p^c, q^b)$ shall be called the *contextual distance* between the two states p^c and q^b , so as to distinguish with the function d as a whole. While a non-contextual distance is measured between two points in space, a contextual distance is measured between two states, i.e. between two pairs of a location and a context. A crucial point of differentiation between the contextual and the non-contextual distance resides in the fact the addition of two measurements must take into account the context: the Triangle Inequality is only assumed when the ending context of the first term is equal to the starting context of the second term, as in figure 3.

The geographical interpretation is that the contextual distance $d(p^c, q^b)$ is meant to express the least possible disutility (e.g. least possible length traveled, time spent, monetary cost, etc.) needed to be endured when starting at location p with context c , to attain location q with context b . We will often speak of “length” instead of disutility, as we want to think about the enriched space geometrically, but it should be remembered that it can be measured in various units.

6 What is new for Geography?

Despite its central role as a concept, and compared to other concepts like space, distance is relatively under-studied in geography. Geographical distance is generally unequivocally associated to empirical measures in kilometres, time or cost, or to a mathematical formula of a metric in a spatial model that will care for spatial remoteness.

Our proposal, contextual distance, includes inside the very concept of geographical distance a set of elements, that were previously not associated to it. A context will be information to be added to the location and that will influence the available choices of movement from this location and their costs. Instead of considering movement solely into physical space, we will recast it into an *enriched space* whose elements are called *states* and compound location and context; in particular, this enables us to express distance to a location not only in terms of the starting location, but also in terms of the context in which the traveler finds herself or himself.

Contextual distances can cover:

- Distances of resource dependant movements:
 - A limited available budget could force the traveller to avoid toll express ways, or to opt for a cheaper transport mode, altering distances measurement
 - Considering fuel gauge level, battery level, or food and water supply will imply the need for a break to reload movement, altering distances measurement
 - The *motility* approach to understanding mobility Flamm and Kaufmann 2006 stresses the role of individual skills and cognitive factors; these factors influence many travel choices, and could be considered in a *contextual distance* approach
- Distances of time dependant movements:
 - A fixed starting time, and
 - a desired arrival time, will have an impact on distance through scheduled transport means, or congestion

The two facets of the enriched space, geometry and context, do not play symmetric roles. The geometry is immutable, at least at the timescale involved in the movement, while the context will usually evolve as the movement proceeds. The geometry only takes into account places (starting location and target location), while context denotes all other movement-related circumstances.

Our framework being flexible, contexts can also be used to model differences between travelers (e.g. differences in wealth or in preferences, i.e. differences that are intrinsic to the travelers and not related to their locations, and thus are ignored in traditional models of geographic distances). Beyond a purely physical approach of space, what we model is closer to space *as it is experienced by the traveler*: different travelers, with different constraints and incentives, will make different choices and this means that they will each have their own “set of states”, with possibly only the spatial component in common.

In our framework, distance, as the solution of an optimization problem, is not universal: it actually depends on the constraints – linked to the context – under which minimization occurs. In this sense we follow the views – inspired by Poincaré’s reflections (1908, p. 116) – of modern human geographers (Blaut 1961; Gatrell 1983; Bailly 1985; Couclelis 1999). This approach moves away from an absolute – Newtonian and Kantian – conception of space and supports a more relative and subjective conception.

7 Acknowledgements

The author thank Benoît R. Kloeckner and Thomas Richard.

References

- [1] Antoine Bailly. “Distances et Espaces: Vingt Ans de Géographie Des Représentations”. In: *L'espace géographique* 3 (1985), pp. 197–205. DOI: [10/ggw564](https://doi.org/10/ggw564).
- [2] J.M. Blaut. “Space and Process”. In: *Professional geographer* 13.4 (1961), pp. 1–7. DOI: [10/fcxkt7](https://doi.org/10/fcxkt7).
- [3] Helen Couclelis. “Space, Time, Geography”. In: *Geographical Information Systems: Principles, Techniques, Management, and Applications*. Wiley. Vol. 1. 1999, pp. 29–38. URL: [http://bookfel.org/download/english\(2\)/g%C3%A9ographie/Space,%20time,%20geography.pdf](http://bookfel.org/download/english(2)/g%C3%A9ographie/Space,%20time,%20geography.pdf) (visited on 08/28/2015).
- [4] Michael Flamm and Vincent Kaufmann. “Operationalising the Concept of Motility: A Qualitative Study”. In: *Mobilities* 1.2 (June 2006), pp. 167–189. ISSN: 1745-0101. DOI: [10/c5x2mj](https://doi.org/10/c5x2mj).
- [5] Anthony C. Gatrell. *Distance and Space: A Geographical Perspective*. Clarendon Press Oxford, 1983. URL: <http://www.getcited.org/pub/102287206> (visited on 05/27/2013).
- [6] Stan C. M. Geertman and Jan R. Ritsema Van Eck. “GIS and Models of Accessibility Potential: An Application in Planning”. In: *International Journal of Geographical Information Systems* 9.1 (1995), pp. 67–80. ISSN: 0269-3798. DOI: [10/dxkxn4](https://doi.org/10/dxkxn4).
- [7] Benoît R. Kloeckner, Alain L'Hostis, and Thomas Richard. “Contextual Metrics a Mathematical Definition for a Comprehensive Approach of Geographical Distances”. en. In: *Geographical Analysis* n/a.n/a (2020). ISSN: 1538-4632. DOI: [10/gjqhf5](https://doi.org/10/gjqhf5).
- [8] Alain L'Hostis. “All Geographical Distances Are Optimal”. en. In: *Cybergeo : European Journal of Geography* (June 2020). ISSN: 1278-3366. DOI: [10/gg3rbx](https://doi.org/10/gg3rbx).
- [9] Alain L'Hostis. “Detour and Break Optimising Distance, a New Perspective on Transport and Urbanism”. In: *Environment and Planning B: Planning and Design* 44.3 (2017), pp. 441–463. DOI: [10/gddwqw](https://doi.org/10/gddwqw).
- [10] Alain L'Hostis. “Misunderstanding Geographical Distances: Two Errors and an Issue in the Interpretation of Violations of the Triangle Inequality”. In: *Cybergeo : European Journal of Geography* 793 (2016). DOI: [10/ggw58h](https://doi.org/10/ggw58h).
- [11] Henri Poincaré. *Science et méthode*. FR. Flammarion, 1908. URL: <https://gallica.bnf.fr/ark:/12148/bpt6k9691658b> (visited on 01/16/2020).
- [12] William Warntz. “Global Science and the Tyranny of Space”. en. In: *Papers of the Regional Science Association* 19.1 (1967), pp. 6–19. ISSN: 1435-5957. DOI: [10/dz96kp](https://doi.org/10/dz96kp).
- [13] Yi-Hwa Wu and Harvey J. Miller. “Computational Tools for Measuring Space-Time Accessibility within Dynamic Flow Transportation Networks”. In: *Journal of Transportation and Statistics* 4.2/3 (2001), pp. 1–14.
- [14] George Kingsley Zipf. *Human Behaviour and the Principle of Least-Effort*. Cambridge, USA: Addison-Wesley, 1949. URL: <http://www.citeulike.org/group/2050/article/1283136> (visited on 06/03/2013).

An extension of the Tobler's law?

Influence of the metric to compute the distance between an optimal facility and its demand points

Didier JOSSELIN¹

¹UMR ESPACE CNRS, Avignon Université

84029 AVIGNON Cedex, France

+33607406938, Didier.josselin@univ-avignon.fr

Submitted to Special Session SS05 "Geographical Distances"

Keywords: Tobler's First Law, distance, 1-facility location problem, L_p -norms, spatial compromise

This research work presents an extension of the Tobler's first law (TFL), that includes the metric used to measure the distance between a center and its demand points, in an optimal facility location problem framework (Hakimi, 1964). It provides a short overview of the works derived from the TFL and opens on the definition of a second law, less geographical and more mathematical. Indeed, its seems the distance, that is to say in our case a L_p -norm (a generalized Minkowski distance) considered to estimate a proximity between two points, has a strong weight on the meaning and on the shape of the TFL.

We organise our talk in several steps. We first recall the foundations and the geographical meaning of the Tobler's first law (1970, 1971) which is still considered as a solid and reliable base in many disciplines (for instance, see the introduction in Cressie & Moores, 2021). This led to many citations of this law, which is indeed a geographically verified assumption rather than a demonstrated law.

We first draw a short history of this law (Cauvin & Reymond, 1991, Miller, 2004, Sui, 2004), provide a few examples of its usages (Grasland, 2009, Hecht & Moxley, 2009) and its extensions, especially in the context of the Santa-Barbara's research group (Couclelis, 1996, Montello et al. 2003). We list and summarize several attempts to design a complementary or alternative second law in quantitative geography (Goodchild in Sui, 2004, Hecht & Moxley, 2009, Foresman & Luscombe, 2017, Josselin et al., 2017).

Then, in a second section, we move on transport and mobility concerns. We define the different purposes (efficacy, equity) of the principal centers used in geography (Beguin, 1989) and optimal location problem solving: k -median and k -center notably, to which we add the k -barycenter (modelling equality), that has interesting intermediate properties. We link these centers to their well-known objective functions.

In a third section, we focus on the 1-facility optimal location, a refined formal framework to explain how we generalize the optimal center through L_p -norms (Peeters & Thomas, 2000). Using sensitivity analysis (Drezner & Hamacher, 2004), we find out by computation a generalized fonction linking center location sensitivity, demand point influence and distance between the center and the demands (Ciligot-Travain & Josselin, 2013, Josselin et al., 2016).

We use this function to draw new balanced optimal locations, respectively with $p=1.5$ and $p=3$ in the L_p -norm. For those, we map the demand point spatial influence of the center location. We develop more deeply the L_3 -norm, which catches our attention: it shows a particular property of a linear relation between the demand points weights and their distance to the optimal center. As we did previously with well-known k -median and k -center metrics, we associate new optimal centers a semantic, respectively “effility” and “equaquity” to those two compromises.

In conclusion, we suggest how, in practical urban planning, our theoretical results can be potentially used to tune the effect of the distance on the demand point influence on the center. Finally, we draw up a comparison between our geographical law and the TFL: is our law an extension of the TFL or a second complementary law? We open a discussion about the proximity and distance collapse or reinforcement, as H. Couclelis put forward in 1996 and which is still a hot issue, from our perspective.

References

- Beguin H., 1989, Efficacité et équité en aménagement du territoire, *L'Espace Géographique*, Vol. 4, pp.335 – 336
- Cauvin C., Reymond H., 1991, Interaction spatiale et cartographie : les solutions de W. Tobler, *Espace Populations Société*, Vol. 3, pp. 467-485 .
- Couclelis H., 1996, The death of distance, *Environment and Planning B: Planning and Design*, Vol. 23, pp. 387-389, Editorial.
- Cressie N., Moores M. T., 2021, Spatial Statistics, arXiv:2105.07216v1[stat.ME].
- Drezner Z, Hamacher HW, 2004, *Facility locations. Application and theories*, Springer.
- Foresman T., Luscombe R., 2017, The second law of geography for a spatially enabled economy, *International Journal Of Digital Earth*, 10(10), pp. 979–995, <https://doi.org/10.1080/17538947.2016.1275830>, Taylor & Francis
- Grasland C., 2009, Spatial Analysis of Social Facts, Bavaud F. & Mager C. (Eds), *Handbook of Quantitative Geography*, University of Lausanne.

- Hakimi S.L., 1964, Optimum Locations of Switching Center and the absolute Center and Medians of a Graph, *Operations Research*, Vol. 12, pp. 450-459.
- Hecht B., Moxley E., 2009, Terabytes of Tobler: Evaluating the First Law in a Massive, Domain-Neutral Representation of World Knowledge, in Stewart Hornsby et al. (Eds.), COSIT 2009, LNCS 5756, pp. 88–105, 2009, Springer-Verlag Berlin.
- Josselin D., Ciligot-Travain M., 2013, Revisiting the optimal center location. A spatial thinking based on robustness, sensitivity, and influence analysis. *Environment and Planning B: Planning and Design*, SAGE Publications, 40 (5), pp.923-941. (10.1068/b38036). (hal-01091354).
- Josselin D., Rojas-Mora J., Blanke D., Gourion D., Ciligot-Travain M., et al., 2014, Influence of the metrics on discrete facility location. Toward a pertinent Lp norm targeting a transport objective, Proceedings of Transport Research Arena 2014, Apr 2014, Paris, France. ([hal-01100471](#)).
- Josselin D., Rojas-Mora J., Ciligot-Travain M., 2017, Spatial Function of Influence on Center Optimal Location based on Lp-norms. ICCSA 2017, GEOANMOD, Computational Science and Its Applications , Jul 2017, Trieste, Italy. (hal-01674560).
- Miller J., 2004, Tobler's First Law and Spatial Analysis, *Annals of the Association of American Geographers*, Forum, Vol. 94, No. 2, pp.284-289, Taylor & Francis.
- Montello D. R., Fabrikant S. I., Ruocco M., Middleton Harvey R. S., 2003, Testing the First Law of Cognitive Geography on Point-Display Spatializations, in Kuhn W., Worboys M.F., and S. Timpf (Eds.): COSIT 2003, LNCS 2825, pp. 316–331, Springer-Verlag Berlin.
- Peeters D, Thomas I., 2000, Distance predicting function and applied location-allocation models, *Geographical Systems*, Vol. 2, pp. 167 – 184
- Sui D. Z., 2004, Tobler's First Law of Geography: A Big Idea for a Small World?, *Annals of the Association of American Geographers*, 94(2), pp. 269–277, Blackwell Publishing, U.K.
- Tobler W. R., 1970, A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography*, Vol. 46, Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods (Jun., 1970), Clark University, pp. 234-240.

Parallel Session PSC4

4 November 11:20 am-1:20 pm (GMT)

Rural to Urban

Settling the challenges of a sustainable city through the dynamic representation of an urban agriculture system

F. SARKISSIAN^{1,2}; JP. ANTONI¹; T. LOYER²

¹ThéMA - UMR 6049

Université de Bourgogne

4 Boulevard Gabriel 21000 Dijon, France

+33 (0)3 80 39 55 41, fanny_sarkissian@etu-u.bourgogne.fr

²GéoHabitat

13 rue du Palais, 21000 Dijon, France

+33 (0)9 81 09 63 38, teddy@geohabitat.fr

Keywords: urban agriculture system, sustainable cities, systemic and dynamic approach, model

1. Introduction

Urban agriculture is seen as a lever for sustainable cities (Smit *et al.*, 1997; Deelstra and Girardet, 2000; Lovell, 2010). Theoretically, it partially responds to the challenges of social justice, sustainable economies, and viable environment (European Conference on Sustainable Cities and Towns, 1994). But a gap between theory and reality is observed: urban agriculture often does not go beyond the project stage and is often temporary. What does explain this difference between the social, economic, and environmental utility presented in the literature, and the real projects implemented? This question is the focus of this research.

Studying urban agriculture is a lever to understand its mechanism, apply it to the reality, and help its planification. Using a typology for observing urban agriculture has been proven to be incomplete. Indeed, this analytical approach has limits, while the systemic approach expands possibilities, “observing agriculture in its entirety, studying all the elements composing it at the same time” (Sarkissian, Loyer and Antoni, 2021). Different authors suggested this approach of agriculture through an ecosystem (Artmann and Sartison, 2018; Abu Hatab *et al.*, 2019). The mobilization of an interactive environment simulation software such as VENSIM has been chosen to study urban agriculture through a systemic approach. This mobilization comes through two axes: detailing the urban agriculture system’s mechanisms and then its behavior through dynamic modelling.

2. Detailing the urban agriculture system's mechanisms

According to the Food and Agriculture Organization (1999), "Urban and Peri-Urban Agriculture (AUP) refers to agricultural practices in and around cities that use resources - land, water, energy, labor - that can also be used for other uses to satisfy the needs of the urban population". In other words, urban agriculture production can be tangible and intangible. According to Duchemin, Wegmuller and Legault (2008), urban agriculture is multifunctional, and can produce eight distinguished functions.

Material production brings **food security** (Armar-Klemesu *et al.*, 2000; Orsini *et al.*, 2013). Defining the urban agriculture system gives the opportunity to study practices beyond the process of production. Three types of distribution can be distinguished: in situ, transported to places of processing, restaurants and host tables, or distributed in commercial services (Koç, Centre (Canada) and Security, 1999; ADEME, 2015; Zeeuw and Drechsel, 2015).

Urban agriculture also produces intangible functions and affects cities and population through **environment** (Viljoen, Bohn and Howe, 2005; Ackerman *et al.*, 2014; Artmann and Sartison, 2018), **social interactions** (Ackerman *et al.*, 2014; Dieleman, 2017), **education** (Azunre *et al.*, 2019), **urban planning** (Artmann and Sartison, 2018; Fosse, 2018), **economic development** (FAO, 1999; Ferreira *et al.*, 2018), **leisure activities** (Bouvier-Daclon and Sénécal, 2001; Lindemann-Matthies and Brieger, 2016), and **health** (Copenhagen: WHO Regional Office for Europe, 2016).

Another important element in the construction of the urban agriculture system is to define the actors interacting with and within urban agriculture productions, functions, and distributions. It can lead to a higher comprehension of services, influence and constraints leading to a fluctuating development.

3. Studying the urban agriculture system's behavior through dynamic modelling

The DYNAMO language (DYNAmic MOdels) is a concept developed by Forrester (1978) used to translate interactions taking place within a closed boundary around a system. Here, DYNAMO language is mobilized to represent the urban agriculture system. The interest is to introduce the formalization and quantification of all the elements interacting with or within this system. Once this demonstration given, the software VENSIM is used to represent this translate and to run different simulations, by studying its ability to maintain its balance and to regulate itself. Indeed, a system can show variety to adapt and survive. In this purpose, an example of representation can be done by running the simulation with data extracted from a theoretical city representative of European cities.

4. Conclusion and perspectives

This presentation aims to detail the elements interacting with and within the urban agriculture system and translate it into a dynamic system. Choosing not to rely on a field of study makes it possible to free oneself from the biases of the reality of a territory. The first results of this systemic and dynamic modeling are encouraging.

The limits of the model currently developed are its focus on the elements entering into account in the urban agriculture system. To fully answer the question posed by this research, the first necessity is to fully measure the durability of the various elements mentioned, by quantifying the functions of urban agriculture. Sustainability is difficult to measure, and it is not consensual (Hély and Antoni, 2019). Here, it will need a coupling of pre-existing indicators (Azunre *et al.*, 2019). The second step is the spatialization of this model. Currently theoretical, it is necessary to confront this model with the reality on the ground, and this confrontation must be done in part through spatialized data. It is planned to reflect on the link to be created between VENSIM and GIS software such as QGIS.

References

- Abu Hatab, A., Cavinato, M.E.R., Lindemer, A. and Lagerkvist, C.-J. (2019) 'Urban sprawl, food security and agricultural systems in developing countries: A systematic review of the literature', *Cities*, 94, pp. 129–142. doi:10.1016/j.cities.2019.06.001.
- Ackerman, K., Culligan, P., Plunz, R., Sutto, M.-P. and Whittinghill, L. (2014) 'Sustainable Food Systems for Future Cities: The Potential of Urban Agriculture*', *The Economic and social review*, 45, pp. 189–206.
- ADEME (2015) 'La transition énergétique en action', *ADEME & vous Le Mag*, p. 16.
- Armar-Klimesu, M., Bakker, N., Dubbeling, M., Gündel, S., Sabel-Koschella, U. and Zeeuw, H.D. (2000) 'Urban agriculture and food security, nutrition and health.', *undefined* [Preprint]. Available at: /paper/Urban-agriculture-and-food-security%2C-nutrition-and-Armar%E2%80%90Klimesu-Bakker/6fc7d0380108b6c3eb5e2ca1eda0f47fe4191fae (Accessed: 20 April 2021).
- Artmann, M. and Sartison, K. (2018) 'The Role of Urban Agriculture as a Nature-Based Solution: A Review for Developing a Systemic Assessment Framework', *Sustainability*, 10(6), p. 1937. doi:10.3390/su10061937.
- Azunre, G.A., Amponsah, O., Peprah, C., Takyi, S.A. and Braimah, I. (2019) 'A review of the role of urban agriculture in the sustainable city discourse', *Cities*, 93, pp. 104–119. doi:10.1016/j.cities.2019.04.006.
- Bouvier-Daclon, N. and Sénécal, G. (2001) 'Les jardins communautaires de Montréal : un espace social ambigu', *Loisir et Société*, 24(2), p. 507. doi:10.7202/000193ar.
- Copenhagen: WHO Regional Office for Europe (2016) *Urban green spaces and health review evidence*. Available at: http://www.euro.who.int/__data/assets/pdf_file/0005/321971/Urban-green-spaces-and-health-review-evidence.pdf?ua=1 (Accessed: 26 June 2019).
- Deelstra, T. and Girardet, H. (2000) 'Urban agriculture and sustainable cities', *Growing Cities, Growing Food: Urban Agriculture on the Policy Agenda*, pp. 43–65.

- Dieleman, H. (2017) 'Urban agriculture in Mexico City; balancing between ecological, economic, social and symbolic value', *Journal of Cleaner Production*, 163, pp. S156–S163. doi:10.1016/j.jclepro.2016.01.082.
- Duchemin, E., Wegmuller, F. and Legault, A.-M. (2008) 'Urban agriculture: multi-dimensional tools for social development in poor neighbourhoods', *Field Actions Science Reports. The journal of field actions* [Preprint], (Vol. 1). Available at: <http://journals.openedition.org/factsreports/113> (Accessed: 9 January 2021).
- European Conference on Sustainable Cities and Towns (1994) 'Aalborg charter', p. 8.
- FAO (1999) *Focus: Agriculture urbaine*. Available at: <http://www.fao.org/Ag/fr/magazine/9901sp2.htm> (Accessed: 7 December 2017).
- Ferreira, A.J.D., Guilherme, R.I.M.M., Ferreira, C.S.S. and Oliveira, M. de F.M.L. de (2018) 'Urban agriculture, a tool towards more resilient urban communities?', *Current Opinion in Environmental Science & Health*, 5, pp. 93–97. doi:10.1016/j.coesh.2018.06.004.
- Fosse, J. (2018) 'Les agricultures urbaines : potentiel de développement et impacts sur l'environnement et l'aménagement des territoires', p. 15.
- Hély, V. and Antoni, J.-P. (2019) 'Combining indicators for decision making in planning issues: A theoretical approach to perform sustainability assessment', *Sustainable Cities and Society*, 44, pp. 844–854. doi:10.1016/j.scs.2018.10.035.
- Jay Wright Forrester (1978) *Dynamique urbaine*. (Economie Publique de l'aménagement et des transports).
- Koç, M., Centre (Canada), I.D.R. and Security, R.P.U.C. for S. in F. (1999) *For Hunger-proof Cities: Sustainable Urban Food Systems*. IDRC.
- Lindemann-Matthies, P. and Brieger, H. (2016) 'Does urban gardening increase aesthetic quality of urban areas? A case study from Germany', *Urban Forestry & Urban Greening*, 17, pp. 33–41. doi:10.1016/j.ufug.2016.03.010.
- Lovell, S.T. (2010) 'Multifunctional Urban Agriculture for Sustainable Land Use Planning in the United States', *Sustainability*, 2(8), pp. 2499–2522. doi:10.3390/su2082499.
- Orsini, F., Kahane, R., Nono-Womdim, R. and Gianquinto, G. (2013) 'Urban agriculture in the developing world: a review', *Agronomy for Sustainable Development*, 33(4), pp. 695–720. doi:10.1007/s13593-013-0143-z.
- Sarkissian, F., Loyer, T. and Antoni, J.-P. (2021) 'Deploying Urban Agricultural System for an Innovative and Sustainable Urban Renewal', in: *7th International Conference on Geographical Information Systems Theory, Applications and Management*, pp. 222–228. Available at: <https://www.scitepress.org/PublicationsDetail.aspx?ID=MXCPHth76uU=&t=1> (Accessed: 25 May 2021).
- Smit, Jac, Ratta A., and Nasr J. (1997) 'Urban Agriculture, Food, Jobs and Sustainable Cities.', *Journal of Nutrition Education*, 29(6), pp. 361–362. doi:10.1016/S0022-3182(97)70254-1.
- Viljoen, A., Bohn, K. and Howe, J. (eds) (2005) *Continuous productive urban landscapes: designing urban agriculture for sustainable cities*. Amsterdam: Architectural Press [u.a.].
- Zeeuw, H. de and Drechsel, P. (2015) *Cities and Agriculture: Developing Resilient Urban Food Systems*. Routledge.

Types of Crimean Rural Settlement

Timur Y. GUSAKOV¹

¹Russian Academy of National Economy and Public Administration

Vernadskogo Prosp., 82, Moscow, 119571, Russia

+79998398366, gusakov.timur@mail.ru

Keywords: human geography, rural settlement, migration, rural area, Crimea

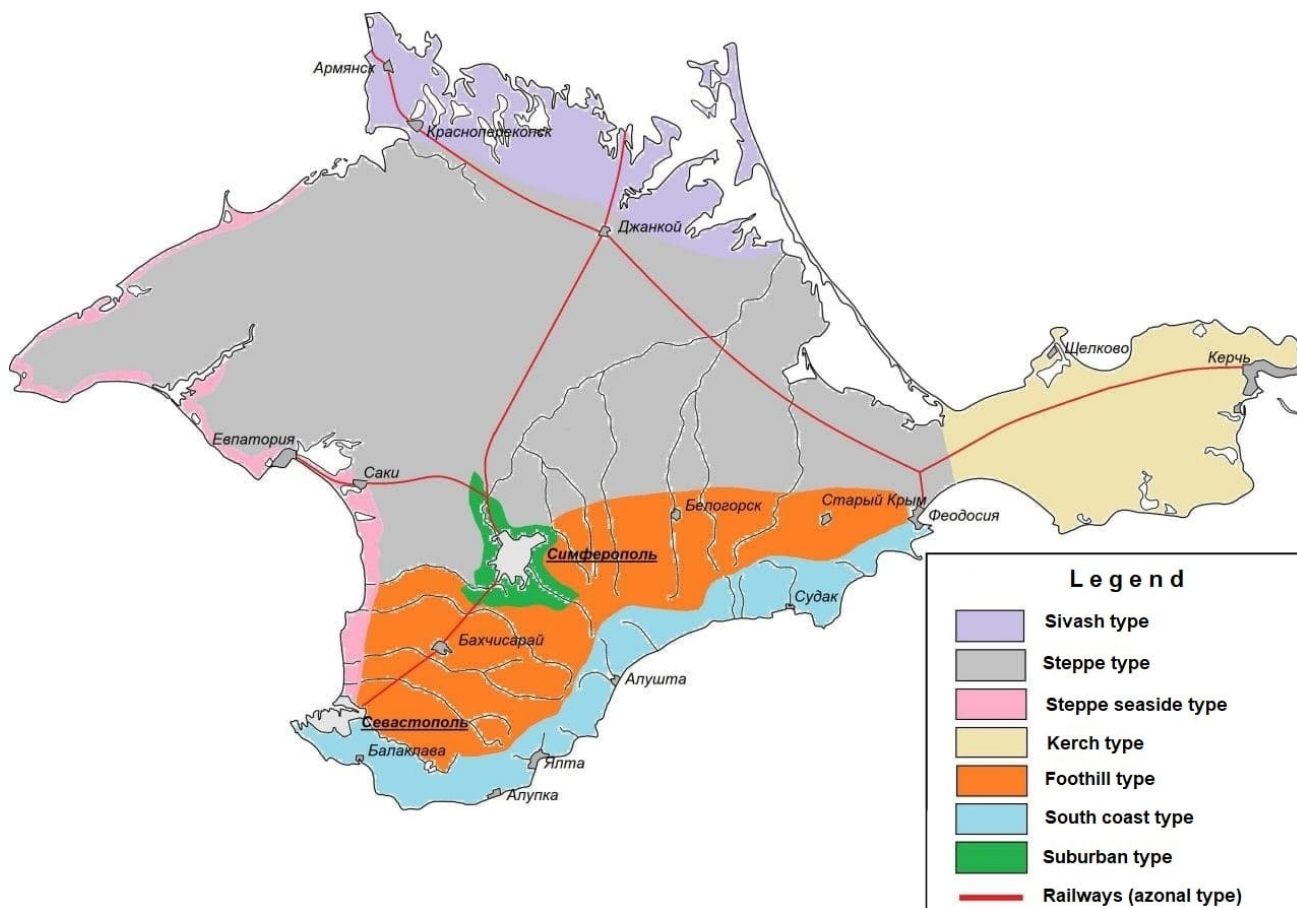
The modern rural settlement of the Crimean Peninsula has spatial differences due to geomorphological, natural, historical features. Considering the set of characteristic features proposed by the researchers of settlement, the following types of rural settlement of the Crimean Peninsula can be distinguished:

Sivash type confined to the gulf of the Sea of Azov Sivash, which is located on the northern shores of the Crimean Peninsula. Includes rural areas of the Armenian urban district, Krasnoperekopsky district and the northern part of Dzhankoy district. This type is characterized by a sparse mosaic rural settlement with small depopulating villages of up to 500 people. These territories were actively populated in the first half of the 20th century with the help of administrative tools – the allocation of state land and the direction of peasants there, later settlers for the establishment of settlements. Due to the spread of salt licks and salt marshes, as well as the aridity of the climate, the area is not very suitable for plant growing, more often for animal husbandry.

Steppe type is typical for most of the peninsula and includes the territories of Krasnogvardeisky, Nizhnegorsky, Pervomaisky, Razdolnensky, Saksy, Sovetsky, Chernomorsky districts, the northern part of Kirovsky and Simferopol districts. Rural settlement of the steppes suffered greatly because of the emigration of the Crimean Tatars during the 19th century (especially after the Crimean War). It received its modern drawing after the economic and resettlement transformations of the 1950s – 1970s. on the enlargement of collective farms, resettlement of small villages or their annexation to large settlements. The area is characterized by a high degree of plowing of soils and the development of animal husbandry more often based on personal subsidiary plots. Large villages with a population of more than 1,000 people are evenly located throughout the territory (former central estates of collective farms, now the centres of village councils), which serve the surrounding small RSs.

The *Steppe seaside subtype* is also distinguished, which includes settlements located on the Black Sea coast. These are villages of up to 1.000 people, like resort towns with a developed service sector, in the summer they perform exclusively recreational functions, in the winter and in the off-season, they are “sleeping” areas.

Kerch type includes a sparse rural settlement of the Kerch Peninsula (the entire territory of the Leninsky district) with small depopulating settlements of up to 500 people. The settlement network suffered greatly during the second exodus of the Crimean Tatars and only partially recovered. In Soviet times, fewer immigrants were sent to these areas due to the unpromising economic development of soils, only in places suitable for agriculture as pastures.



Map 1: Types of rural settlement of the Crimean Peninsula

Foothill type is common on the northern slope of the Crimean Mountains. It was formed at the end of the 19th century and has not undergone significant changes since then. Includes the northern territories of Sevastopol, the south of the Kirov and Soviet districts, Bakhchisarai, Belogorsk districts. Settlements are usually confined to natural objects – rivers, mountain valleys, and therefore have an elongated configuration, there are village-streets and multi-row villages. In economic terms, vegetable gardening, gardening, cattle breeding is widespread, less often – plant growing in rainfed fields.

South coast type was formed in the second half of the 19th – the first half of the 20th centuries. and includes the urbanized coastal territories of the southern coast of Crimea – the urban districts of Yalta, Alushta, Sudak, the southern territories of Feodosia and Sevastopol. It is represented by small urban-

type settlements, less often by villages that mostly perform recreational functions. The settlements have an elongated configuration and are located along the main highways along the Black Sea coast. You can also distinguish the *azonal type*. It was formed thanks to the construction of railways in the late 19th – early 20th centuries and includes a network of large settlements at railway stations. The settlements are characterized by block buildings and the adjoining of the centre of the settlement to the railway station. Initially, settlements developed on one side of the railway, but due to the annexation of neighbouring villages in Soviet times, in modern times they often consist of two parts, separated by a railway.

In connection with the development of agglomerations, a *suburban type* of rural settlement is formed, which includes rural settlements, which are expanding due to the active migration of the rural population to the suburbs from the steppe zone. A similar situation is observed around Simferopol. The adjoining rural settlements are turning into village-streets, stretching along the main directions of highways.

Acknowledgments

The article was written based on the RANEPА state assignment research programme.

References

- Kovalev S.A. Izbrannye trudy [Selected Works]. Smolensk, Oykumena, 2003, pp. 64–68.
- Lyalikov N.I. Tipy sel'skogo rasselenija SSSR [Types of the USSR's Rural Settlement] // Geography in school, 1949, No. 3, p. 4.
- Michurina F.Z. Sel'skoe rasselenie: metodologija i metodika regional'nogo analiza (na materialah Urala) [Rural Settlement: Methodology and Methods of Regional Analysis (Based on the Materials of the Urals)]. Perm, 1998, 199 p.
- O solonchakah Kryma [About the Crimea salt marshes] // Voprosy melioracii soloncov / ed. I.N. Antipov-Karataev. Moscow, USSR Academy of Sciences, 1958, pp. 176–192.
- Saushkin Y.G. Geograficheskoe izuchenie sel'skih naseleennyh punktov Sovetskogo Sojuza [Geographical study of rural settlements of the Soviet Union] // Geography issues, 1947, Vol. 5, p. 55.

IMPACT OF STATIC AND DYNAMIC DRIVERS ON URBAN DENSIFICATION

A. Chakraborty^{1*}, J. Teller¹, H. Omrani²

¹LEMA, University of Liege, 4000 Liege, Belgium

^{1*}+32 465769470, a.chakraborty@uliege.be

¹+32 43669499, jacques.teller@uliege.be

²Urban Development and Mobility, Luxembourg Institute of Socio-Economic Research, Luxembourg

+352 58 58 55 313, hichem.omrani@liser.lu

Keywords: Urban densification, Driving forces, Cadastral Data, Multi logistic Regression Model, Multi-density approach.

Abstract

With the increase in population, urban areas are becoming densely occupied, with building expanding vertically and horizontally. Thus, leading to construction of new housing units within the existing built up area resulting into urban densification. This paper examines and evaluates the static and dynamic drivers for urban densification using Multinomial Regression Model (MLR). The model is developed, calibrated and validated for the area of Brussels capital region, Flemish Brabant and Wallonia Brabant of Belgium. The 100 x 100 built up maps were generated from Belgian Cadastral (CAD) data for the years 2010 and 2016 and grouped into four classes namely, non-urban, low-density, medium-density and high-density. Modelling includes socio-economic, accessibility and topographical drivers to analyse the potential drivers for densification. The findings show that all drivers have significant variations to each level of density classes and thus facilitates a huge scope for planners and policy makers to execute various urban development strategies.

1. Introduction

In the last few decades, urban population has increased swiftly from 2.3 billion in 1991 to 6.6 billion in 2020 and is predicted to surpass 60% by 2030 ([World Population Prospects,2019](#)). The rapid increase in urbanization is directed by various factors such as employment, residences, and dispersion because of overcrowding. Expansion of urban areas has also led to a considerable amount of deterioration in agricultural land, loss of green spaces and environmental degradation resulting into disintegration of habitable areas ([Arsanjani et al.,2013](#)). Urban densification is characterized as a continuous increase in urban density as a result of demand for housing units that can be accommodated within existing neighborhoods ([Broitman and Koomen,2015](#)). Recent research has examined densification as a complex and nuanced process that has drew the attention of policymakers and

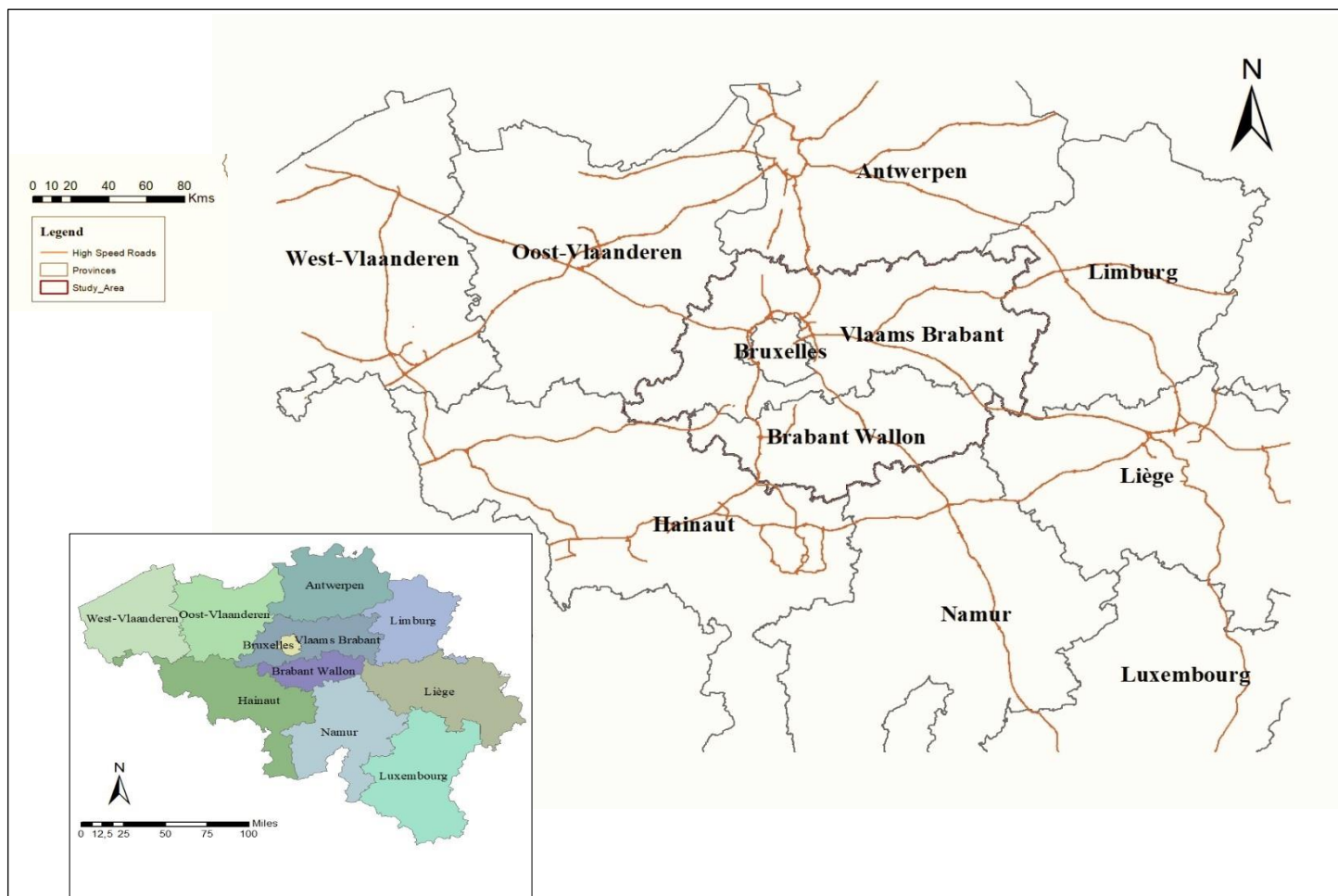
researchers (Wang et al., 2019). Urban densification process is driven by the interaction of several geo-physical factors such as slope and DEM, accessibility indicators as distance to major cities, roads and railways and socio-economic determinants such as population density and household income known as drivers of densification (Table 1) (Mustafa et al., 2018). In this study, an attempt has been taken to analyze the impact of these drivers classified under the category of static and dynamic. This paper also examines the impact of these drivers on different level of urban density using a statistical approach.

2. Material and Methods

2.1 Study area

The study area comprises of the regions of Brussels, Vlaams Brabant and Walloon Brabant of Belgium (Figure 1). This study area occupies 3376 sq. km which is approximately 11% of the total area covered by Belgium. Since the area consists of both vlaams Brabant and Walloon Brabant, so it represents very well the amalgamation of both language - French and dutch. The population size in 2019 was 28,24,813 which was one-fourth of the total population of Belgium.

Figure 1. Study area



2.2 Dependent Variable

The dependent variable for the model is defined by the cadastral data. CAD is a vector dataset, which was provided by the Belgian Land Administration. The construction year is the most important attribute by which the raster built-up maps for the year 2000 and 2016 were created at a cell size of 2 x 2 m. A threshold value of 25 m² was set to consider a cell value as built up according to the average size of a building in Belgium. These raster were further aggregated to a cell value of 100x100 m for speeding up the computing process. The raster was classified further into four density classes based on the threshold value ([Table 1](#)).

Table 1. Range of Built- Up classes in the number of 2*2 cells

Class	Min	Max	% Of cells
Class 0 (non - urban)	0	24	77.21
Class I (Low-density)	25	264	16.38
Class II (Medium- density)	265	735	4.77
Class III (High – density)	736	2500	1.62

2.3 Outline of the Regression model

Over the years, Brussels capital Region and Brabant areas of Wallonia and Flanders have shown rapid urban development due to the affordability and low-cost spaces ([Poelmans and Rompaey,2009](#)). The four urban density classes play a major role for the model as it helps to form a multi density approach. Sixteen major driving factors, acting as the independent variable ([Table 2](#)), were considered and were created as raster grids with a resolution of 100m². Relationship between CAD data and these independent variable was investigated using a multinomial logistic regression (MLR) model. The approach was validated by comparing simulated maps and actual maps for the period of 2000-2016 using the relative operating curve (ROC).

Table 2. Static and dynamic drivers of urban densification

Driver	Name	Unit	Resolution
X1	DEM	m	Cell level
X2	Slope	%	Cell level
X3	Euclidean distance to highways	m	Cell level
X4	Distance to primary roads	m	Cell level
X5	Euclidean Distance to secondary roads	m	Cell level

X6	Euclidean Distance to residential roads	m	Cell level
X7	Euclidean Distance to local roads	m	Cell level
X8	Euclidean Distance to Major cities	m	Cell level
X9	Euclidean Distance to Railways	m	Cell level
X10	Number of households	Number	Statistical level
X11	Mean housing price	€	Statistical level
X12	Population density	inh/km ²	Statistical level
X13	Income per household	Number	Statistical level
X14	Number of employed people	Number	Statistical level
X15	Mean size of housing	Number	Statistical level

3 Results

The transitions from low to medium density or medium to high density classes can be considered as a phenomenon of urban densification. [Table 2](#) shows the results of MLR model for the urban densification process. The model represents odds ratio (OR) which shows the possibilities of these drivers to impact the densification.

The Slope's $OR < 1$ indicates the densification tends to occur largely in flat areas rather than hilly areas. *Distance to cities*, highways and railways contribute negatively towards densification. This means that existing areas within cities, highways or railways does not impact much on future densification process. *Distance to Residential roads* shows a positive impact towards densification whereas *primary roads*, *secondary roads* and *local roads* shows negative impact. This implies that roads directly connected to built up areas are much more substantial in terms of transport network.

The mean *size of housing* and *population* positively impacts phenomenon. This also signifies that the average floor space of a housing is directly proportional to the population within that area. Other socio-economic drivers such as *number of households*, *number of employed people*, *Income per household* and *population density* shows OR close to 1 indicating to a negative impact on densification process for each density classes.

Table 3. Odds ratio for MLR model

Factor	2000-2016	
	Reference Class-I Sample Size: 20000	Reference Class-II Sample Size: 15000
	Class – II	Class- III
DEM	0.9964	0.9871
Slope	0.5875	1.1599
Distance to highways	0.9998	0.9902
Distance to primary roads	0.9348	0.9853
Distance to secondary roads	1.0012	0.9828
Distance to residential roads	1.0013	0.9955
Distance to local roads	0.9976	1.0072
Distance to cities	1.0001	1.0009
Distance to Railways	0.9985	0.9970
Number of households	1.0074	0.9684
Mean housing price	0.9959	0.9726
Population density	0.9992	0.9961
Income per household	0.9999	1.0005
Number of employed people	0.9983	0.9994
Mean size of housing	0.9696	0.9449

4 Discussion and Conclusion

Using a multinomial logistic regression, this paper solely focuses on dynamic of urban densification and the drivers impacting the densification process for the Capital region of Brussels, Flemish and Walloon Brabant of Belgium. In future work, this model can be extended to simulate futuristic densification process in context to residential use or mixed use and by integrating models like cellular automata.

Acknowledgement

This research was funded by the INTER program, co-funded by the Fond National de la Recherche, Luxembourg (FNR) and the Fund for Scientific Research-FNRS, Belgium (F.R.S—FNRS), grant number 19-14016367 — ‘Sustainable Residential Densification’ project (SusDens, 2020–2023).

References

- Arsanjani, J. J., Helbich, M., Kainz, W. & Bolorani, A. D., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation*, Volume 21, pp. 265-275.
- Broitman, D. & Koomen, E., November 2015. Residential density change: Densification and urban expansion. *Computer, Environment and Urban Systems*, Volume 54, pp. 32-46.
- Mustafa, A., Saadi, I., Cools, M. & Teller, J., 2018. Understanding urban development types and drivers in Wallonia: a multi-density approach. *International Journal of Business Intelligence and Data Mining*, Volume 13.
- Poelmans, L. & Rompaey, A. V., January 2010. Complexity and performance of urban expansion models. *Computers, Environment and Urban Systems*, 34(1), pp. 17-27.
- United Nations, Department of Economic and Social Affairs, Population Division, 2019. *World Population Prospects 2019 : Highlights*. [Online] Available at: https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf [Accessed 28 August 2021].
- Wang, L. et al., January 2020. Analysis on residential density dynamics in USA-a case study in southeast Wisconsin. *Sustainable Cities and Society*, Volume 52.

A bottom-up approach for delineating urban areas: Simultaneous optimisation of criteria for building interval and built cluster size minimising the connection cost of buildings by roads

Hiroyuki Usui¹

¹ Urban Engineering, School of Engineering, The University of Tokyo
7-3-1, Hongo, Bunkyo, Tokyo, 113-8656, Japan
81-3-5841-6259, usui@ua.t.u-tokyo.ac.jp

Keywords: urban area, building, nearest neighbour distance, simultaneous optimisation

The management of urban public infrastructure is important for sustainable urban space. As urban populations increase, urban areas are also expanded to provide residents with places to live. In general, the expansion of urban areas is accompanied by a large number of road networks to provide residential places (buildings) with urban public infrastructure, such as a water supply system, sewage system and electric power supply, alongside access to anywhere along the road network. Thus, urban public infrastructure along road networks is simply called *road networks*.

In the late 20th century, Japan experienced this type of growth and was facing urban sprawl, characterised by sets of sparsely distributed buildings along road networks without following urban planning. However, since the beginning of the 21st century, Japan has entered the depopulation era and today faces urban shrinkage. In 2014, the Japanese central government started to require local governments to make their urban areas spatially compact to reduce the management costs of urban public infrastructure.

However, we encounter the following two problems. First, given that contemporary urban areas are officially delineated by top-down approaches merging predetermined basic spatial units (e.g. census units) that meet the population density criterion (e.g. metropolitan statistical areas, MSA, and densely inhabited districts, DID), they depend on how basic spatial units and the population density criterion are set. More importantly, urban areas delineated in this way do not reflect the composition of road networks and the buildings along them (Finance & Swerts, 2020). Rather, urban areas should be delineated based on a bottom-up approach focusing on the locations of buildings, enabling us to compare the present and the future image of urban areas in a consistent way and evaluate whether or not the latter will be made more spatially compact than the former (Usui and Perez, 2020).

As alternatives to conventional definitions of urban areas based on populations, several methods based on road networks have been proposed, such as *Natural Cities* (Jiang & Jia, 2011; Jiang & Liu, 2012) and *Hierarchical Percolations* (Arcaute et al., 2016). Based on the locations of buildings, the following methods for delineating urban areas are among those to have been proposed: a graph-based approach incorporating computation of the local Moran's I statistic (Caruso et al., 2017); density-based spatial clustering of applications with noise (DBSCAN), a clustering approach (Arribas-Bel et al., 2019); *MorphoLim*, a fractal approach (Tannier et al., 2011); a CCA and percolation-based approach (Behnisch et al., 2019); and *Natural Cities* applied to the locations of buildings (Jiang, 2019; Usui, 2019). Montero et al. (2021) have compared methods of delineating urban areas by applying *Natural Cities*, *MorphoLim* and *Hierarchical Percolations* to building centroids and road network intersections.

However, these methods fail to account for the relationship between delineated urban areas and the average total management costs of the road networks connecting buildings. Hence, we lack a norm that can be compared not only with the present urban areas in terms of cost efficiency. This is the second problem. In general, the more spatially compact urban areas are in terms of the road networks and buildings along them, the more their management costs can be reduced. In sprawling urban areas, a set of sparsely distributed buildings emerges as a small settlement and the nearest neighbour distance (NND) of a building here tends to be longer than in compact urban areas, which means that urban management is inefficient because the management costs of road networks in sprawling urban areas are higher. Conversely, as the NND and settlement size are shorter and greater, delineated urban areas are spatially compact and cost-efficient.

To evaluate whether or not current urban areas are spatially compact and cost-efficient, a consistent method for delineating urban areas needs to be developed, not only focusing on the locations of buildings and the road networks connecting them, but also considering their management costs. Hereafter, management cost is simply called *cost*.

In general, road networks are classified into either global or local networks (Bertaud, 2018; Usui, 2019). National and prefectural road networks are classified as global networks. This is because they play an important role in connecting cities and settlements based on regional- and national-scale planning. Hence, the total lengths of national and prefectural road networks do not depend on the number of buildings and their locations. On the other hand, road networks other than national and prefectural road networks are classified as local networks and play an important role in connecting buildings that are adjacent to one another in a settlement. Therefore, the costs of global and local road networks can be regarded as fixed and variable, respectively.

Following this classification, Usui (2019) has proposed a two-step method based on the concept of *Natural Cities* for delineating urban areas, clustering buildings whose interval (the NND of each building) is shorter than a criterion for the maximum NND of each building as built clusters (equivalent to settlements) and then delineating built clusters whose size (the number of buildings in its cluster) is greater than the minimum criterion (average size). Subsequently, focusing on a trade-off relationship between the average fixed and variable costs of global and local road networks, the criterion for the NND is given as the optimal solution that can minimise the average total cost of road networks.

However, the criteria for the NND of each building and the built cluster size have yet to be simultaneously optimised in terms of the average total cost of the road networks connecting buildings. Therefore, the objective of this study is to answer the following question: *how can one simultaneously optimise the criteria for the maximum NND of each building, r_c , and the minimum built cluster size, m_c , that minimise the average total cost of the road networks connecting buildings?*

The main conclusions of this study are as follows: (1) there exists a pair of r_c and m_c that can minimise the average total cost and optimal urban areas can be delineated in terms of cost efficiency; and (2) in the empirical study region (Chiba prefecture, to the east of Tokyo), the optimal r_c and m_c are 48 metres and 42, respectively. Such optimal urban areas are expected to provide urban planners with a norm that can be compared not only with the present but also the future image of urban areas. To make present urban areas spatially compact, m_c as well as r_c is important in removing small built clusters, where buildings are sparsely distributed as a result of urban sprawl.

Acknowledgments

This work was supported by Japan Society for the Promotion of Science (20H02383).

References

- Arcaute, E., Molinero, C., Hatna, E., Murcio, R., Vargas-Ruiz, C., Masucci, A. P., & Batty, M. (2016). Cities and regions in Britain through hierarchical percolation. *Royal Society Open Science*, 3(4), 150691.
- Arribas-Bel, D., Garcia-López, M. À., & Viladecans-Marsal, E. (2019). Building(s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics*, 103217.
- Behnisch, M., Schorcht, M., Kriewald, S., & Rybski, D. (2019). Settlement percolation: A study of building connectivity and poles of inaccessibility. *Landscape and Urban Planning*, 191, 103631.
- Bertaud, A. (2018). *Order without design: How markets shape cities*. MIT Press.

- Caruso, G., Hilal, M., & Thomas, I. (2017). Measuring urban forms from inter-building distances: Combining MST graphs with a Local Index of Spatial Association. *Landscape and Urban Planning*, 163, 80–89.
- Finance, O., & Swerts, E. (2020). Scaling laws in urban geography: Linkages with urban theories, challenges and limitations. In D. Pumain (Ed.), *Theories and models of urbanization geography, economics and computing Sciences* (pp. 67–96). Springer.
- Jiang, B. (2019). Natural cities generated from all building locations in America. *Data*, 4(2), 59.
- Jiang, B., & Jia, T. (2011). Zipf's law for all the natural cities in the United States: A geospatial perspective. *International Journal of Geographical Information Science*, 25(8), 1269–1281.
- Jiang, B., & Liu, X. (2012). Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. *International Journal of Geographical Information Science*, 26(2), 215–229.
- Montero, G., Tannier, C., & Thomas, I. (2021). Delineation of cities based on scaling properties of urban patterns: A comparison of three methods. *International Journal of Geographical Information Science*, 35(5), 919–947.
- Tannier, C., Thomas, I., Vuidel, G., & Frankhauser, P. (2011). A fractal approach to identifying urban boundaries. *Geographical Analysis*, 43(2), 211–227.
- Usui, H. (2019). A bottom-up approach for delineating urban areas minimizing the connection cost of built clusters: Comparison with top-down-based densely inhabited districts. *Computers, Environment and Urban Systems*, 77, 101363.
- Usui, H. & Perez, J. (2020). Are patterns of vacant lots random? Evidence from empirical spatiotemporal analysis in Chiba Prefecture, east of Tokyo. *Environment and Planning B: Urban Analytics and City Science*. (published online 20201006).

Special Session SS09

4 November 4:00 am-5:40 pm (GMT)

Special Session in honour of Martin Charlton

Proponents

Chris Brunsdon, Lex Comber, Paul Harris and Richard Kingston

Whats in a Name? A Spatial Analysis of British Public House Names

Chris Brunsdon*

National Centre for Geocomputation, Maynooth University, Ireland

1 Introduction

This is a short paper illustrating how spatial data visualisation and analysis can be used to explore patterns in the naming of pubs in Britain. It draws on a number of computational tools including those for spatial data modelling, visualisation and text processing, all of which are available in R. The focus here is mainly exploratory, although some hypothesis testing and modelling is considered. As well as identifying trends in the naming of pubs spatial patterns are also identified, where certain pub names or more general styles in pub naming are clustered around certain parts of the country. Martin Charlton was a lover of pubs, and of spatial data analysis - it is hoped this paper would meet with his approval.

*christopher.brunsdon@mu.ie

2 Data source

The data here was obtained from <https://www.getthedata.com/open-pubs> - a web portal for a number of data sets created from open data providers - some are mixtures from more than one provider. This data comes from the Food Standard Agency's Food Hygiene Ratings (<https://ratings.food.gov.uk/open-data/>) and the ONS Postcode Directory (<https://geoportal.statistics.gov.uk>). It is a data frame with columns as listed in table 1.

Field Name	Possible Values	Comments
<code>fsa_id</code>	int	Food Standard Agency's ID for this pub.
<code>name</code>	string	Name of the pub.
<code>address</code>	string	Address fields separated by commas.
<code>postcode</code>	string	Postcode of the pub.
<code>easting</code>	int	X coordinate of pub. GB National grid.
<code>northing</code>	int	Y coordinate of pub. GB National grid.
<code>latitude</code>	decimal	Latitude
<code>longitude</code>	decimal	Longitude
<code>local_authority</code>	string	Local authority this pub falls under.

Table 1. Variables provided in FSA data set

There is some redundancy here; the `easting` and `northing` variable pair and the `latitude` and `longitude` pair both contain point locations for each of the pubs. This table can be represented as a point 'SimpleFeatures' spatial data object. Also used - and obtained from <https://geoportal.statistics.gov.uk> is a set of boundaries for counties and unitary authorities in the UK. In figure 1 the locations of the pubs (as points) are drawn over county boundaries with an alpha value of 0.02 - giving an overview of their geographical distribution.

3 A Visual Exploration

This exploration will consider different the words used in the names of pubs, and their relative popularity throughout England, Scotland and Wales. A simple way to consider this is to map the most commonly occurring words in pub names by local authority in Great Britain. This can be achieved in R by making use of the `tidytext` and `stringr` libraries. Together, these tools can remove *stopwords* from pub names - these are words such as 'the,' 'and,' 'of' and so on which appear very commonly in all types of speech and text, but convey little contextual information, and then count the number of occurrences of each word in the name of each pub. The counting is done on separately for each local authority, and the results are mapped in figure 2. From this it can be seen that most of the popular words are terms for a pub - such as 'inn,' 'bar' and 'hotel' - the later often used when an establishment primarily functions as a pub, but does also offer guest rooms. Some spatial patterns are apparent, perhaps most notably that popularity of the term 'bar' seems to cluster in Scottish local authorities, whereas in much of England and Wales 'inn' is more popular. In England places where 'bar' is more popular tend to be in urban areas, such as some

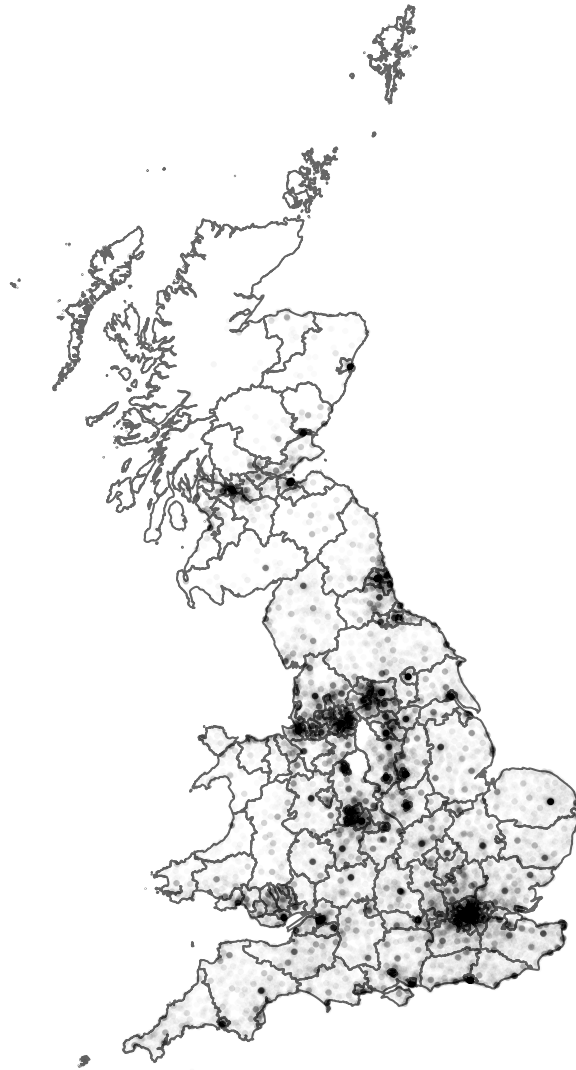


Figure 1: Locations of pubs in Great Britain

of the London boroughs - and also in Tyne and Wear. There is also a strip from a little north of London through Cambridge and into East Anglia, where the word 'house' is popular.

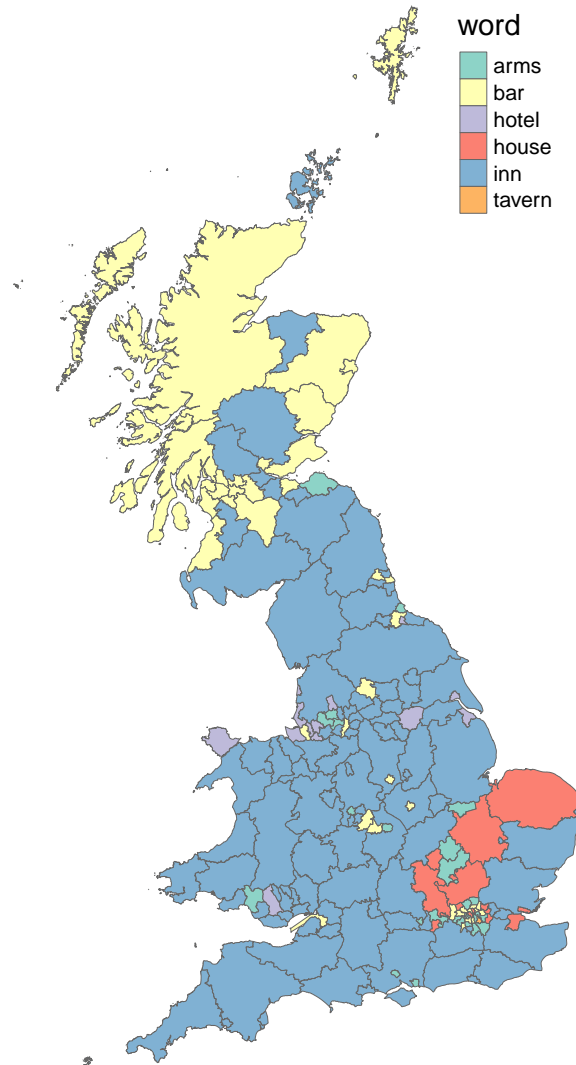


Figure 2: Most common word in pub name by Local Authority

Some further explorations can also be carried out. For example a number of British pubs have the word 'king,' 'queen,' 'prince' or 'princess' in the title. It may be interesting to investigate this spatial distribution. An initial exploration here detected any pub names containing 'king,' 'queen' 'prince' or 'princess.' From the full sample of 41127 pubs, there are 1394 containing these words. A simple visualisation is to draw a dot map of the locations of the subset, and compare this to a dot map of the complementary set (ie non-royal pub names). Each dot map draws the dots with a degree of transparency so that areas of higher density can be seen. For the complementary subset dot map, a random subsample of the same size as the number of royal named pubs is used, so that intensity levels of the dot maps are comparable. This is shown in figure 3.

Although in general the dot patterns reflect urban areas, differences in intensity are noticeable - in particular 'Royal' pubs are more prevalent in East Anglia, and less so in North East England.

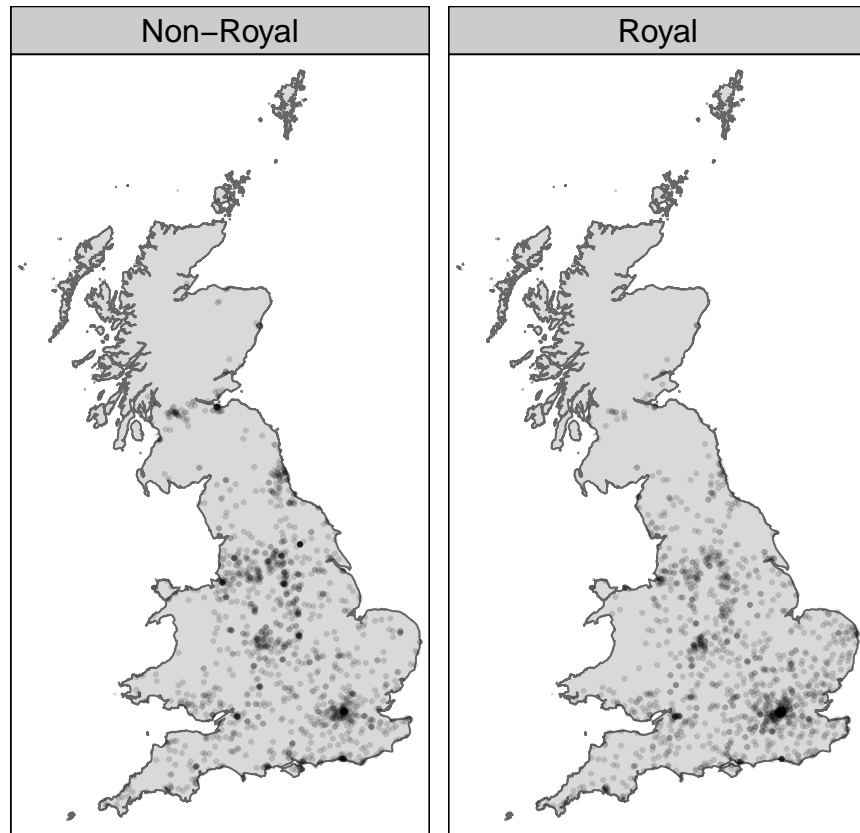


Figure 3: Dot Density Maps for Pubs With 'Royal' Names vs. general distribution.

4 A Statistical Analysis Example

This can be analysed in more detail. A possible model for this is to consider a binary variable `royal` for each pub in the data set, where a value of 1 implies ‘king,’ ‘queen’ or ‘prince’ or ‘princess’ in the title, and a value of zero implies none of these words appear. A statistical model of the form

$$\text{Royal} \sim \text{Binomial}(\theta(x, y))$$

is proposed where (x, y) are the easting and northing of the pubs location. Mapping $\theta(x, y)$ gives an indication of regional trends in the likelihood of a pub have a royal-themed name. Here, a non-parametric approach is taken, where

$$\text{logit}(\theta(x, y)) = f(x, y)$$

and the function f is generated by a Gaussian process. The Gaussian process part of the model is similar to the assumptions made about the underlying trend surface being estimated using Kriging (Wackernagel 2003).

Here, the model is calibrated using the `gam` function in the R package `mgcv` (Wood 2003). Once calibrated, the value of $\theta(x, y)$ can be sampled at points on a hexagonal lattice, and mapped to visualise the underlying pattern - this is done in figure 4.

The shows that the probability that a pub name contains a royal word varies from around 0.01 to 0.08. There is generally a trend away from royal names as one heads into Scotland - and a notable trend in favour of such names in s strip from the north of London into East Anglia. The latter was noticed in the dot plots shown earlier. This trend also coincides geographically with a trend in the general most popular words - where ‘inn’ sees less favour in a similar region stretching towards East Anglia. Also in common is the distinction between England and Scotland. Finally, the South West of England shows a trend in favour of royal names, and to a lesser extent so does the North West.

Finally, the validity of the trend surface can be considered, by comparing the Akaike Information Criterion (AIC) (Akaike 1974) of the Gaussian process model to the ‘benchmark’ model where $\theta(x, y)$ is a constant. Doing this, we have

Model	AIC	Δ AIC
Benchmark	1723.2	178.6
Gaussian Process	1544.6	

Table 2. AIC values for the Gaussian and Benchmark Models

Thus, the AIC of the Gaussian process model is considerably lower than that of the benchmark model, suggesting it has notably better prediction power.



Figure 4: Gaussian Process-Based Trend Surface for Probability of a 'Royal' Pub Name

5 Closing Comments

In this abstract, I have included one example of the kind of analysis that can be carried out, as well as a more general exploration. My intention in the talk is to try to make it reasonably close to something Martin could regard as an ideal talk, by including further examples, making reference to both trains and cathedrals.

References

- Akaike, H. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19: 716–23.
- Wackernagel, Hans. 2003. *Multivariate Geostatistics*. Springer.
- Wood, S. N. 2003. "Thin Plate Regression Splines." *Journal of the Royal Statistical Society (Series B)* 65 (1): 95–114.

A spatial analysis of prescribing data: revisiting the ‘Geographical Analysis Machine’

Lex Comber, Paul Harris, Chris Bunsdon

September 2021

1. Introduction

The Geographical Analysis Machine (GAM) was proposed by Openshaw et al (1987) in response to the growth of spatial data and the dearth of spatial methods for handling such data. The GAM aimed to make best use of the growing number of spatial datasets, to avoid making ill founded assumptions about spatial data and to develop effective and unbiased search techniques capable for testing and generating hypotheses—i.e. to “exploit both the growing richness and size of spatial datasets and the great computational power of modern computers.” Of relevance to the quantitative geography community it sought to overcome the problem that “purely statistical methods are increasingly being viewed as unable to provide either accurate or unbiased scientific answers to the problem of detecting clusters”. Thus the development of the GAM reflected a desire to:

- test all possible hypotheses for a specific geographical problem;
- exclude all bias by removing any (spatial) selectivity;
- enumerate the universe of all possible hypotheses, that by definition encapsulate all (spatial) knowledge about a problem, known and unknown.

It does this by treating all locations equally in order to identify any meaningful patterns through the examination of different cluster related spatial hypotheses. It seeks to identify clusters in the data without suggesting locations around which a focused analysis could be undertaken.

The key idea of the GAM was that many clusters (circles in the original implementation) at different spatial scales of analysis would be found in locations where an excess of the process being examined was present.

2. Data and Methods

Open prescribing data have been available for a number of years now in England. This records each prescriptions issued by GP practice, its BNF code, cost, quantity as well the month and year, but contains no information about the patient. Additional data describe the details for each GP practice, including its address, which allows the surgery to be located. The spatial distribution of the home locations of patients registered at each GP surgery can be obtained from the NHS data portal. A full description of the data and methods for its analysis are provided in Comber et al (2001) who used it to examine the Geography of Misery via antidepressant prescribing patterns, and in so doing unpicked the differential impacts of the post 2008 austerity experiment in the UK.

Here a similar approach is applied to examine the prescribing of small anal plugs, only because Martin suggested it in our discussions and first examinations of the prescribing data, which eventually led to the the Comber et al 2021 paper.

Data for Anal Plugs were extracted from the prescribing data for the years 2011 to 2018. The main products are the Peristeen® Anal Plug and the Renew® Insert - see <https://openprescribing.net/dmd/?q=2202> for details. These come in 2 sizes (small and large) and the BNF codes allow the different sized product to be

identified. The number of small and large items prescribed by each GP practice were determined and the proportion of patients registered in each LSOA for each GP practice was calculated. These proportions were then used to allocate counts of both types of anal plug to each LSOA. The sum of small and large plugs in each LSOA form the underlying population, and examining the distribution of small anal plugs (indicating tight arses) is the purpose of this analysis¹.

The Observed and Expected distributions of prescriptions for small sized anal plugs are shown in Figure 1, identifying some highly localised areas with high Expected and or Observed counts.

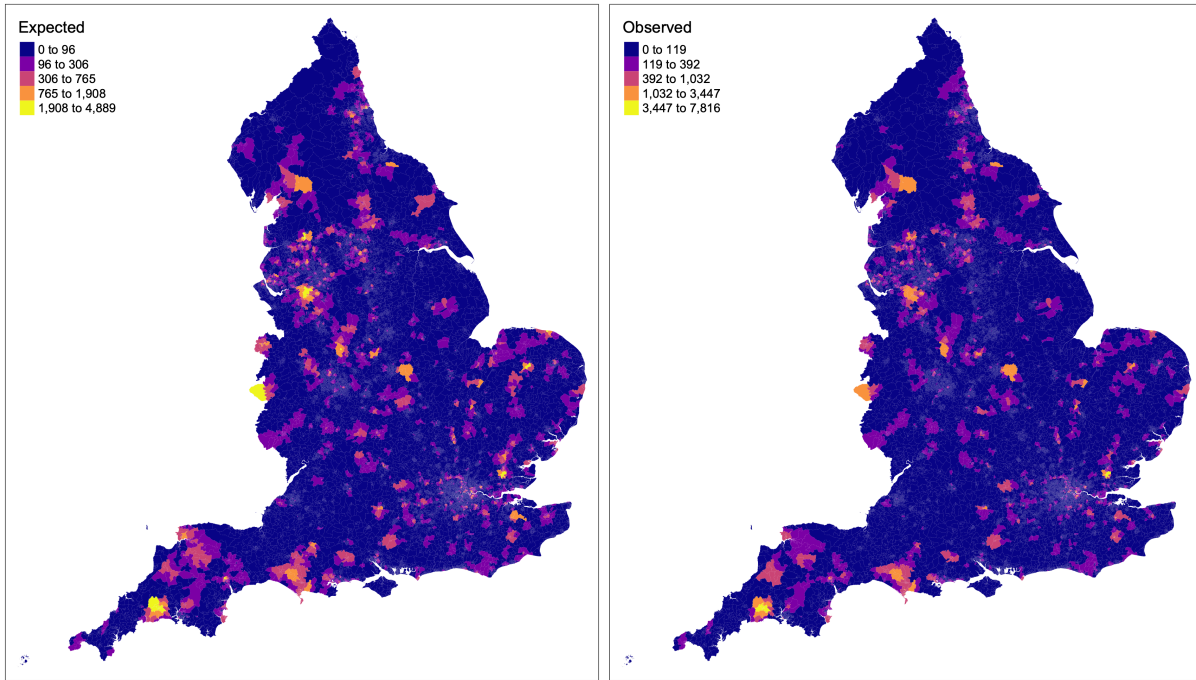


Figure 1: The distribution of Expected and Observed small anal plugs (2011-2018).

3. Analysis and Results

The GAM is an exploratory approach for identifying spatial clustering in the incidence of some phenomenon relative to some form of expected distribution, and a threshold of statistical significance, here 0.002 to minimise the Type I errors. A basic GAM can be implemented in the `DCluster` R package (Gómez-Rubio et al., 2015) and this was used to evaluate the Expected and Observed populations of small size anal plugs under a sequence of search windows, 1-20km in steps of 1km, and the locations identified as significant were merged. These are shown in Figure 2.

Clearly there are distinct spatial patterns on the South coast, in the North (running from Liverpool in the North West, Manchester, Leeds and up to Newcastle in the North East) and in various places in the Midlands (Gloucester, Hereford, Leicester) and to the East of the London area. These could be associated with lifestyle, age (retired people on the south coast for example) and other populations could be investigated in depth - perhaps in relation to the characteristics of the underlying population in relation to the causes of anal seepage: irritable bowel syndrome, inflammatory bowel disease, proctitis, faecal overflow, pelvic floor

¹In arsecons (see <https://www.angelfire.com/home/mkazic/Profanisaurus.htm> a source that was frequently referred to by Martin), tight arse is indicated with (!), slack arses requiring larger plugs as (_o_), and of course normal arse is (_!_), lard arse (_!_) and sore arse (_*_)

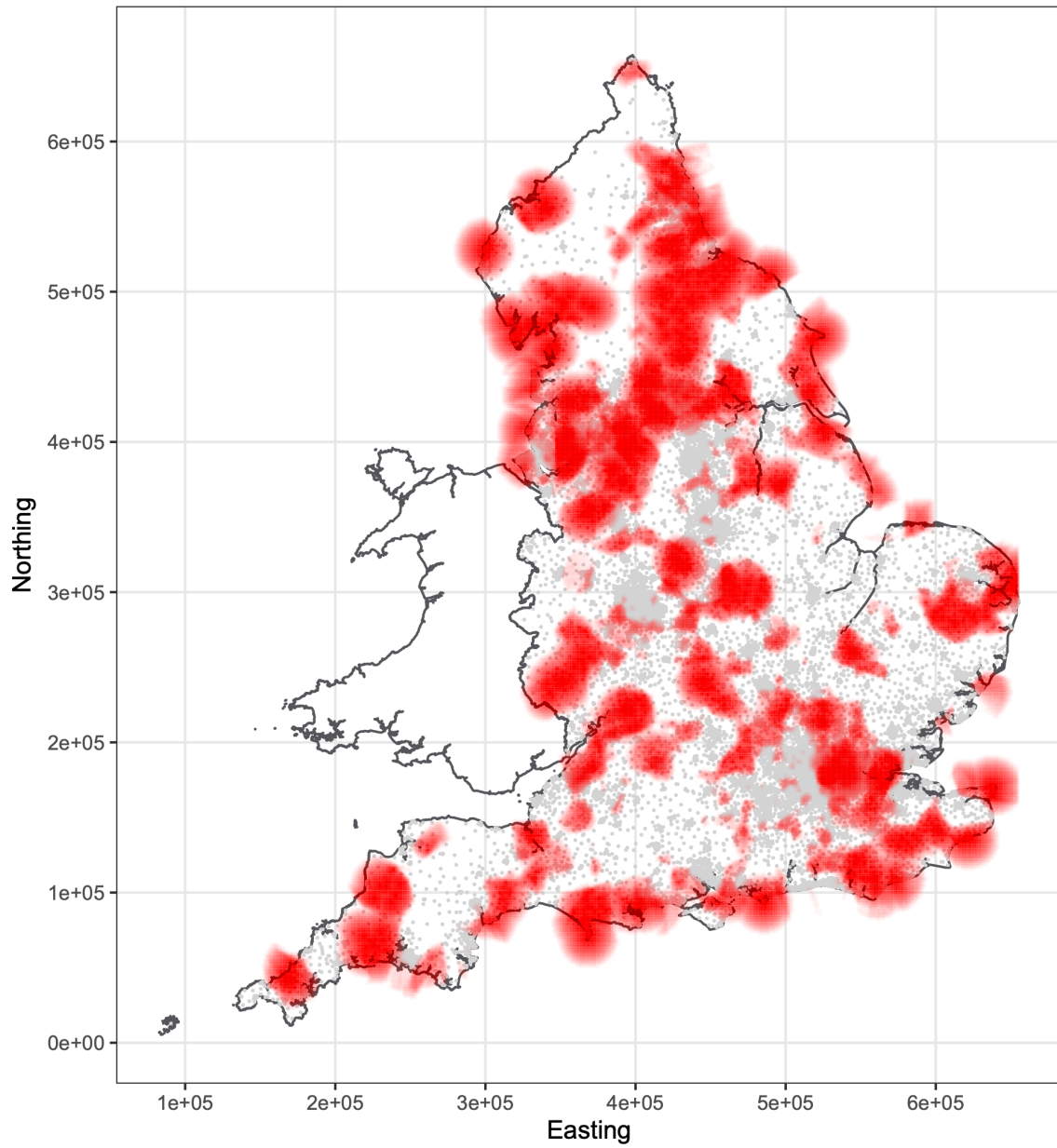


Figure 2: The clusters identified by the GAM for *small* anal plugs (2011-2018), over an Expected distribution of *all* anal plugs.

weakness, rectum emptying disorders, muscle and nerve damage as well as neurological disorders, loss of rectal storage capacity, haemorrhoids, rectal prolapse, etc.

4. Reflection

The reason for selecting an old but at the time novel approach to spatial analysis in this submission was because it encapsulates some key considerations when undertaking analyses of spatial data: exploratory analysis, investigation of scales of influence, spatial hypothesis testing, local analyses. The GAM and the 1987 paper was not without criticism with little control of multiple testing, the at-risk population could be difficult to define for variations (e.g. by age or gender). But it was the first approach to identify clusters of a rare disease, leukaemia, with the data available. It avoided any pre-selection bias, could analyse fine spatial resolution data, and avoided both fixed-size and fixed population sub-regions.

The original 1987 GAM paper was included in the *Classics from IJGIS* and both the paper and Martin's commentary of it are freely available online². The latter gives a real sense of how he approached his research and an insight into the man. The paper and commentary when read today provide a relevant counter narrative to current obsessions over often inference- and theory-free approaches to data mining, ever fancier non-linear regression approaches and the pursuit of ever tighter model fits. The importance of the need for spatial exploration of spatial data was central to Martin Charlton's work. His work with Stan Openshaw on the GAM was an early contribution, before he was involved in the development of Geographically Weighted Regression (Brunsdon et al, 1996) and its refinements. Such considerations were very much to the fore in the last piece of research (unfunded) we worked on together (Comber et al 2021) that has provided the inspiration for this contribution.

Vale Martin.

References

- Brunsdon C, Fotheringham AS, Charlton ME (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 28:281-98.
- Charlton ME (2006). A mark 1 geographical analysis machine for the automated analysis of point data sets: twenty years on, pp.35-40 in *Classics from IJGIS: twenty years of the International Journal of Geographical Information Science and Systems*. CRC Press, London.
- Comber A, Brunsdon C, Charlton M and Cromby J (2021). The changing geography of clinical Misery in England: lessons in spatio-temporal data analysis, pp. 64-77 in Birkin M., Clarke G.P., Corcoran J., Stimson R. (eds), *Big Data Applications in Geography and Planning: An Essential Companion*, Edward Elgar, London.
- Gómez-Rubio V, Ferrándiz-Ferragud J, López-Quílez A (2015). *DCluster: functions for the detection of spatial clusters of diseases*. <https://CRAN.R-project.org/package=DCluster>, R package version 0.2-7
- Openshaw S, Charlton M, Wymer C, Craft A. (1987). A mark 1 geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information System*, 1:335-58.

²<http://ndl.ethernet.edu.et/bitstream/123456789/50204/1/18pdf.pdf>

Training Sets and Train Sets: How organs, pubs and railways inter-relate

Steve Carver and Chris Brunsdon

Introduction

In this paper we deal with topics that were close to Martin's heart... organs, pubs and trains. By integrating spatial analysis and Boolean logic with data on the relative locations of cathedrals, pubs and railway stations we demonstrate how a simple spatial decision support system can be used to optimise the combination of these three seemingly disparate topics. We draw on a variety of computational tools and methods including traditional GIS analyses, terrain models and networks, Boolean logic and text processing in R. Results are presented as a series of maps and choices of watering hole determined by walking distance between origin (cathedral) and destination (railway station) and the time available between end of organ practice and last train home. This paper forms a natural progression to the paper by Brunsdon, Harris and Comber "What's in a name?" in this same series by putting the pub data to a good use with other data via an integrated analysis to create a spatial logic to pub search and selection that incorporates space-time optimisation and pub-name aesthetic.

Data sources

The data used here were obtained from multiple sources and include a dataset showing the names and locations of pubs in the UK, the location of cathedrals and their organs, the location of railway stations and railways, a road and footpath network, open water areas (including rivers, lakes and coastal areas) and a digital terrain model. All data are open source and freely available for download. These are summarised in Table 1.

Table 1. Data sources

Description	Entity type	Source	Link
Pubs	Point	Food Standard Agency's Food Hygiene Ratings	https://ratings.food.gov.uk/open-data/
Cathedrals	Point	Wikipedia List of cathedrals in the United Kingdom	https://en.wikipedia.org/wiki/List_of_cathedrals_in_the_United_Kingdom
Railway stations	Point	Ordnance Survey VectorMap District	https://www.ordnancesurvey.co.uk/business-government/products/vectormap-district
Railways	Line	Ordnance Survey VectorMap District	https://www.ordnancesurvey.co.uk/business-government/products/vectormap-district
Roads and footpaths	Line	Open Street Map	https://www.geofabrik.de/data/shapefiles.html
Open water	Polygon	Open Street Map	https://www.geofabrik.de/data/shapefiles.html
Terrain	Raster	Ordnance Survey Terrain 50	https://www.ordnancesurvey.co.uk/business-government/products/terrain-50

Personal knowledge and visualisation of the pattern of pubs in the FSA pub dataset suggest that the pubs data are incomplete, with several pubs of note - for example The Clachaig Inn, Glencoe, The Vaynol Arms, Llanberis, The Fox House near Burbage, and The Old Dungeon Gill, Langdale - all being missing from the list. This seems particularly to be the case in more remote rural areas. One reason for this is that the database is based on hygiene checks, but some pub results may not have been recorded, and there may be omissions. In urban areas this may result in a small number of missing pubs, but in rural areas, where population density – and also pub density – is lower, omission may lead to the only pub in a village (or a remote location) being missing. In addition, some pubs may be closed (even if temporarily) again resulting in omission from the database. Nonetheless, the pubs database appears more reliable in the urban areas around cathedrals and railway stations where this analysis is focused.

Methods

Brunsdon, Harris and Comber (this issue) have shown how an analysis of pubs and their names can be used to visualise patterns in the popularity of pub names relative to selected geographies. The aim here is to build on this work and combine the pub data with additional information on cathedrals and railway stations to develop a simple spatial decision support system (SDSS) to assess the location of pubs in the database relative to origin (cathedral) and destination (railway station) points and the time available to enjoy them. The basic logic behind such a model is to examine the overlap between time taken to walk from the origin and to the destination and the relative locations of pubs enroute together with a version of the textual analysis of pub names described by Brunsdon, Harris and Comber when choosing which pub to visit. The number of pubs frequented between origin and destination is taken to be a function of distance (as time taken to walk) and time available (between end of organ practice and last train home). A simple Boolean overlay model is used to determine the number and location of pubs visited.

Walking times from origin and destinations are modelled using an implementation of Naismith's Rule (1892) in ArcGIS (Carver and Fritz, 2000). This well-known walker's rule of thumb assumes a speed of 3mph on the flat and an additional hour for every two thousand feet of climbing. Langmuir's Correction (1984) is used to compensate for increased speed walking downhill by subtracting 10 minutes for every 300 feet of descent. A terrain model is used here together with a road and footpath network to calculate walking times from specified origin/destination points using the PATHDISTANCE tool. Data on rivers, lakes and coastal waters are used as barriers (NoData) to constrain walking routes by forcing solutions to "go around" and cross only at bridges. The resulting surface predicts walking times and can be used with a backlink raster to model least cost paths or shortest routes should this be necessary. This approach has frequently been used to model remoteness as time taken to walk from the nearest road in wilderness quality and character mapping (Carver et al., 2012; Carver et al., 2013) and route choice fell-running and orienteering races (Scarf, 2007) by incorporating variations in walking speeds depending on land cover. The approach is adapted here for urban walks using the road and path network in place of land cover and assuming a uniform walking speed of 5 km/hr to define two isochrone surfaces, one showing walking time from the nearest cathedral, and one showing walking times from nearest railway station.

If we assume a search area within a maximum walking time of half an hour from the origin (cathedral) and quarter of an hour to the destination (station) as defined by the above approach, we can define two (hopefully) overlapping envelopes wherein to search for suitable watering holes from

the pub database. A novel application combining time available with these isochrone envelopes can be used to select which and how many pubs to visit using simple Boolean logic as shown in Figure 1.

Pubs within these overlapping regions or envelopes are selected using point-in-polygon overlay. Taking the example of Brunsdon, Harris and Comber (this session) it seems logical that several of the pubs within the specified origin/destination envelopes ought to have appropriate ecclesiastical or railway-related names and might add further logic in the aesthetic of pub choice according to name when multiple choices are available. This merits further investigation. In R it is possible to search a database of pub names using the ‘stringr’ package, which can identify character values meeting regular expression patterns (for example ‘station’, possibly with and s at the end, and with upper or lower case characters). For example, railway-related pubs may contain patterns such as ‘railway’ ‘station’ ‘train’ ‘engine’, while ecclesiastically named pubs may have names matching ‘church’ ‘abbey’ ‘priory’ ‘chapel’ ‘monastery’ ‘vicarage’ ‘vicar’ ‘mitre’ ‘bell’ ‘bishop’ ‘monk’ ‘angel’ ‘nun’ ‘pilgrim’ ‘pope’. There are also a few one-off names punning on trains such as the Piston Broke, Shoreham, or the ‘Head of Steam’ chain of pubs that are in or near to railway stations. There may also be some outliers – possibly reflecting historical events, such as the Beeching closure of railway lines and stations in the early 1960s in the UK. This has occasionally resulted in ‘ghost’ railway-named pubs that were once close to a railway station, but where the station is no longer there.

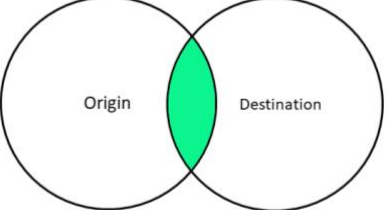
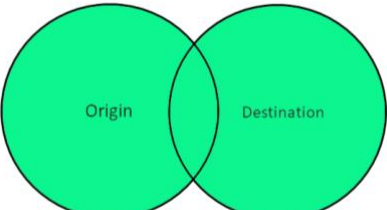
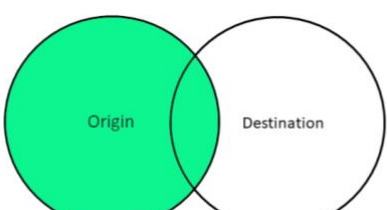
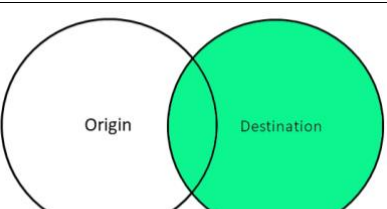
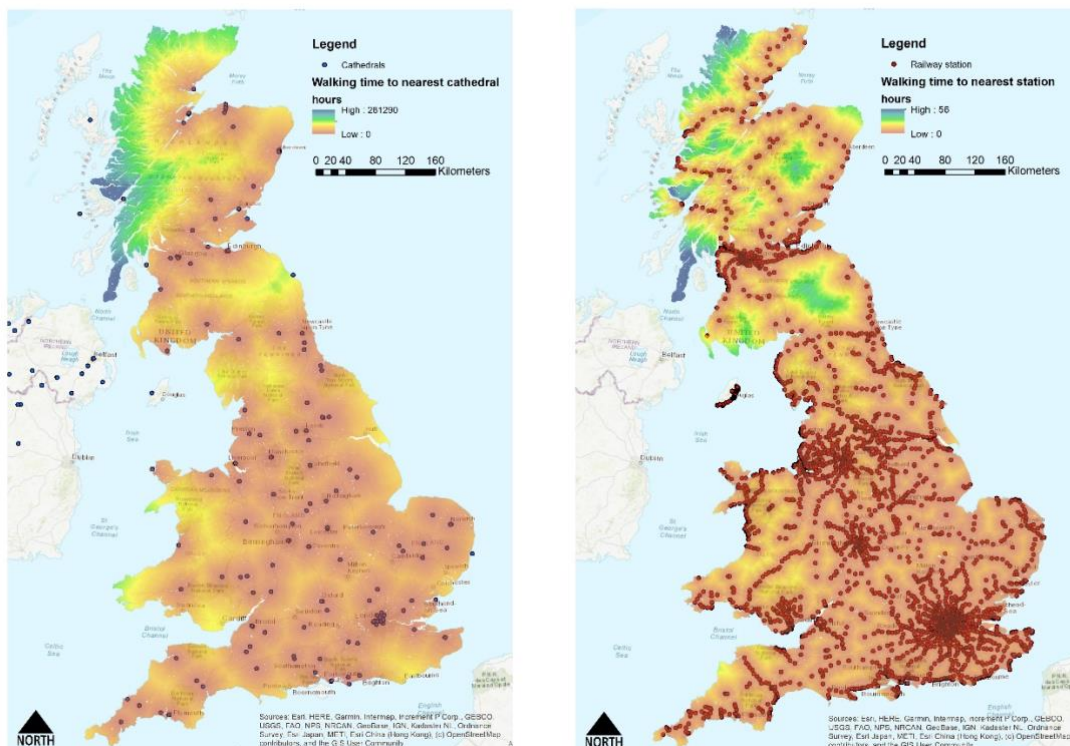
Rule	Venn diagram	Time available
Boolean AND		Only time for the one
Boolean OR		Session
Boolean NOT (option 1)		A couple with the verger
Boolean NOT (option 2)		One for the road

Figure 1. Boolean logic applied to pub selection between origin and destination

Results

The maps in Figure 2 show the isochrone surfaces for walking times from cathedrals in the UK (origin points) and railway stations (destination points). These have been reclassified using maximum walking times of 30 and 15 minutes, respectively to define the origin and destination search envelopes shown in Figure 3 using a regional example for clarity. Pubs falling into these search envelopes are shown and ecclesiastical and railway-themed pub names highlighted.



Figures 2 and 3. Isochrone surfaces for walking times from cathedrals and railway stations

Figure 4 shows an example of the Boolean decision logic applied to these data and how it can be used as a simple SDSS in making choices as to the pub(s) in which to imbibe depending on time available between end of organ practice and the last train home.

Conclusion

In this paper we have tried to show how open data can be combined using spatial analysis methods in GIS and R to develop a simple SDSS aiding frequency and choice of pub taking both walking distance (expressed time from and to origin/destination points) and the aesthetic of pub names relative to origin and destination. While not perfect, the data available on pub and cathedral locations when applied with data on road/footpath networks, terrain and railways stations with some simple Boolean logic can generate a range of feasible solutions to this often tricky and ill-structured decision problem.

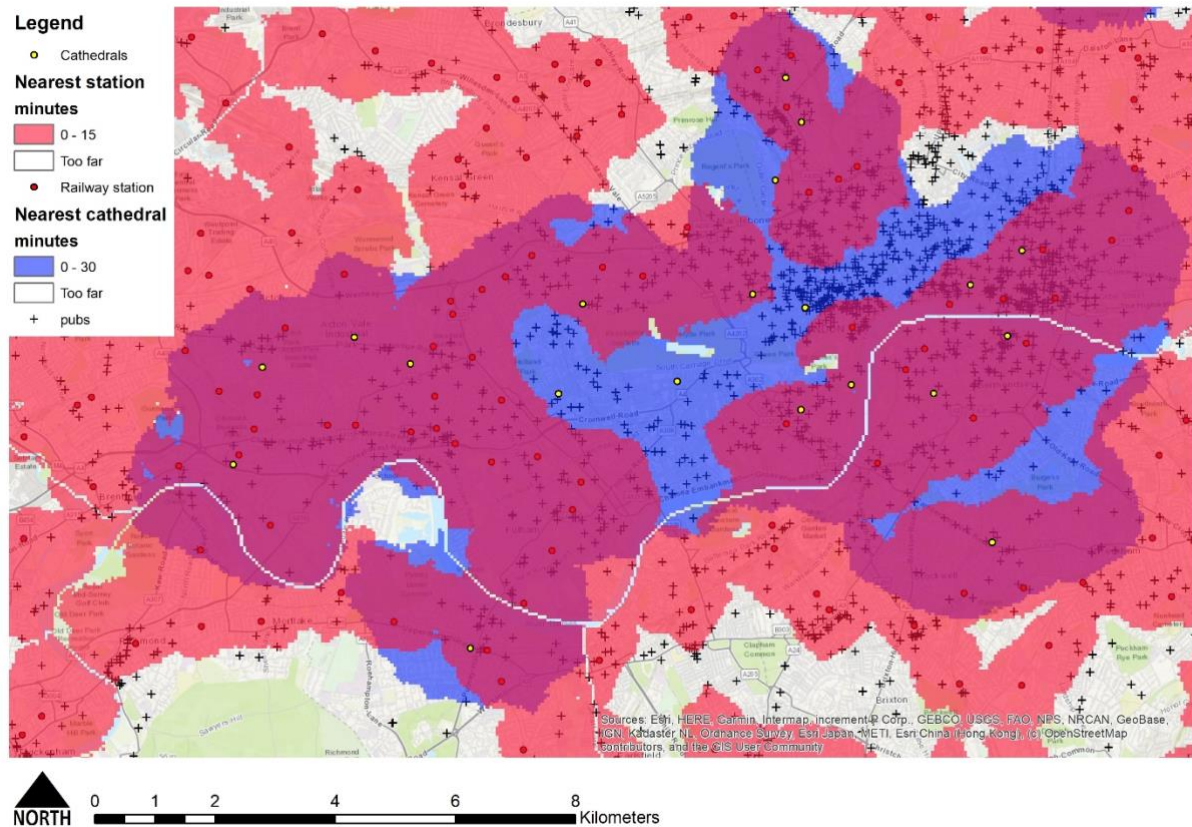


Figure 4. Spatial intersection of organs, pubs and stations

Future work could incorporate information and reviews from CAMRA webpages linked to the pubs database as a means of further enhancing the SDSS with crowd sourced information on range of beers available, ambiance, opening times and other essential information. Ideal routes could be plotted using backlink rasters to determine least cost paths, minimising the time taken walking between origin, pub(s) and destination. Other applications might be found in both urban and rural pub walking tours and managing liquid refreshments on recognised long-distance footpaths such as the Coast to Coast and Pennine Way.

References

- Carver, S. and Fritz, S., 2000. Munro-bagging with a computer...? Naismith's rule and the long walk in. *The Scottish Mountaineering Club Journal*, 36, p.190.
- Carver, S., Comber, A., McMorran, R. and Nutter, S., 2012. A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and urban planning*, 104(3-4), pp.395-409.
- Carver, S., Tricker, J. and Landres, P., 2013. Keeping it wild: Mapping wilderness character in the United States. *Journal of environmental management*, 131, pp.239-255.
- Scarf, P., 2007. Route choice in mountain navigation, Naismith's rule, and the equivalence of distance and climb. *Journal of Sports Sciences*, 25(6), pp.719-726.

Parallel Session PSD2

4 November 4:00 am-5:40 pm (GMT)

Energy and Industry

Studying social inequality in urban energy use in the UK

Fernando BENITEZ-PAEZ¹; Mark BIRKIN¹

¹The Alan Turing Institute

British Library, 96 Euston Road, London, England, NW1 2DB, United Kingdom

+44 (0)20 3862 3352, fbenitez@turing.ac.uk

Keywords: inequality, urban energy use, social policy, urban analytics

The coronavirus pandemic has highlighted the fragility of societies and economies around the world. Political instability, economic uncertainty and climate change are as likely as health to trigger the next global crisis. In light of such threats, the creation of greater resilience in social systems is an urgent priority for all nations. As part of a major programme to develop and demonstrate the importance of AI and data science for government and planning systems, the Alan Turing Institute has established a research programme in “Shocks and Resilience” which seeks to exploit the power of behavioural modelling and spatial analysis to increase the robustness of social systems (Alan Turing Institute 2021). With more people being forced to stay at home, their energy use (electricity and gas) increased as well as the social, economic, and environmental consequences. Based on this context, in times where societies can face multiple crises, there is also a need for transboundary approaches, where threats like climate change and the consequences of the current pandemic outline numerous challenges. One of those is inequality, the unfair distribution of resources relative to need (Zenk SN 2006). We focus our effort on a particular type of unequal distribution of resources. This paper describes the research undertaken to understand better the potential inequality in urban energy use in the UK and how social determinants like income, ethnicity and other urban conditions might drive different levels of energy use that later define spatial and behavioural strategies for fair distribution of energy efficiency programs and social programs of urban well-being.

1. Introduction:

With health care systems under severe pressure, in particular during the first and second wave of the crisis, and escalating economic impacts, there has been evidence of increasing poverty levels and rising inequality especially in urban places (Whitehead M 2021). With all kinds of inequalities, the current crisis has provided fuel to expose the weaknesses in many social policies. Higher-paid workers were able to work from home while lower-paid workers most likely did not have that option. Poorer households have been heavily affected economically, lack of green access in urban areas, fewer than 1% of children from poor households reported they have access to the internet for distance learning in

a survey of 37 countries (Save the children, 2020). The economic sectors that have suspended their activities are mostly represented by low paid workers; richer countries were able to provide financial support to those seriously affected but most of the poorer countries did not have that capacity. As people were forced to stay at home and attempt to run their activities (e.g., educate children, work from home), their energy demand raised to a higher level, with the associated cost adding another struggle to the burden of lockdowns.

The global report authored by the International Resource Panel Cities highlights the important issue of inequitable resource distribution especially in urban areas (IRP, 2018). Cities like New York or Boston have started to find methods and metrics to provide a data driven approach to shape new investments in energy efficiency programs by analysing the inequality distributions and then the spatial aspects of social equity to inform more equitable distribution of their investments (City of Boston 2019, NYC Mayor's Office of Sustainability, 2016). To date, we are not aware of any study showing how the energy demand and the consequent costs have impacted those more susceptible in the UK, even prior to the pandemic.

This research attempts an improved understanding of spatial distributions in energy use, and an examination of the relationships with level of income and ethnicity. Initially energy consumption will be integrated with building characteristics, social demographics, and energy performance certificates data at the finest scale possible. The spatial disparities in energy use intensity will then be calculated, and the compute the Gini coefficient (Jang S, et al. 2017) to measure the inequality distribution of the energy use. Finally, computation of further statistical analysis will identify any spatial correlation between the most vulnerable population by income, ethnicity and energy consumption.

2. Method:

This research examines social inequality in urban energy use and associated equity implications of the spatial distributions of the energy (electricity and gas) consumption in the UK.

There are four stages taken in the studying of the social inequality of urban energy use in the UK. As first stage we have the data gathering process where we integrated and mapped the spatial distribution of energy consumption from the Department for Business, Energy, and Industrial Strategy (BEIS), the domestic energy performance of building certificates (EPCs) from the Ministry of Housing, Communities, and Local Government (MHCLG). Datasets including population, age, ethnicity, gender income, and household characteristics were based on the Office of National Statistics (ONS), from the 2011 census. The second stage is related to the calculation of Energy Use Disparities (EUD), by income

and ethnicity where we investigate the impact of ethnicity structure across different income strata. During the third stage we compute the inequality measures across the scale of data aggregation by applying the Gini coefficient, which can be used to quantify inequality. The Gini coefficient assesses the distribution of resources (mostly applied in economic analyses such as income inequality, or in this case energy use) across the population. As a final stage, additional statistical analysis is required to determine any statistically significant clustering of energy use disparities and how the energy consumptions vary in different income groups and ethnicity population distributions.

3. Expected results:

This abstract illustrates the proposed method to study the inequality in the urban energy use to quantify the social inequality. The data integration process is already in progress. The purpose of this abstract to shape the foundations of the defining metrics to measure social inequality in urban energy use which will form the basis for mapping the spatial distribution of such inequality metrics. There are four expected results to come from this research. 1. Integrated dataset. We could not find any literature that considers the spatial distribution of energy use to analyse potential inequalities by income or ethnicity in the UK. A comprehensive and integrated dataset would provide room to other researchers to analyse the energy use at the scale of the data. 2. Energy Use Disparities (EUD) by income and ethnicity considering several metrics including energy use per household, by income and by ethnicity and per capita. Domestic energy readings held securely by the Consumer Data Research Centre (CDRC) (CDRC 2015) will be accessed for this purpose. This dataset will provide energy consumption at fine scale and therefore permit computation of the EUD per capita. 3. Spatial inequality coefficients. We will provide the Gini coefficient values at the scale of the integrated data. And 4. Additional statistical support to analyse the spatial correlation between the spatial inequality's values by income and ethnicity.

Acknowledgments

This research is conducted at the Alan Turing Institute with funding by Wave 1 of The UKRI Strategic Priorities Fund under EPSRC Grant EP/W006022/1.

References

Alan Turing Institute (2021). <https://www.turing.ac.uk/research/research-projects/shocks-and-resilience>. Accessed 14 September 2021.

(City of Boston 2019). City of Boston, "Climate Action Plan 2019 Update". https://www.boston.gov/sites/default/files/embed/file/2019-0/city_of_boston_2019_climate_action_plan_update_4.pdf. Accessed 14 September 2021.

The Consumer Data Research Centre. Domestic Energy Provider - Aggregated data for 2015. <https://data.cdrc.ac.uk/dataset/domestic-energy-provider-aggregated-data>. Accessed 14 September 2021.

IRP (2018). The Weight of Cities: Resource Requirements of Future Urbanization. Swilling, M., Hajer, M., Baynes, T., Bergesen, J., Labbé, F., Musango, J.K., Ramaswami, A., Robinson, B., Salat, S., Suh, S., Currie, P., Fang, A., Hanson, A. Kruit, K., Reiner, M., Smit, S., Tabory, S. A Report by the International Resource Panel. United Nations Environment Programme, Nairobi, Kenya. Accessed 14 September 2021.

(NYC Mayor's Office of Sustainability, 2016). Office of Sustainability, "One New York: The plan for a strong and just city". <http://www.nyc.gov/html/onenyc/downloads/pdf/publications/OneNYC.pdf>. Accessed 14 September 2021.

Save the children (2020). Protect a Generation: The impact of COVID-19 on children's lives. <https://resourcecentre.savethechildren.net/library/hidden-impact-covid-19-children-global-research-series>. Accessed 14 September 2021.

Jang S, An Y, Yi C, et al. (2017) Assessing the spatial equity of Seoul's public transportation using the Gini coefficient based on its accessibility. *International Journal of Urban Sciences* 2017;21(1):91-107. doi: 10.1080/12265934.2016.1235487. Accessed 14 September 2021.

Whitehead, M., Taylor-Robinson, D., & Barr, B. (2021). Poverty, health, and covid-19. In *The BMJ* (Vol. 372). British Medical Journal Publishing Group. <https://doi.org/10.1136/bmj.n376>. Accessed 14 September 2021.

World Economic Forum (2020). COVID-19 is increasing multiple kinds of inequality. Here's what we can do about it. <https://www.weforum.org/agenda/2020/10/covid-19-is-increasing-multiple-kinds-of-inequality-here-s-what-we-can-do-about-it/>. Accessed 14 September 2021

Zenk SN, Tarlov E, Sun J. Spatial equity in facilities providing low-or no-fee screening mammography in Chicago neighborhoods. *Journal of Urban Health* 2006;83(2):195- 210. Accessed 14 September 2021.

Integrating stated preferences, landscape metrics and attitudes to assess visual impact of wind turbines.

Marion NATIVEL¹; Gilles VUIDEL¹; Jean-Christophe FOLTETE¹; Yohan SAHRAOUI^{1,2};
Samy YOUSOUFI¹

¹UMR 6049 ThéMA, CNRS

32 Rue Mégevand, 25000 Besançon, France

+33 7 61 54 48 34, marion.nativel@univ-fcomte.fr

²UMR 8586 Prodig, CNRS, IRD

12 Pl. du Panthéon, 75231 Paris

Keywords: wind turbines, visual impact assessment, landscape perception, landscape visibility metrics, attitudes.

In order to reduce the emission of greenhouse gasses, energy goals across countries have been focused on developing energy projects with a low carbon footprint. Wind turbine facilities may be one of the renewable energies technologically and economically mature chosen by politics to tackle these energy transition challenges. In France, the deployment of wind turbine facilities is promoted by the government through financial incentives that benefit both companies developing projects and local territories where these projects are taking place. In parallel, wind turbine projects generate local communities' resistance reclaimed by associations who bring the landscape visual impacts among other disturbance. Hence, characterising the visual impact of wind turbines can meet some issues like fostering social acceptance (Warren et al., 2005, Wolsink, 2007) or supporting decision making process for planning further implantations (Rodrigues et al., 2009, Lothian, 2008).

In this context, the aim of this study is to assess the people's landscape preferences associated with landscapes concerned by wind turbine deployments. These preferences depending on both the characteristics of the landscapes and the observers (Kaltenborn et al., 2002, Molnarova et al., 2012), we seek to highlight to what extent landscape preferences depend on respondents' attitude toward wind turbine facilities (and energy transition as a whole) or landscape (landscape configurations, differentiated impacts of wind turbines, etc.)?

We adopted here a quantitative approach to tackle this issue. 32 photographs showing various real wind turbine facilities configurations in Burgundy-Franche-Comté landscapes (France) have been taken. Landscape visibility GIS metrics characterising both landscape configuration and the presence of wind turbines (their number and visual pervasiveness depending on their distance) have been computed with the PixScape software (Sahraoui et al., 2018) for each landscape scene. Those photographs constituted the visual stimuli of an online photo-based survey completed by 554

participants among France. Respondents' preferences have been assessed using pairwise comparisons (Courcoux and Semenou, 1997). They were asked to select their favourite landscape scene between two photographs pick-up randomly in the set of photos. On a second time, respondents were asked about their experience and general attitude toward wind turbine facilities.

The material collected consists of three data tables containing (1) the participants' preferences for each landscape scene; (2) the participants' profile according their experience and general attitude toward wind turbine facilities; and (3) the landscape scenes features with visibility metrics describing the landscape configuration, its composition and the configuration of the wind turbine parc in the landscape scene. The co-inertia structure of this data set has been analysed using the RLQ-method (Dodédec et al., 1996). The RLQ analysis, is commonly used in ecological studies (Sitzia et al., 2017) and has however still been very little mobilized in geography. The results of this study will be presented and discussed in this colloquium.

Acknowledgments

We'd like to thank P. Savary for his backing in the RLQ-method application and S. Dray for his help in the interpretation of results.

References

- Courcoux, P., & Semenou, M. (1997). Preference data analysis using a paired comparison mode. *Food Quality and Performance*, 8(5), 353–358.
- Dodédec, S., Chessel, D., ter Braak, C.J.F., Champely, S., (1996). Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* 3, 143–166. <https://doi.org/10.1007/BF02427859>.
- Kaltenborn, B. P., Bjerke, T., (2002). Associations between environmental value orientations and landscape preferences. *Landscape and Urban Planning*, 59(1), 1-11. [https://doi.org/10.1016/S0169-2046\(01\)00243-2](https://doi.org/10.1016/S0169-2046(01)00243-2).
- Lothian, A., (2008). Scenic Perceptions of the Visual Effects of Wind Farms on South Australian Landscapes. *Geographical Research*, 46, 196–207. <https://doi.org/10.1111/j.1745-5871.2008.00510.x>.
- Molnarova, K., Sklenicka, P., Stiborek, J., Svobodova, K., Salek, M., Brabec, E., (2012). Visual preferences for wind turbines: Location, numbers and respondent characteristics. *Applied Energy*, 92, 269–278. <https://doi.org/10.1016/j.apenergy.2011.11.001>.

- Rodrigues, M., Montañés, C., Fueyo, N., (2010). A method for the assessment of the visual impact caused by the large-scale deployment of renewable-energy facilities. *Environmental Impact Assessment Review*, 30, 240–246. <https://doi.org/10.1016/j.eiar.2009.10.004>.
- Sahraoui, Y., Vuidel, G., Joly, D., Foltête., J.C., (2018). Integrated GIS software for computing landscape visibility metrics. *Transactions in GIS*, 22(5), 1310-1323. <https://doi.org/10.1111/tgis.12457>.
- Sitzia, T., Dainese, M., Krüsi, Bo., McCollin, D., (2017). Landscape metrics as functional traits in plants: perspectives from a glacier foreland. *PeerJ* 5, e3552–e3552. <https://doi.org/10.7717/peerj.3552>
- Warren, C.R., Lumsden, C., O'Dowd, S., Birnie, R. V., (2005). ‘Green On Green’: Public perceptions of wind power in Scotland and Ireland. *Journal of Environmental Planning and Management*, 48(6), 853-875. <https://doi.org/10.1080/09640560500294376>.
- Wolsink, M., (2007). Wind power implementation: The nature of public attitudes: Equity and fairness instead of ‘backyard motives’. *Renewable and Sustainable Energy Reviews*, 11, 1188–1207. <https://doi.org/10.1016/j.rser.2005.10.005>.

Wood trade regions of Russia

**Nikita SINITSYN¹; Mikhail ELYUTIN²; Andrei KOROTKOV³; Artem KUSHLEVICH⁴;
Roman MENYALSCHIKOV⁵; Dmitry RUSAKOV⁶;**

¹Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 925 704 7066, nicksinus@yandex.ru (correspondent author)

²Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 962 958 3294, mikhail.elyutin@student.msu.ru

³Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 910 474 7213, korotkov.andrey@geogr.msu.ru

⁴Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 919 724 2683, artem.kushlevich@yandex.ru

⁵Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 961 026 9060, roman.menyalschikov@yandex.ru

⁶Lomonosov Moscow State University
1, Leninskie Gory, Moscow, 119234, Russia
+7 999 280 1536, d_c_rusakov@mail.ru

Keywords: Russia, wood industry, trade, regions, graph clustering

In 2015 Unified State Automated Information System (USAIS from Russ. EGAIS) ‘Accounting Timber and Transactions with It’ (2015) was launched. Legal entities and sole traders have to provide information about their wood trade (timber and lumber) transactions. Significant part of data is open: transaction id, ITN or TIN and organization name for both buyers and sellers, date, deal volume in cubic meters.

Data analysis was performed using R programming language.

We used information about deals with wood between Russian companies in 2020. For each entity (by ITN/TIN) registration address and primary occupation by NACE were extracted from Federal Tax Service of Russia. Addresses were geocoded using OSM Nominatim.

Table with deals was transformed into a graph with organizations as vertices and transactions as edges. The edge weight is the volume of wood deal.

Then the graph was clustered using Leiden algorithm (Traag, Waltman and van Eck, 2006). This algorithm, as well as Louvain algorithm, has a problem of resolution limit. These algorithms tend to

combine some medium and small clusters into large. We solved this issues in the following way. Every Leiden-created cluster was drawn in abstract space using Fruchterman-Reingold layout (Fruchterman and Reingold, 1991, pp. 1129-1164). Then we extracted clusters from layout using meanshift algorithm (Fukunaga and Hostetler, 1975, pp. 32-40). Such approach helped us to divide many Leiden-created clusters into component parts. Also, a special method was developed to restrict uncertainty of clustering – both Leiden algorithm, and Fruchterman-Reingold layout are non-deterministic algorithms.

Wood-trade clusters have different size and content. They are rather autonomous. The average share of intra-cluster trade is 93 %. The modularity score of our partition is 0,863.

Some of them look like classical Porters's clusters (1990): many companies are linked in the technological chains. There is a competition in such clusters – considerable part of companies inside them has the same specialization. Some other clusters look like Kolosovskiy's spatial-industrial complexes (SIC) (2006). SIC clusters usually consist of several large manufacturers with unique specialization inside cluster and their small multiple "satellites". It's important to stress that the main participants of SIC clusters don't compete, they could only cooperate with each other within production chains.

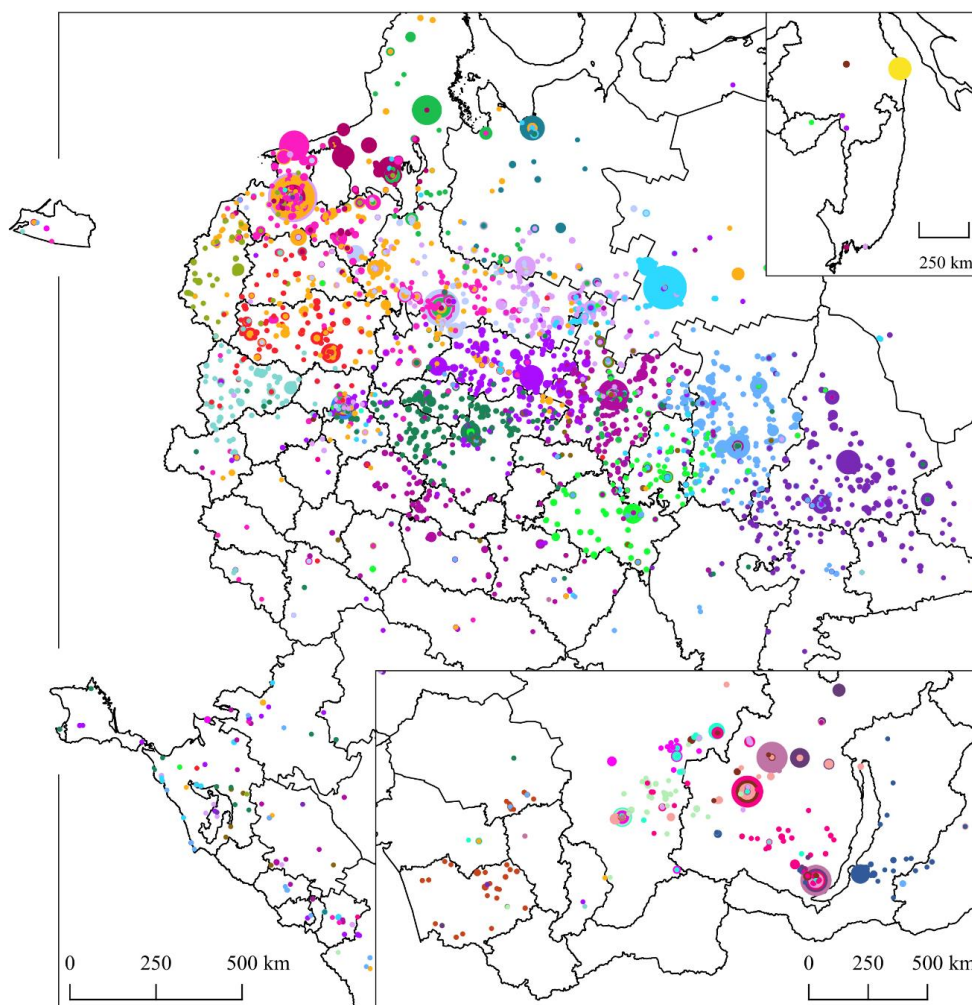
We defined 5 morphological types of clusters. Vertical monocentric consist of one large enterprise and many medium-sized and small firms. Vertical polycentric consist of several (usually 2 or 3) large enterprises and many medium-sized and small firms. All links in vertical clusters are between small or medium-sized firms and large enterprise. Small and medium-sized firms have no or very few links between each other. Vertical clusters often look like SICs. Horizontal clusters are well-connected groups of enterprises, which have different sizes. Its density of links usually is very high. Sometimes they contain several subclusters. Horizontal clusters often look like Porter's clusters. Dendritic clusters are linear chains of enterprises (or little groups of enterprises), which relate to each other sequentially. They are widespread in Siberia. Most of such clusters are situated on the underdeveloped territories. They may be kernels or consistent parts of a future vertical or horizontal clusters. Simple clusters are very small groups of enterprises. The layout of its graph looks like a little chain, triangle, square or something like that. Their nature is very different – little local group (several little firms from neighbor towns), a pair of huge plants somewhere in the middle of Siberian periphery, special wood-trade company of Russian railroads, etc.

There are three factors of clustering in wood industry of Russia. The fist is production chains – direct sale and purchase of raw materials by wood industry enterprises without traders. There are 4 types of production chains – pulp, plywood, lumber and chipboard chains. The second is demand chains - sale and purchase of raw materials by wood industry enterprises via traders. There are 3 types of demand

chains – redistribution of raw material between manufactures within the cluster, accumulation of wood volumes for some enterprises outside the cluster, wood trading between traders within the cluster (perhaps, it's a kind speculation). The third is common holder. There are a plenty of clusters with groups of enterprises, which belongs to the same holders or group of holders.

The last step of our study was to plot clusters on the map – most clusters create rather compact areas. So, such spatial projections of wood trade clusters we called wood trade regions. Its borders often overlap with existing administrative boundaries. Sometimes its areas occupy a part of some region or several regions simultaneously (fig. 1).

Turnover and affiliation of settlements to russian wood trade clusters



Wood trade turnover
of settlement, tsd cbm

Every cluster has it's own colour.
Colours were selected by random order.

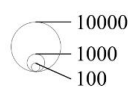


Fig. 1. Wood trade regions of Russia.

So, the proposed approach of economic sector segmentation, based on the graph clustering, gave some interesting result. The structure of Russian wood trade was described. Most clusters have clear interpretation. We defined factors of clustering of Russian wood industry. Also, we defined morphological types of clusters. Maps of clusters shows that its enterprises tend to concentrate in some compact areas. The borders of the biggest wood trade regions are rather close to the borders of the official regions of Russia.

References

- Kolosovsky, N.N. (2006). *Izbrannye Trudy*. Edited by N.N. Kazanskiy et al. Smolensk: Oykumena.
- (2015) *Accounting timber and transactions with it*. Available at: <https://lesegais.ru/open-area/deal> (Accessed: 01 March 2021).
- Fruchterman, T.M.J. and Reingold, E.M. (1991). 'Graph drawing by force-directed placement', *Software: Practice and Experience*, 21, 1129-1164.
- Fukunaga, K. and Hostetler, L. (1975). 'The estimation of the gradient of a density function, with applications in pattern recognition', *IEEE Transactions on Information Theory*, 21/1, 32-40.
- Porter, M. (1990). *The Competitive Advantage of Nations*, Free Press: New York.
- Traag, V.A., Waltman, L. and van Eck, N.J. (2019). 'From Louvain to Leiden: guaranteeing well-connected communities', *Scientific Reports*, 9, 5233.

The Luxembourgish Space Industry – Evolving Between Local and Global Forces

Denisa NAIDIN¹; Sabine DÖRRY²

¹University of Luxembourg and Luxembourg Institute for Socio-Economic Research

11 Porte des Sciences, L-4366 Esch/Alzette, Luxembourg

denisa.naidin@liser.lu (correspondent author)

²Luxembourg Institute for Socio-Economic Research

11 Porte des Sciences, L-4366 Esch/Alzette, Luxembourg

sabine.doerry@liser.lu

Keywords: Innovation Networks, Sectoral Dynamics, Clusters

The small country Luxembourg has emerged as a surprising new hub for the space industry that aims to challenge established European space nations. We employ large-scale web-sourced information on industry actors and joint market-oriented R&D from key space funders to analyse this curious case by linking agglomeration and sectoral dynamics across multiple geographical scales. Empirical data on space industry actors in Luxembourg shows that the sector has moved from being heavily reliant on large European aerospace organisations towards a more localised base. The organisations that can benefit from the connections within and outside of Luxembourg to develop innovations tend to be older and located in one of three geographical clusters. Recently incorporated companies tend to be poorly connected, which raises questions on the sustainability of these developments. Sectoral connections to globally relevant players outweigh agglomeration dynamics. Data suggest this is partly due to the thin and inexperienced system of connectors that would normally give newly established businesses a longer networking reach. The causes for the innovation network's shift towards a more localised base cannot be fully explained by sectoral or cluster transformations, raising the need for further research.

1. Introduction

Over the past 20 years, Luxembourg has emerged as a surprising new hub for the space industry. It did so in spite of its small size and limited industrial base in this sector. What has enabled this change? Have global sectoral forces driven this evolution or were local agglomeration dynamics and national transformations behind it? This exploratory research mobilises concepts from innovation networks, localization economies and evolutionary economics to explain the behaviour of the space industry in Luxembourg. It aims to explore the tension between local context and the global systems in which organisations evolve by considering innovation networks both territorially and sectorally. It is a perspective well adapted to the space industry, as it combines both established MNEs heavily concentrated in a handful of clusters and New Space born-global businesses that rely on heterogeneous collaboration networks.

2. Literature Review

Sustained economic growth has long been associated with innovation (Malmberg and Maskell, 2002) resulting from collaborations (Singh, 2005). Scholars in industrial development have employed two lenses to analyse it: the territorial view, considering place-specific factors that drive path development (Tödling and Trippel, 2018; Coenen et al, 2015), and the sectoral view, focused on how relations between organisations within the same industry, which often transcend geographical boundaries, shape the evolution of that industry (Rohe, 2020; Binz and Truffer, 2017). More and more authors are underlining the need to combine both in order to get a more accurate picture of how industries transform within and beyond space (Binz et al, 2020; Boschma et al, 2017). The study of the structure and density of innovation systems can bridge geography and sector, help predict the direction an industry might grow in and highlight system elements that can support sustainable economies.

3. Methodology & case study overview

Without a clear-cut definition of the space sector and no database of joint innovations, we use large-scale web-sourced information (see Kinne and Axenbeck, 2020) on industry actors and joint market-oriented R&D from key space funders¹ to build temporal innovation networks for the space sector in Luxembourg. Focusing on their structural properties (as in Broekel and Boschma, 2011) and evolution, we aim to show key organisations and temporal system transformations. The sectoral networks are subsequently connected to spatial effects through members' clustering behaviour and connectivity within/between clusters, linking agglomeration spill-overs and place-based effects to sectoral change. Luxembourg's space industry was traditionally organised around a large government-supported player (satellite-operator SES). Its structure has evolved with the recent addition of a significant number of new companies (see figure 1)². This follows generic trends in the sector worldwide, but also coincides with new national policies³ and the development of a critical mass in the local innovation system.

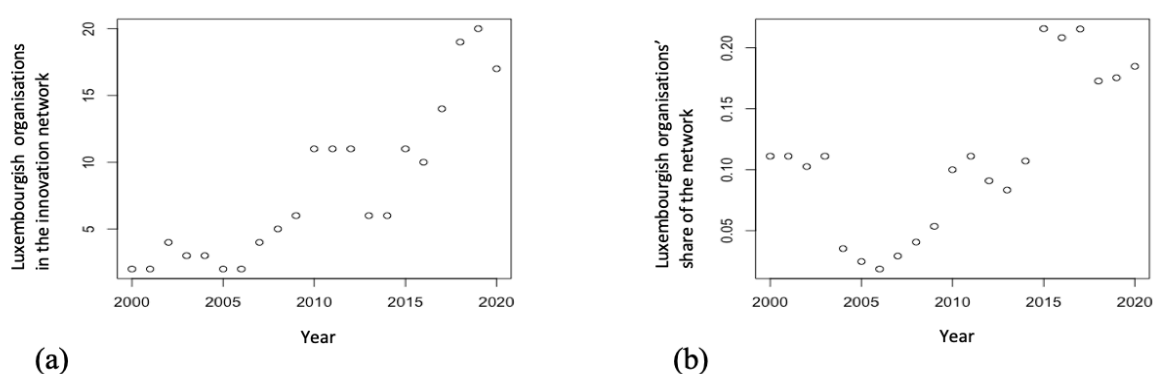


Fig 1: Luxembourgish organisations' connectivity in (a) absolute terms and (b) as a proportion of all connected organisations

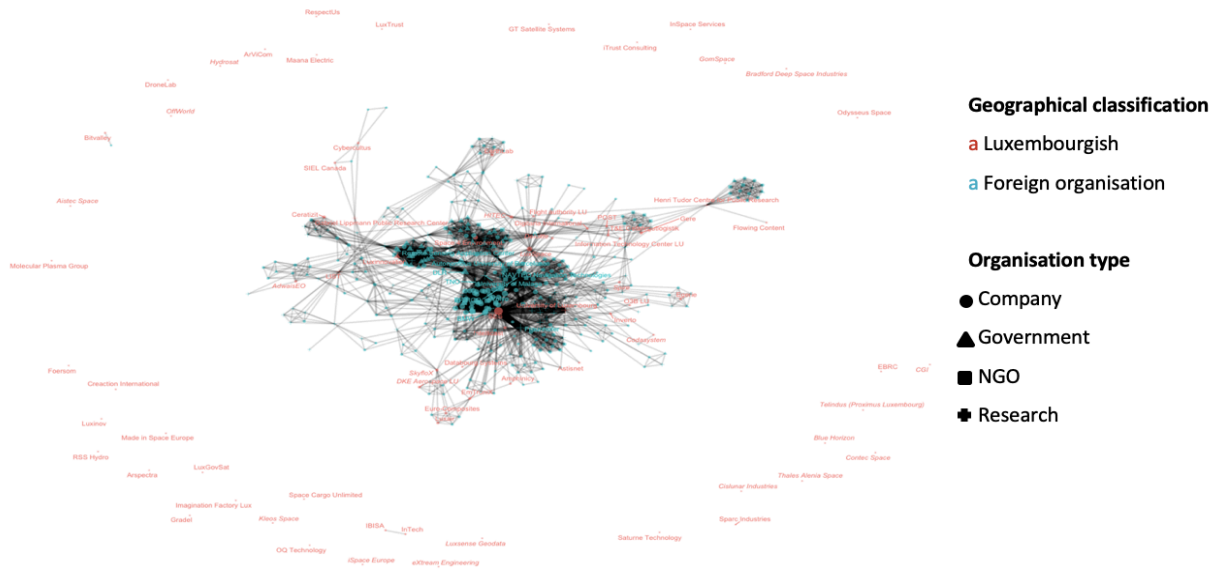
¹ We employ data from the Luxembourg Space Agency (LSA), European Space Agency (ESA) and the European Union (EU). The last two are the largest funders of space innovation in Europe, as underlined in the European Investment Bank's Future of European Space report. We focus on organisations funded by space-dedicated calls to avoid the challenge of there not being a clear-cut definition of what constitutes a space company (see the OECD Handbook of Measuring the Space Economy). We only consider joint projects that involve at least one partner in Luxembourg (our geographical anchor) and at least one commercial entity, since we are interested in market-oriented innovation. The LSA does not offer funding but puts together a yearly directory of space Luxembourgish companies, augmenting the list of organisations to consider.

² In the last 5 years, Luxembourg has seen the incorporation of as many companies as in the 60-year period before its accession to the European Space Agency (in 2005). 40% of new incorporations in the sector in the period 2015-2020 represents local branches of global space actors.

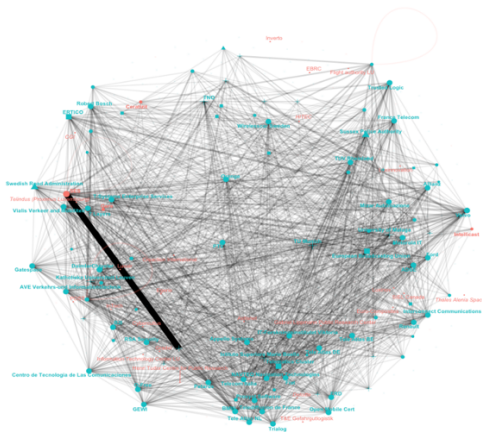
³ Beyond the near-unique legislation on space resources that recognises outer-planet property rights for all Luxembourg-based firms, this includes MoUs with other space nations, the establishment of a dedicated research centre, support programmes for space businesses etc.

4. Empirics

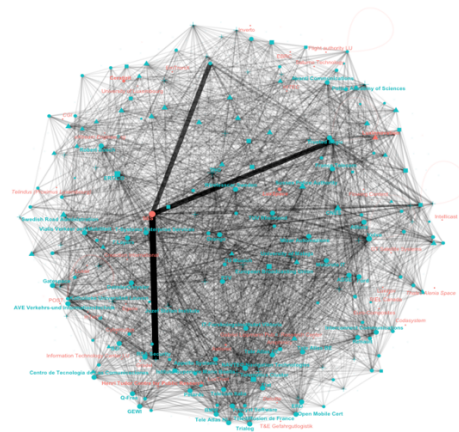
Luxembourg’s space innovation network remains thin, centred on the legacy actor SES and heavily reliant on foreign knowledge and networking (figure2). However, it evolved towards higher density and a stronger local base following two connectivity spikes: involvement in large-scale EU space networks (late 2000s) and increased local dynamism (2015-2020).



(a) 2000 - 2020



(b) 2000 - 2004



(c) 2005 - 2009

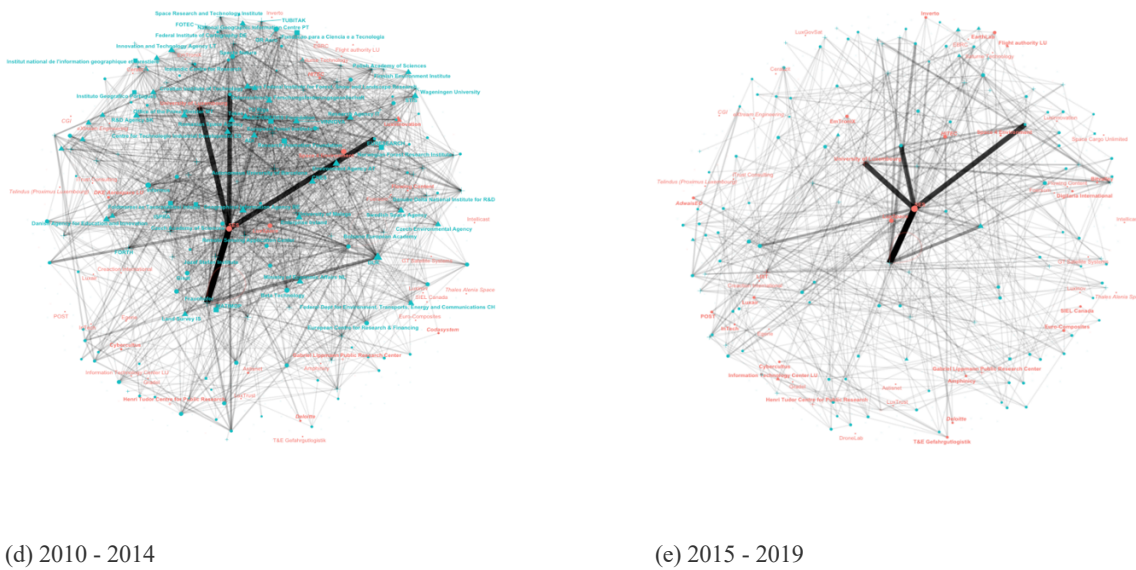
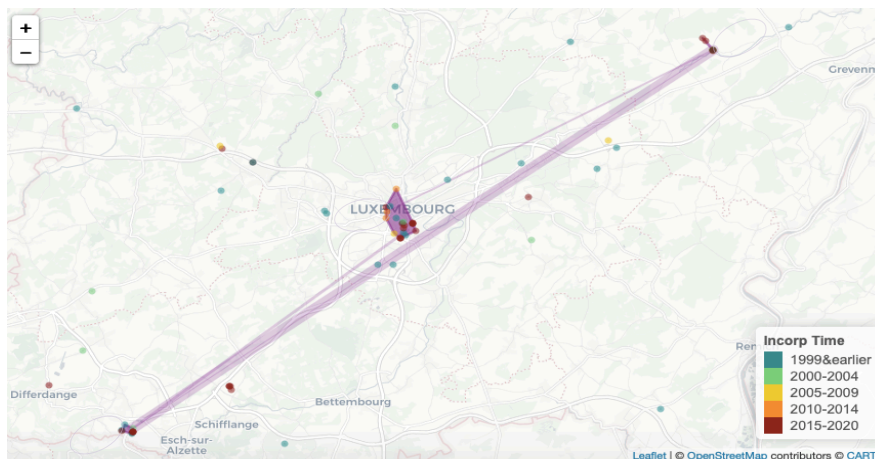


Fig 2: Temporal sectoral innovation networks for the space industry in Luxembourg

Links represent joint participation in space CR&D projects.

The historic space actor SES remains heavily anchored into the international environment, collaborating with large global partners in established space nations. Within the national innovation network, it has reinforced its centrality, in turn reflected geographically in the establishment of a cluster around its main office. Conversely, 40% of newly incorporated companies did not connect into either the local or the wider European sectoral innovation system, irrespective of whether they were subsidiaries of global groups or home-grown start-ups. The geographical clusters that contain these organisations present a mix of older and newer members, consistently connecting outwards rather than inwards (map1).



Map 1: Collaborations between/within the 3 space clusters

5. Discussion

Sectoral network structures highlight the self-reinforcing central role of large established companies and the national innovation system’s dependence on foreign organisations that act as connectors. Legacy players capitalise on connections to domestic and global industry leaders to build a track-record of innovation, while newly-incorporated companies rely on limited international collaborations⁴. Their lack of local connectivity could be explained by the thin support system⁵. Recently, the rise in incorporations has been accompanied by a significant expansion in the number of local companies participating in joint innovation (figure3) and a small increase in network density.

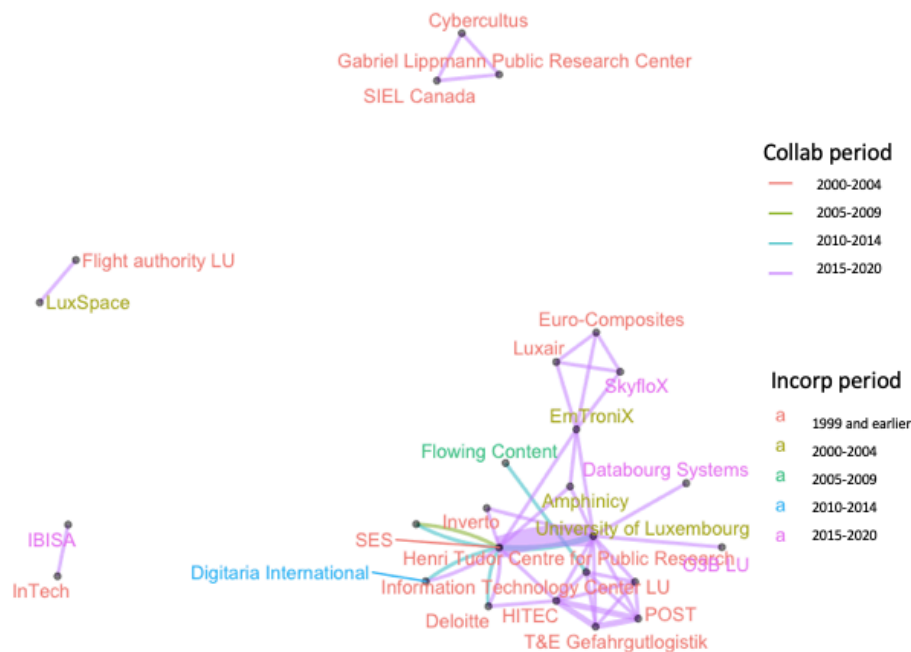


Fig 3: Timeline of collaborations between Luxembourg-based organisations

⁴ This partly confirms the characteristics of New Space companies as born-global actors. Nonetheless, the level of international connectivity among organisations established in the period 2015-2020 is overall lower than for those incorporated between 2000 and 2010.

⁵ The key foreign institutions that Luxembourgish companies connected into provided a mix of knowledge and networking activities (most of them were applied research, innovation and technology transfer organisations: Fraunhofer, DLR, TNO etc). With the establishment of Luxembourg’s own equivalents, that role has been diminished but did not completely disappear, especially for businesses with larger networks. Some of the new local institutions have networked widely across the international environment to build their own knowledge base (e.g. LIST), while others have built privileged relations with established local companies (e.g. the university predominantly works with SES). However, that connectivity gap is still there for smaller and younger companies, insufficiently well connected internationally.

The geographical structure of innovation networks has also changed: younger space actors prefer urban areas to the peri-urban settings of the established space industry (map1, conforming to global locational trends in highly complex industries, see Gomez-Lievano & Patterson-Lomba, 2019). Yet, the agglomerations that emerge do not boost CR&D⁶, raising the need for further explanation.

6. Conclusions

Our research shows emerging patterns of a relatively recent industry clustered in a small country that aims to rival top space nations worldwide. While agglomeration dynamics are present, what matters more for collaborative innovation in Luxembourg's space sector is international connectivity. Considering that almost half of the latest wave of incorporations has remained unconnected to the national and supranational innovation system, important questions should be explored to increase economic sustainability, and for which this research is an important stepping stone.

Although global players remain important in the innovation network, the national level did become a lot more prominent in the period 2015-2020. This shift is uncorrelated to the number of players joining/active in the network and cannot be explained by the establishment of new locally-active connectors. One potential explanation is the implementation of a number of national policies targeting the space industry. More research is needed to establish causality and illuminate these mechanisms.

References

Binz, C., Coenen, L., Murphy, J.T., et al. (2020) Geographies of transition—From topical concerns to theoretical engagement: A comment on the transitions research agenda. *Environmental Innovation and Societal Transitions* 34, 1-3.

Binz, C., Truffer, B. (2017) Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts. *Research Policy* 46(7), 1284-1298.

⁶ As the wide majority of connections happen between organisations in different clusters, rather than within. See map1 for further details.

Boschma, R., Coenen, L., Frenken, K., Truffer, B. (2017). Towards a theory of regional diversification: combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies*, 51(1), 31-45.

Broekel, T., Boschma, R. (2011). Aviation, Space or Aerospace? Exploring the Knowledge Networks of Two Industries in The Netherlands. *European Planning Studies*, 19:7, 1205-1227.

Coenen, L., Moodysson, J., Martin, H. (2015) Path Renewal in Old Industrial Regions: Possibilities and Limitations for Regional Innovation Policy. *Regional Studies* 49(5), 850-865.

European Investment Bank (2019) The future of the European space sector, doi: 10.2867/484965

Gomez-Lievano, A.; Patterson-Lomba, O. (2018) The drivers of urban economic complexity and their connection to urban economic performance. *Papers 1812.02842*, arXiv.org, revised Mar 2019.

Kinne, J., Axenbeck, J. (2020) Web mining for innovation ecosystem mapping: a framework and a large-scale pilot study. *Scientometrics* 125, 2011–2041.

Malmberg, A., Maskell, P. (2002) The Elusive Concept of Localization Economies: Towards a Knowledge-Based Theory of Spatial Clustering. *Environment and Planning A: Economy and Space*, 34(3), 429-449.

OECD (2012) Handbook on Measuring the Space Economy, ISBN 978-92-64-16862-6

Rohe, S. (2020) The regional facet of a global innovation system: Exploring the spatiality of resource formation in the value chain for onshore wind energy. *Environmental Innovation and Societal Transitions* 36: 331-344.

Singh, J. (2005) Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, 51(5), 756-770.

Tödting, F., Tripl, M. (2018) Regional innovation policies for new path development – beyond neo-liberal and traditional systemic views. *European Planning Studies*, 26(9), 1779-1795.

Parallel Session PSD3

4 November 4:00 am-5:40 pm (GMT)

Perception of Urban Space

Simulating pedestrian's visual perception: Dynamic 3D quantitative analyses

Yosifof Roei & Fisher-Gewirtzman Dafna

Faculty of Architecture & Town Planning, Technion, IIT

Technion city, Haifa, 3200003, Israel

+972-52-4203531, ryosifof@gmail.com

+972-54-7833331, ardafna@ar.technion.ac.il

Keywords: 3D simulation, walkability, pedestrian visual perception, visibility analysis, urban analysis

1. Introduction

Spatial assessment of the urban fabric generates insights on urban attributes and their respective influences on the urban setting and peoples' perceptions, behavior and choices. This influence may remain unclear until urban design qualities can be defined, quantified, measured, and tested (Yin, 2014). Researchers have endeavored to describe the complexity of the urban structure with analytical models and tools designed to represent this complexity by measuring wide range of data (Batty, 2013). The aspiration for quantifiable assessment methods generates opportunities for examining the performance of specific urban attributes based on empirical, data-driven methodologies.

Walkability has attracted the attention of urban planners, architects, and researchers from various disciplines in recent decades, due to the advantages to residents. In terms of effects, walkable neighborhoods have been shown to have a positive impact on physical activity and health, reduction of energy use, economic value, and social connection (Speck, 2012; Lee and Talen, 2014).

Yet walkability is an elusive concept, which is tied with numerous physical and social qualities (Forsyth, 2015). A walkable environment is influenced by the characteristics of the urban fabric and its built environment. Accordingly, the term walkability is interpreted in many different ways in the scientific literature. Forsyth illustrates various themes and attributes related to walkability - environmental conditions which are associated with physical dimensions, social outcomes which are related to the availability and attractiveness of public activity on the streets, and proxy for a better design where walkability serves as an indicator of better urban areas which tend to have high liveability.

2. Walkability assessment tools

The complexity and interdisciplinarity of the walkability concept in the urban setting pose some difficulty in measuring it. Previous research has defined five main urban physical qualities under which

a walkable environment can be evaluated: Human Scale, Enclosure, Imageability, Transparency and Visibility, and Complexity (Ewing and Handy, 2009) in addition to Connectivity and Thermal comfort. Each urban quality is composed of many different attributes and requires different assessment methods. Researchers developed different indexes, whose methodology vary greatly based on the specific field of research and desired attribute to be measured. The various attributes can be divided into different scales – micro, intermediate and macro and different research methods – qualitative or data-driven quantitative methods.

The micro-scale focuses on qualities that affect the direct experience of the pedestrian, such as human scale, streetscape features, vegetation, compactness, etc. It is often associated with public health disciplines. Walkability indexes designed to operate on the microscale are based on audits and surveys conducted by pedestrians who walk along the street. An example of a micro-scale index is the Minnesota-Irvine index which contains 160 items spanning over 20 topics that measure characteristics of streets, buildings, amenities, architectural design features, and other elements (Boarnet et al., 2006). Macro-scale indexes focus on large data sets, such as road networks, land-use, statistics, and GIS analysis. Some examples of macro-scale assessment methods are among the space-syntax tools (Hillier, 1996), and Walkscore (Duncan et al., 2011). The macro-scale is often associated with transport and urban planning disciplines.

Between the micro and macro-scales is the intermediate scale, which attempts to bridge the gap between the level of detail of the micro-scale, and the automated rapid methods of the macro-scale. The intermediate scale focuses on the physical urban qualities related to walkability, the built environment, and its features.

The current research project focuses on the development of analytical models of quantifiable urban physical attributes, which can assess the performance of urban fabrics in the fields of walkability, human scale, and simulate the pedestrian dynamic visual perception and experience in the intermediate scale.

Human scale refers to the size, texture, and articulation of physical elements that proportionally fit humans' size and walking speed, such as buildings' geometry, materials, details, vegetation, or street furniture (Ewing and Handy, 2009; Speck, 2012; Gehl, 2013).

Enclosure refers to the degree to which streets and other public spaces are visually defined by buildings, walls, trees, and other vertical elements. The enclosure rate is usually measured by a ratio between the building's height to the street width. Enclosure rate between 0.5 to 1.5 is considered a comfortable ratio. (Jacobs, 1993; Ewing and Handy, 2009).

Transparency refers to the degree to which people can perceive what lies beyond the edge of a street and the degree to which people can see or perceive humans and their potential interaction. (Ewing and Handy, 2009).

3. Introducing the Duo analysis: DVA & DESSA, simulating pedestrian perception of the urban scape.

The research proposes integrating two analytical models simulating pedestrians’ perception in the urban settings: the DVA (Dynamic Visibility Analysis) and a novel analytical model: the DESSA (Dynamic Enclosure Street Section Analysis).

The DVA (Fisher -Gewirtzman, 2017), simulates and predicts the visual perception and wellbeing of pedestrians moving along an urban path. This analytical model was assessed in an experiment conducted in an immersive virtual environment, using variant urban path morphologies. It takes into account the integrated effect of the geometry of the environment, as well as variant elements of the view such as the sky, trees and vegetation, buildings’ uses, roads, water, etc. The DVA analysis is calculating the accumulated length of lines of sight to all target surfaces, from each viewpoint, predefined along a designated path. Figure 1A is simulating lines of sight aimed at all surrounding surfaces from a pedestrian’s viewpoint. 1B is illustrating a coloured environment according to building uses. Figure 1C is illustrating the visibility calculations for each of the 100 viewpoints along the designated path.

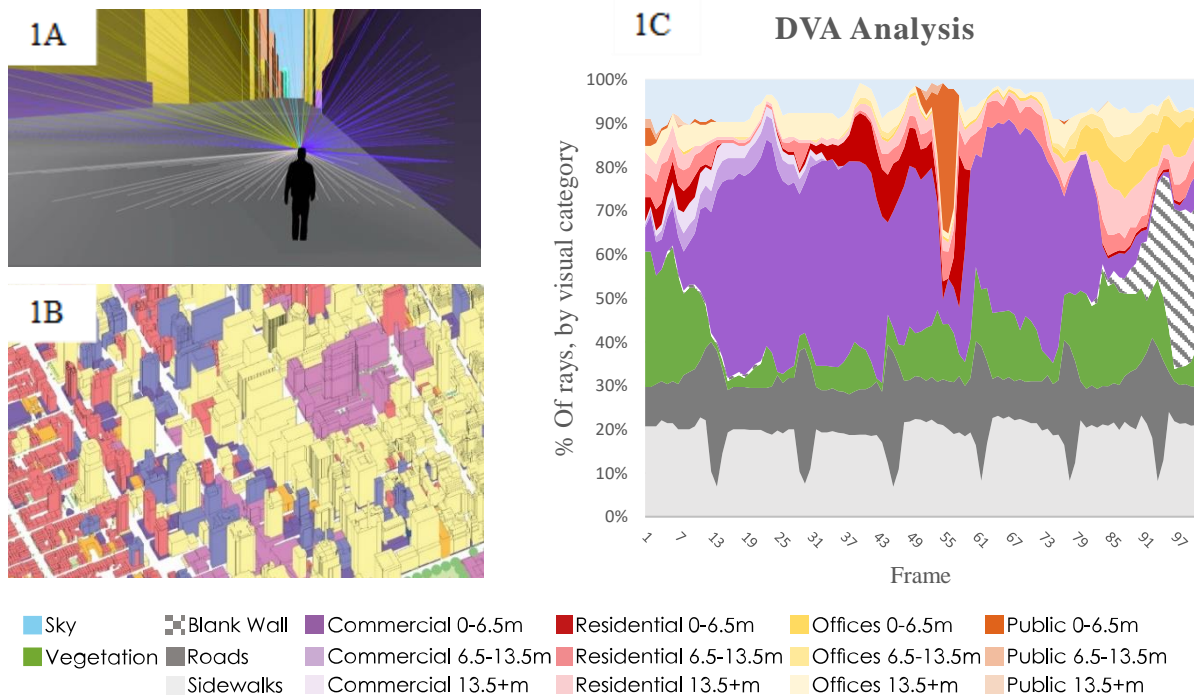


Figure 1: DVA model, simulation and results.

The DESSA (Dynamic Enclosure Street Section Analysis), is a novel analytical model, generating a street section analysis along a designated path in a selected environment. For every sequential section along the path vertical edges (such as buildings) heights are measured in addition to the distance between the building to the sidewalk, sidewalk width and road width. The collected data is used to calculate the enclosure ratio of each section side- left and right, as well as the ratio between the space allotted to the sidewalk and the road (Figure 2)

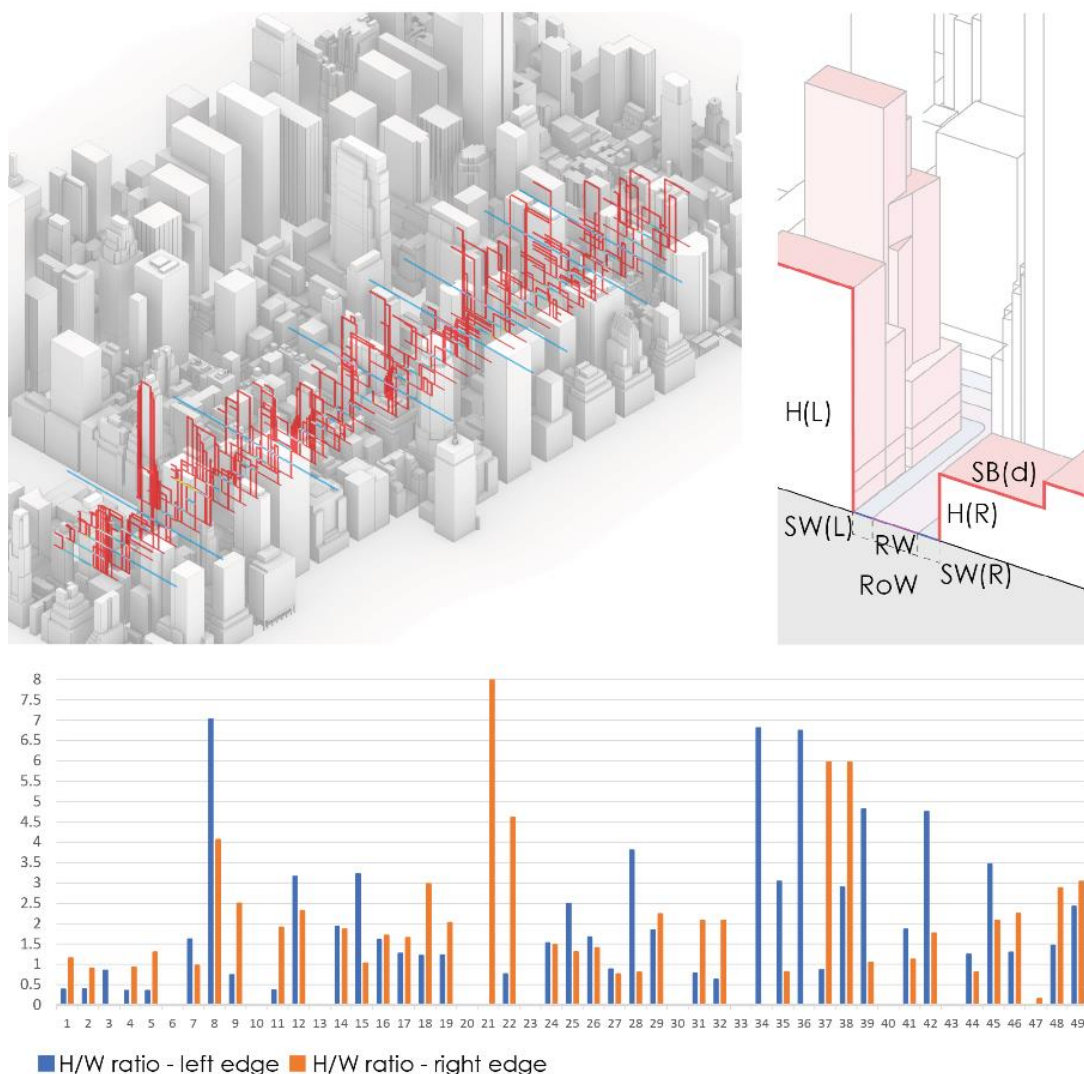


Figure 2: DESSA analysis of the enclosure ratio along 5th avenue, Manhattan, NYC. Enclosure ratio left (blue), right (orange) edges.

Both the DVA and DESSA models can operate synchronously in a 3D virtual environment. They focus on the core morphological aspects of the streetscape based on simplified virtual models, thus allowing an in-depth, quantitative exploration of the spatial relationship between the street, its defining edges, and potential visible space. Each holds an important role in the streetscape and the pedestrian's visual perception. A comparative analysis of different urban environments may shed light on better urban

settings for pedestrians and highlight the morphology, geometry, greenery, potential interaction locations and other streetscape characteristics best suited for pedestrians' wellbeing.

The combination of DVA and DESSA is an attempt to create an analytical model simulating the pedestrians' dynamic experience in a 3D environment. The DVA analyses the pedestrian's visual perception while the DESSA adds information by analysing the edges of the built environment that encloses the pedestrian's path (Figure 3). Operating in a 3D virtual environment, the tools can be used by architects and planners to evaluate design alternatives in the early design stages while making more conscious design decisions considering the pedestrian's point of view.

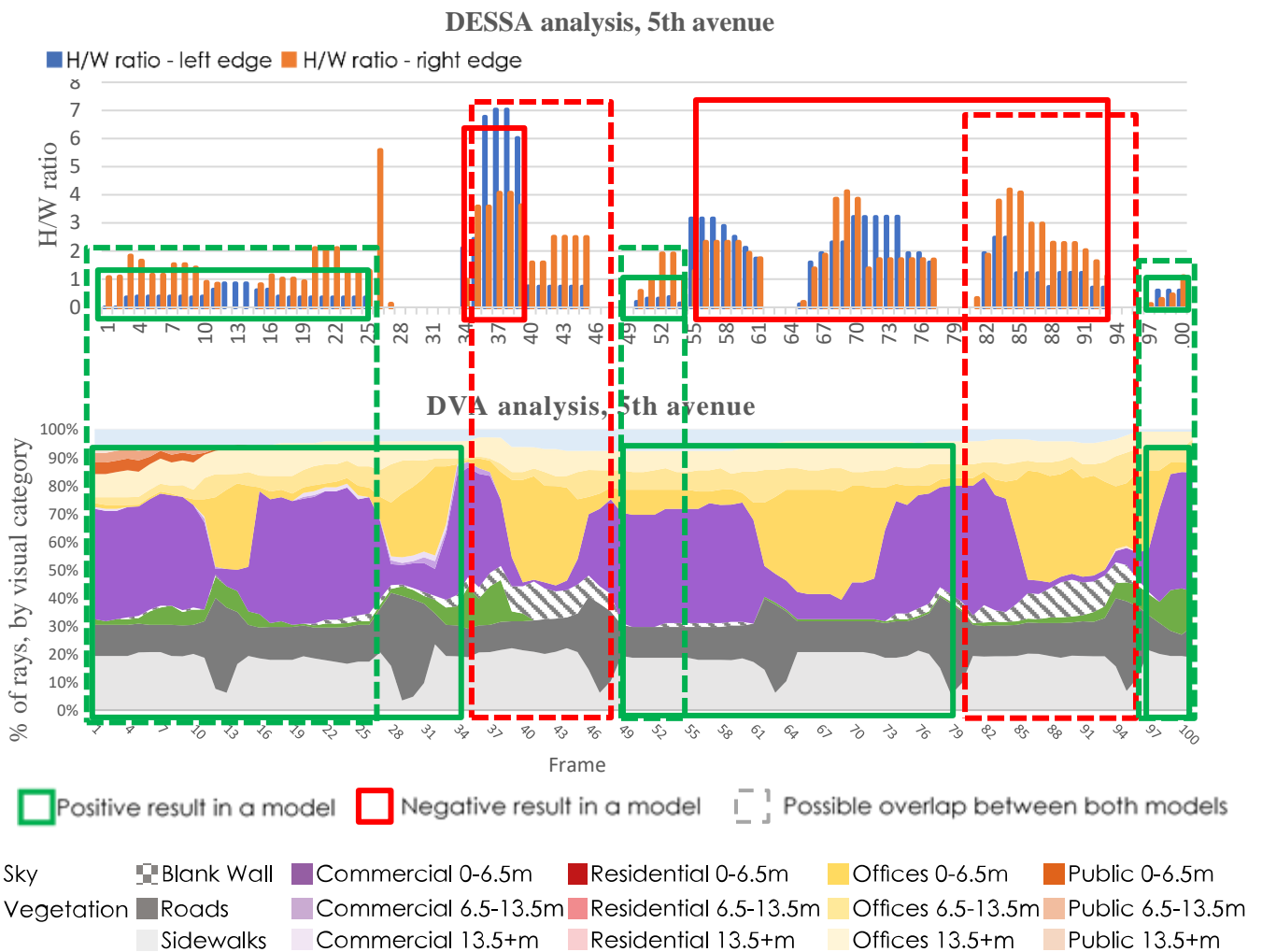


Figure 3: DVA and DESSA analyses along 5th avenue.

References

- [1] Batty, M. (2013) *The new science of cities*. Cambridge, Massachusetts: MIT Press.
- [2] Boarnet, M.G., Day, K., Alfonzo, M., Forsyth, A. and Oakes, M., (2006). The Irvine–Minnesota inventory to measure built environments: reliability tests. *American journal of preventive medicine*, 30(2), pp.153-159.
- [3] Duncan, D.T., Aldstadt, J., Whalen, J., Melly, S.J. and Gortmaker, S.L., 2011. Validation of Walk Score® for estimating neighborhood walkability: an analysis of four US metropolitan areas. *International journal of environmental research and public health*, 8(11), pp.4160-4179.
- [4] Ewing, R. and Handy, S. (2009) 'Measuring the unmeasurable: Urban design qualities related to walkability', *Journal of Urban Design*, 14(1), pp. 65–84.
- [5] Fisher-Gewirtzman, D. (2017), July. Can 3D Visibility Calculations along a Path Predict the Perceived Density of Participants Immersed in a Virtual Reality Environment? In Proceedings of the 11th Space Syntax Symposium, Lisbon, Portugal (pp. 3-7).
- [6] Gehl, J., (2013). *Cities for people*. Island press.
- [7] Hillier, B. (1996) *Space is the machine: a configurational theory of architecture*. Cambridge: Cambridge University Press.
- [8] Jacobs, A. B. (1993) *Great streets*, MIT Press.
- [9] Lee, S. and Talen, E., (2014). Measuring walkability: a note on auditing methods. *Journal of Urban Design*, 19(3), pp.368-388.
- [10] Speck, J. (2012) *Walkable city: how downtown can save America, one step at a time*, North Point Press.
- [11] Yin, L. (2014) 'Book Review: Measuring Urban Design: Metrics for Livable Places', *Journal of Planning Literature*, 29(3), pp. 273–274.

Perceived Variables in Virtual Reality: A Methodological Analysis

Ron BAR-AD¹; Geoffrey CARUSO²; Nuno PINTO³; Markel VIGO⁴

¹University of Manchester Data Centre of Doctoral Training
ron.bar-ad@postgrad.manchester.ac.uk (correspondent author)

²Luxembourg Institute of Socio-Economic Research

³University of Manchester School of Environment, Education and Development

⁴University of Manchester Department of Computer Science

Keywords: virtual reality; perceived density; method analysis; city modelling; research methods.

1. Background

1.1 Perceived Density

Although it has a measured effect on inhabitant well-being (Haifler and Fisher-Gewirtzman 2020), research examining perceived urban density from a data-driven, quantitative perspective are few and far between. Many touch on the subject (see Mavridou 2012) or adjacent subjects like walkability (see Adkins et al. 2012) or greenness (see Gilchrist et al. 2015), or assess perceived density from a qualitative stance (see Rapoport 1975). Those that do use a quantitative approach may still have small, often specialised samples (Fisher-Gewirtzman 2018) or less than conclusive results.

1.2 Virtual Reality

Virtual environments have the potential to reconcile the trade off between internal and external validity in experiment design. By creating an experience as close as possible to actually being somewhere, while also allowing close control of the dependent factors, well-made immersive virtual environments (IVEs) can give researchers the best of both worlds (see Birenboim et al. 2019). The use of IVEs and VR in research is still burgeoning, but has a lot of potential for research in human-environment interaction (Bishop 2001). Many such studies in VR have already taken place (see Sanchez et al. 2017; Shushan et al. 2016), and many more show the potential of VR as a research tool (Kuliga et al. 2015; Yuan and Zhang 2014; Orland et al. 2001; Ball et al. 2008; Silvestri et al. 2010).

2. Research Questions

In this research we aim to answer these questions:

- How can immersive virtual environments be utilised in user-centred urban environment research?

- What are the comparative benefits of using today's virtual environment technology for urban environment research in contrast to existing methods?

By answering these questions, we would be able to address the following open challenges:

- How accurate are laypeople's perceptions of urban population density?
- What factors most influence a layperson's perception of urban population density?

3. Methods

The analysis will address these questions in two ways. First, a robust and comprehensive literature review covering the state of the art in regard to (a) perceived density; (b) virtual reality; and (c) similar methodologies in research. Second, a thorough evaluation of methods and approaches for designing a study of perceived variables in virtual reality.

3.1 Literature Review

The review will try to encompass as much of the available literature as possible, drawing its conclusions from articles, chapters, and books to create a complete image of the state of the art. It will include an assessment of the measured or predicted effects and affects of perceived and actual density; an evaluation of VR for research and specifically the usefulness of IVEs; and an examination of previous research using similar methodologies.

3.2 Method Analysis

The methods analysis will cover a range of potential approaches to the problem of designing virtual urban research, including:

- (a) Software solutions for city modelling, on three dimensions:
 - a. Immersion (how real the environment feels);
 - b. Control (how much the researcher can manipulate the environment for their needs);
 - c. Availability (how easy it is to access and use the software; including purchase and subscription prices, hardware requirements, and ease of use for new users);
- (b) Research methods for gathering data, including:
 - a. Quantitative data: Likert scales, answers measured against objective measurements;
 - b. Qualitative data: unstructured interviews, eye- and motion-tracking, voice recording;
 - c. Evaluation of the feasibility and usefulness of these methods or a mix of them;
- (c) A decision-making framework for those intending to pursue the same type of research.

The decision-making framework will attempt to direct researchers to the best methods and approaches available based on their resources and intentions. For example, a researcher looking to study urban beauty in IVEs, but that has only a short amount of time and limited funding, may be directed towards a cheap, easy-to-use tool with high photorealism but low control, and recommended to use qualitative methods in small samples¹. This is an example, of course, but the decision-making framework should aim to cover a wide range of possibilities for researchers, and in doing so expose where gaps exist for software solutions or methodological approaches to fit a certain niche.

4. Timeline

The report is already underway, but may take several months to complete. At time of writing, it is expected to be at its final stages by the end of the year.

Acknowledgments

This paper is part of an ongoing collaboration between the University of Manchester (UoM) and the Luxembourg Institute of Socio-Economic Research (LISER), with funding from LISER and supervision from both LISER and UoM.

References

- Adkins, A., Dill, J., Luhr, G. and Neal, M., 2012. Unpacking walkability: Testing the influence of urban design features on perceptions of walking environment attractiveness. *Journal of Urban Design*, 17(4), pp.499-510.
- Ball, J., Capanni, N. and Watt, S., 2008. Virtual reality for mutual understanding in landscape planning. *International Journal of Social Sciences*, 2(2), pp.78-88.
- Birenboim, A., Dijst, M., Ettema, D., de Kruijf, J., de Leeuw, G. and Dogterom, N., 2019. The utilization of immersive virtual environments for the investigation of environmental preferences. *Landscape and Urban Planning*, 189, pp.129-138.
- Bishop, I.D., Wherrett, J.R. and Miller, D.R., 2001. Assessment of path choices on a country walk using a virtual environment. *Landscape and urban planning*, 52(4), pp.225-237.
- Fisher-Gewirtzman, D., 2018. Perception of density by pedestrians on urban paths: An experiment in virtual reality. *Journal of Urban Design*, 23(5), pp.674-692.

¹ This is an example of the factors the decision-making framework may use when guiding researchers, not an actual recommendation based on literature and methods analysis.

Gilchrist, K., Brown, C. and Montarzino, A., 2015. Workplace settings and wellbeing: Greenspace use and views contribute to employee wellbeing at peri-urban business sites. *Landscape and Urban Planning*, 138, pp.32-40.

Haifler, T. and Yaala and Fisher-Gewirtzman, D., 2020. Urban Wellbeing, As Influenced by Densification Rates and Building Typologies-A Virtual Reality Experiment.

Kuliga, S.F., Thrash, T., Dalton, R.C. and Hölscher, C., 2015. Virtual reality as an empirical research tool— Exploring user experience in a real building and a corresponding virtual model. *Computers, environment and urban systems*, 54, pp.363-375.

Mavridou, M., 2012. Perception of three-dimensional urban scale in an immersive virtual environment. *Environment and Planning B: Planning and Design*, 39(1), pp.33-47.

Orland, B., Budthimedhee, K. and Uusitalo, J., 2001. Considering virtual worlds as representations of landscape realities and as tools for landscape planning. *Landscape and urban planning*, 54(1-4), pp.139-148.

Rapoport, A., 1975. Toward a redefinition of density. *Environment and Behavior*, 7(2), pp.133-158.

Sanchez, G.M.E., Van Renterghem, T., Sun, K., De Coensel, B. and Botteldooren, D., 2017. Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. *Landscape and Urban Planning*, 167, pp.98-107.

Shushan, Y., Portugali, J. and Blumenfeld-Lieberthal, E., 2016. Using virtual reality environments to unveil the imageability of the city in homogenous and heterogeneous environments. *Computers, Environment and Urban Systems*, 58, pp.29-38.

Silvestri, C., Motro, R., Maurin, B. and Dresch-Langley, B., 2010. Visual spatial learning of complex object morphologies through the interaction with virtual and real-world data. *Design Studies*, 31(4), pp.363-381.

Yuan, S., Song, S. and Zhang, Y., 2014. Experimental research in urban spatial cognition by using virtual reality technology. *Athens Journal of Technology Engineering*, 1(1), pp.19-32.

On Tel Aviv's Roofs

Osher Ann Lapid¹; Efrat Blumenfeld Lieberthal¹

¹Azrieli School of Architecture

The Yolanda and David Katz Faculty of the Arts

Tel Aviv University

Ramat Aviv, 6997801, Israel

972-3-6405556, osherlapid@gmail.com

Keywords: Roofs, Fifth front, 3DGIS, Building typologies

The roof, or "fifth front," has been part of our built environment since man built his first shelter. While the basic purpose of a roof is protection from natural forces, throughout history and a variety of cultures, people have used their roofs for other purposes as well.

In recent years, we have witnessed many cities that have re-discovered the roofs and utilized them for a variety of purposes, in order to improve life in dense urban environments; Current trends in these terms are using roofs for green space and urban agriculture (Pons *et al.*, 2015, p. 441). Tel Aviv seems to have all the conditions to offer additional living space and opportunities for multi-use activities on its roofs, as most of the buildings have flat roofs that facilitate extensive activities, and the mild weather allows people to enjoy open spaces during most of the year.

However, the roofs space in Tel Aviv has not yet realized its potential (this is also true in comparison to other cities around the world) (Tsarfati and Shafran, 2018, p. 69). The rooftops in Tel Aviv have been neglected over the years, covered with tar, satellite dishes, antennae, and solar water heaters. Today, these are the typical components of the roofs in Tel Aviv (Tsarfati and Shafran, 2018, p. 62). Recently, due to the Covid-19 pandemic and the lockdowns that accompanied it, the residents of Tel Aviv got a glimpse of the benefits the city's roofs may offer.

Urban roof spaces are, in fact, huge public-private spaces¹, which become more attractive as city densities increase and land in the city becomes more expensive - but they are rarely mapped. While a few cities around the world have mapped their green roofs to understand the magnitude of the phenomenon and its contribution to urban climate (City of Melbourne, 2019), the only roof mapping that exists in Tel Aviv is that of solar roofs constructed by the city government (Tel Aviv Yafo, 2020). In my research, I use digital visualization tools to analyze the roof spaces of shared residential buildings in Tel Aviv. This mapping includes both morphological and functional components, in addition to the relationship between the two. Based on the mapping, I identify different types of urban roofs in Tel Aviv and analyze their characteristics, concerning socioeconomic, social, constitutional, and constructive issues, as well as their connections and effects.

The purpose of the study is to uncover the potential that lies within city roofs for various purposes. Furthermore, the mapping allows for a deeper understanding of this unique and unexplored space in the context of the existing situation, its sources, and its effects, and may assist planners in finding ways to realize its potential.

My research question is whether there are any recognizable patterns concerning the use of Tel Aviv roofs, and if so, how can they be characterized?

The importance of the research resides in the broader understanding it gives of Tel Aviv rooftop space, which does not exist in the city today, and may contribute to the development of planning tools with benefits to the city's citizens. Often, the roof of a residential building in the city, which is a private open space, but which is shared with the building's occupants, becomes a no man's land. A roof can be an unused space complete with solar water heaters, pipes, and systems, or it can be a more vivid place,

¹ It is a private area for the occupants of the residential building, yet it is also a public space for all occupants of the building. Additionally, there are roofs whose ownership is tied to a specific tenant and are part of his apartment. Furthermore, the roof itself is a public-private space due to the fact that it is a visible outdoor space in the urban environment.

for tenants' organizations and social gatherings, and more recreational activities. The ambiguity of the roof's ownership, often creates tension between tenants and between tenants and residents in the neighborhood, concerning the level of impact that the use of the roof has on the environment - for better or for worse. Understanding the spatial distribution of the typologies and the activities that take place on the city roofs may help planners, law enforcement, and anyone who wishes to formulate policies or guidelines that will encourage tenants to live in this no man's land in ways that benefit as many parties as possible, while coexisting in the city and under the same roof.

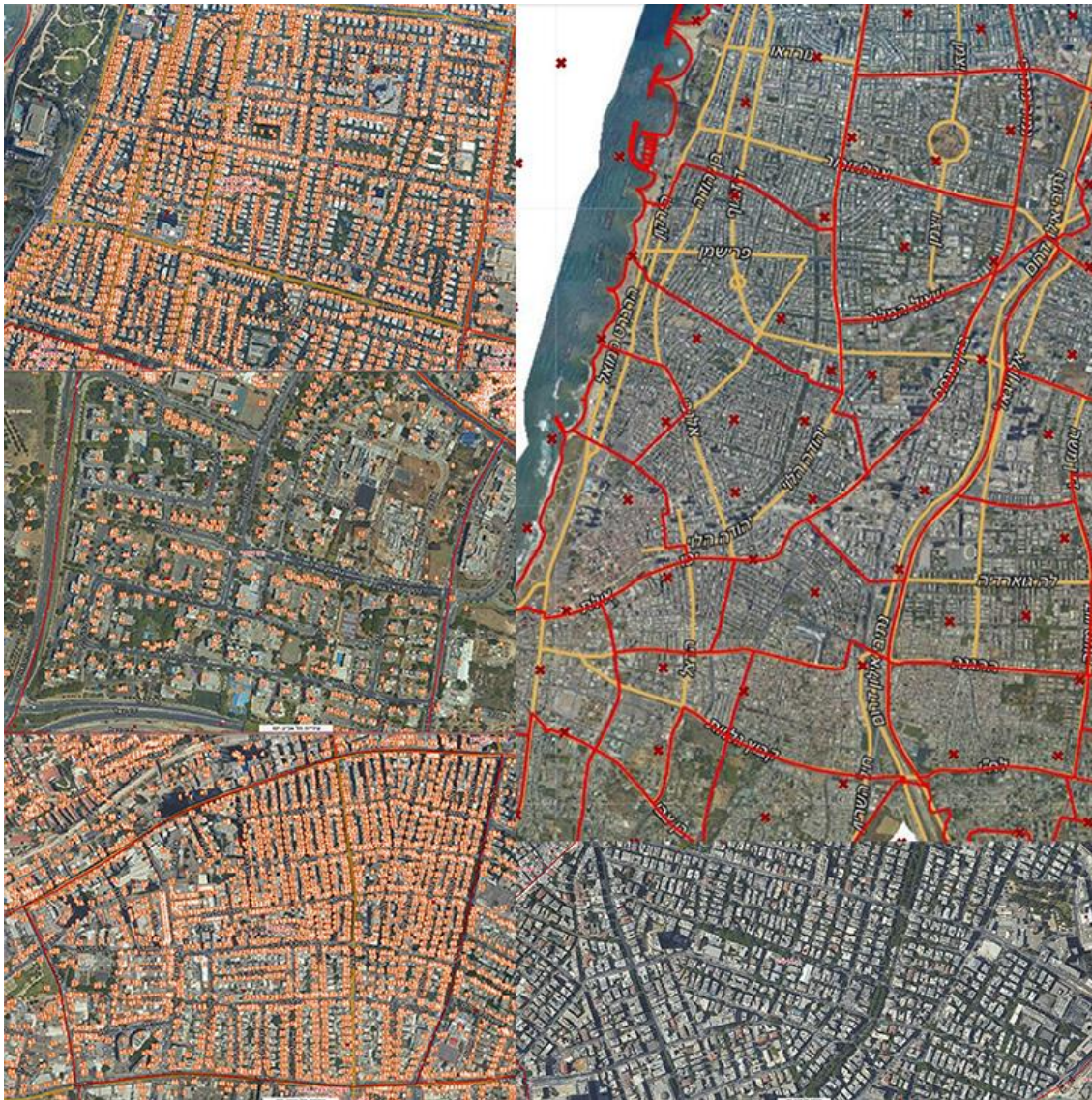
1. Methodology

This study focuses only on shared residential buildings and ignores public ones, as the use of public roofs for various purposes is already more common in the city and is less complex for regulatory implementation (since it is driven by decision-making made by the top-down forces, i.e. the municipality itself, and not by bottom-up forces that represent users). Additionally, there is no tension between the private and public on these roofs.

The mapping includes two major parameters: The morphological parameter (position in the field, the height of the structure, level of exposure to the environment, proportions, proximity to neighboring structures, and more), and The functional parameter - in what way the roof is used (or not used). The mapping is carried out using two main tools. The first is the Tel Aviv Municipality's GIS system, from which the basic data of the building is derived, as well as other data related to the morphological parameter of the mapping. The second tool is the detailed three-dimensional model of Tel Aviv, developed by *Simplex Mapping Solutions Ltd.* A key advantage of the model is its high resolution (3cm) that enables one to explore the fine details the three dimensions space. As a result of this high resolution, it is possible to utilize the model to map the functional parameters, such as activities that take place on rooftops, as well as other morphological parameters, by examining the roofs and buildings in it.

2. Results

The preliminary results are based on a sample of 80 roofs located at residential buildings throughout Tel Aviv² (Map 1). We employed descriptive statistics for the analysis such as histograms and correlation tests to determine the frequency of observed uses, and the strength of the relationship (positive or negative) between the various morphological parameters and observed uses, and then statistical significance.



Map 1: A total of 40 roofs were randomly selected from all neighborhoods south of the Yarkon River and another 40 roofs from 4 neighborhoods in the city: The Old North, Lev HaIr ,Florentine, and Neve Ofer (10 roofs from each neighborhood).

² The entire photo collection can be viewed at the following URL:
<https://drive.google.com/drive/folders/1vfw8rcR5R5w6PNw23TorsXuhamreHC8?usp=sharing>

The results of the study (Figure 1-2) demonstrate the differences between the various neighborhoods, which are particularly noticeable between each other and in particular between the northern (The Old North, Lev HaIr), typified by high socio-economic levels, and southern neighborhoods (Neve Ofer, Florentine), typified by low socio-economic levels (Schnell, 2009). According to the results, the following parameters were identified as being relevant in the context of how the roofs of the city are used: height of the building, number of floors, the division of the roof into separate areas or not, the existence of balconies in the building, the position of building on the plot (part of a larger block or an individual building) and additional construction on the roof. Furthermore, the study has also revealed that there is a strong connection between the observed uses (that occur on the same roof) in the southern neighborhoods, but that this connection is greatly weakened in the northern neighborhoods.

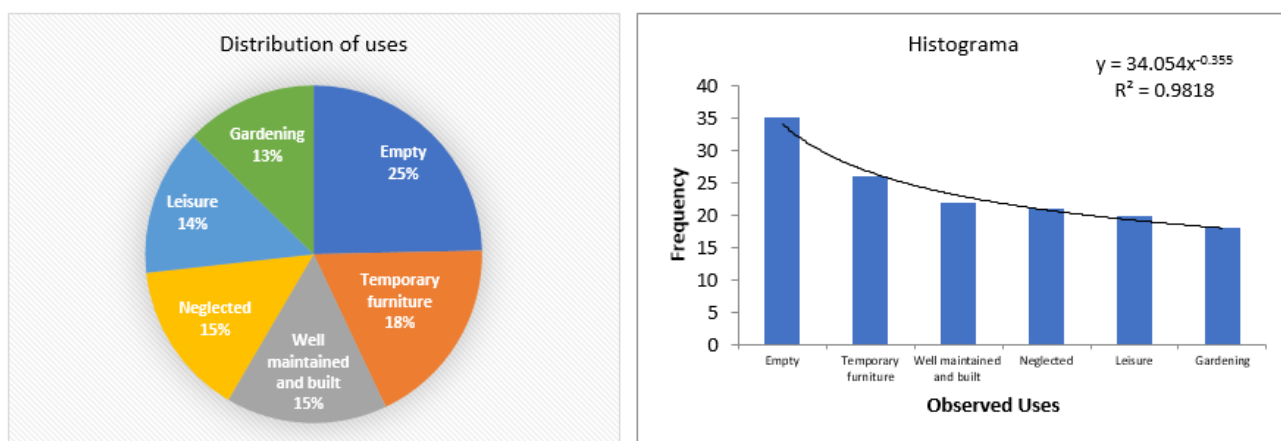


Figure 1: Roofs south of the Yarkon River

Finally, the rooftop space in Tel Aviv is a space with a very wide variety of connections at different levels, that are affected by the complex of life that takes place below and above the rooftops. In the absence of thorough research into this space, there is potential for understanding and developing it optimally - in order to benefit the residents of the city.

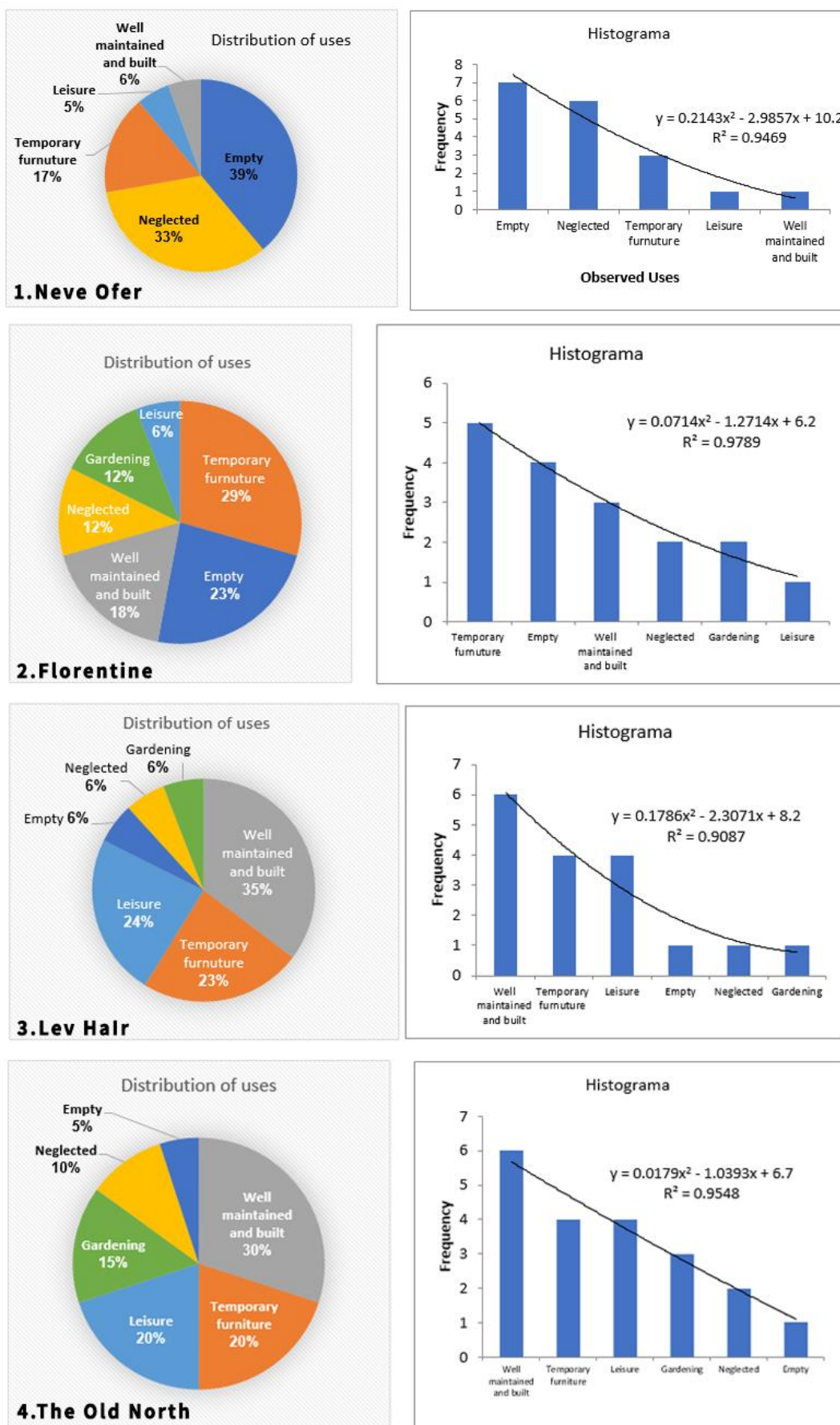


Figure 2: Roofs from four neighborhoods in the city: Neve Ofer, Florentine, Lev HaIr and The Old North

References

City of Melbourne, (2019). *Green Our Rooftop project* [Online]. Available at:

<https://www.melbourne.vic.gov.au/community/greening-the-city/green-infrastructure/Pages/green-our-rooftop-project.aspx> (Accessed: 1 September 2020).

Pons, O. *et al.* (2015) 'Roofs of the Future: Rooftop Greenhouses to Improve Buildings Metabolism', *Procedia Engineering*, 123, pp. 441–448.

Schnell I. (2009) 'The social areas of Tel Aviv-Yafo' in Baruch, K. (ed.) *Tel Aviv-Yafo : from a garden suburb to a world city*. Haifa: Pardes, pp. 339-361.

Tel Aviv Yafo, (2020). Tel Aviv Municipality's GIS system [Online]. Available at: <https://gisn.tel-aviv.gov.il> (Accessed: 1 August 2020).

Tsarfati, M. and Shafran, R. (2018) *Sustainability and the environment as a development tool in local government - a guide for local government*. Israel: Ministry of Interior.

The future of high streets: investigating post-pandemic scenarios in the UK

Anna ZANCHETTA¹; Mark BIRKIN^{1,2}

¹Alan Turing Institute

96 Euston Road, NW1 2DB, London

azanchetta@turing.ac.uk

² School of Geography and Leeds Institute for Data Analytics, University of Leeds

Worsley Building, LS2 9NL, Leeds

Keywords: high streets, spatial interaction models, land use, pandemic

In recent years, commercial activities on and around the UK high streets have experienced an increased economical fragility, exacerbated first by the financial crisis of 2007-2008 and accelerated further by a steady erosion of business to online retailing and by the ongoing global pandemic of coronavirus disease 2019. It is of interest of local and governmental stakeholders and policy makers, to understand the patterns of land use in the UK high streets and what to expect as a possible response to such stressful situations, in order to address the changes through informed decisions and policy design. In this context, tools like geographical and statistical analysis and modelling can help delineate the current and future scenarios in the UK urban areas.

The paper describes the research undertaken to model and analyse such a compelling phenomenon, connecting the current pandemic situation to longer-term issues of digital planning, land use modelling and social forecasting.

1. Introduction

During the course of the XX century, the land use along the UK high streets has become predominantly characterised by activities related to selling goods and services, commonly referred to as ‘retail’. Over the years, the high streets became increasingly “retail-focused” (HCLG, 2019), and this became particularly evident with the financial crisis of 2007-2008, that negatively affected consumer confidence for at least the following five years (Wrigley, 2015). Concurrently, the traditional way of purchasing in physical stores has been profoundly disrupted by the growth of online alternatives like e-commerce, online shopping and ‘click & collect’ systems (HCLG, 2019, Davies, 2018).

In this scenario, governmental and local bodies have invested in understanding how to help high streets and town centres to cope and adapt in order to survive, shifting to models of urban landscape where the social interactions and local identity are enhanced (HCLG, 2019). The current global pandemic of coronavirus disease 2019 (COVID-19) added further stress to this situation, and a general question on the high streets’ retail survival has arisen, as a consequence of policy measures (shop closures,

lockdown, remote working) and consequent people behaviour (increase in online shopping, limited movements, reduced commuting).

The aim of this research is to help delineate the actual and optimal allocations of land use along UK high streets and city centres, considering the current situation and several post-pandemic scenarios. The research is still in its early stages, thus this paper will report on early methodological developments and on the planned activities for the coming future.

2. Methodology and data

The research programme originates in discussions with the Cities and Local Growth Unit of the Ministry of Housing, Communities and Local Government (MHCLG).

The scope of the research is to describe land use and property characteristics in UK high streets in connection with population activity patterns, to identify current and future demand for retail services. Focusing on economic recovery, the idea is to delineate the current allocation of urban land use categories and to model scenarios of changes in provision, against future behaviour patterns and political strategies as a response to the current pandemic situation.

We are concentrating especially on four groups of individuals and their demand for retail: residential, workers, students, visitors (both internal and international).

The methodology employed in this research involves the use of SIMs to investigate the flows of the four groups of interests to the retail centres. At first the analysis is performed on the student population, with the aim to build a methodology that can be applied to the other groups.

2.1 Spatial interaction models

SIMs are statistical models for predictions of flows between origins and destinations, widely applied in disciplines such as social science, transport studies and urban planning (Birkin, 2019). The first formulations of SIMs were developed in the early XX century, but it was only in the 70s that Wilson improved them by using entropy maximisation and providing a better theoretical justification (Birkin, 2019). Wilson introduced four types of SIMs, depending on the pieces of information available on the flows whether on the demand (or production, the origin side) and/or the attraction (the destination side) or none. In our case the origin totals are known, that is the number of individuals living in the i zones, so we can treat the system as a production constrained model (Wilson, 1971), where we want to predict the flows to the j destinations (the retail centres).

The application of this model requires a proper calibration on actual flows data, i.e., an evaluation of the distance decay parameter. As we are currently looking for real flows data for the calibration, the

employment of a variation of this model, the radiation model proposed by Simini et al. in 2012, is also under investigation. This model has the advantage of being parameter free, so it can be applied where no measurements of flows are available (Lovelace, 2016).

2.2 Data

Most of the data used in this research comes from the UK Office for National Statistics (ONS): the demographic data (population characteristics, occupation) at LSOA levels from the Census 2011 data and the geographic data from the Open Geography portal.

The list of UK retail centres, with their boundaries, is provided by the University of Liverpool and is available at CDRC (2021).

In the first instance, population flows will be calibrated against sources including Google mobility reports and Census journey-to-work flows.

2.3 Pandemic and post-pandemic scenarios

To define different scenarios, and their implication on the high streets, we first need to divide retail centres between ‘high streets’ and ‘non-high streets’ centres (for example out of town shopping centres). Subsequently, a calibrated variation of both the origins and the destinations will allow to delineate the scenarios, i.e., changing both the number of individuals per origin and the attractiveness of the retail centres.

2.4 Workflow

The proposed analysis workflow comprises the following steps:

- Generate centroids for the population (origins) and the retail centres (destinations)
- Generate the origin-destination matrix and the other variables necessary to run the model
- Run the model on the student population, then perform calibration
- Run the model on different scenarios for different groups of individuals and perform analysis of the results

3. Preliminary results, discussion and conclusion

3.1 Application to retails and high streets in the UK

A preliminary analysis was run using R and QGIS on the Local Authority District (LAD) of Leeds, for the student population (a map of the results is shown in Fig. 1). The retail centres attractiveness was evaluated on a scale from 1 to 10; for the distance decay parameter a value of 0.43 was adopted

as in Waddington et al. (2019).

We expect that different scenarios will create different patterns. Further statistical analysis will then allow to draw conclusions on the several scenarios for each group of individuals.

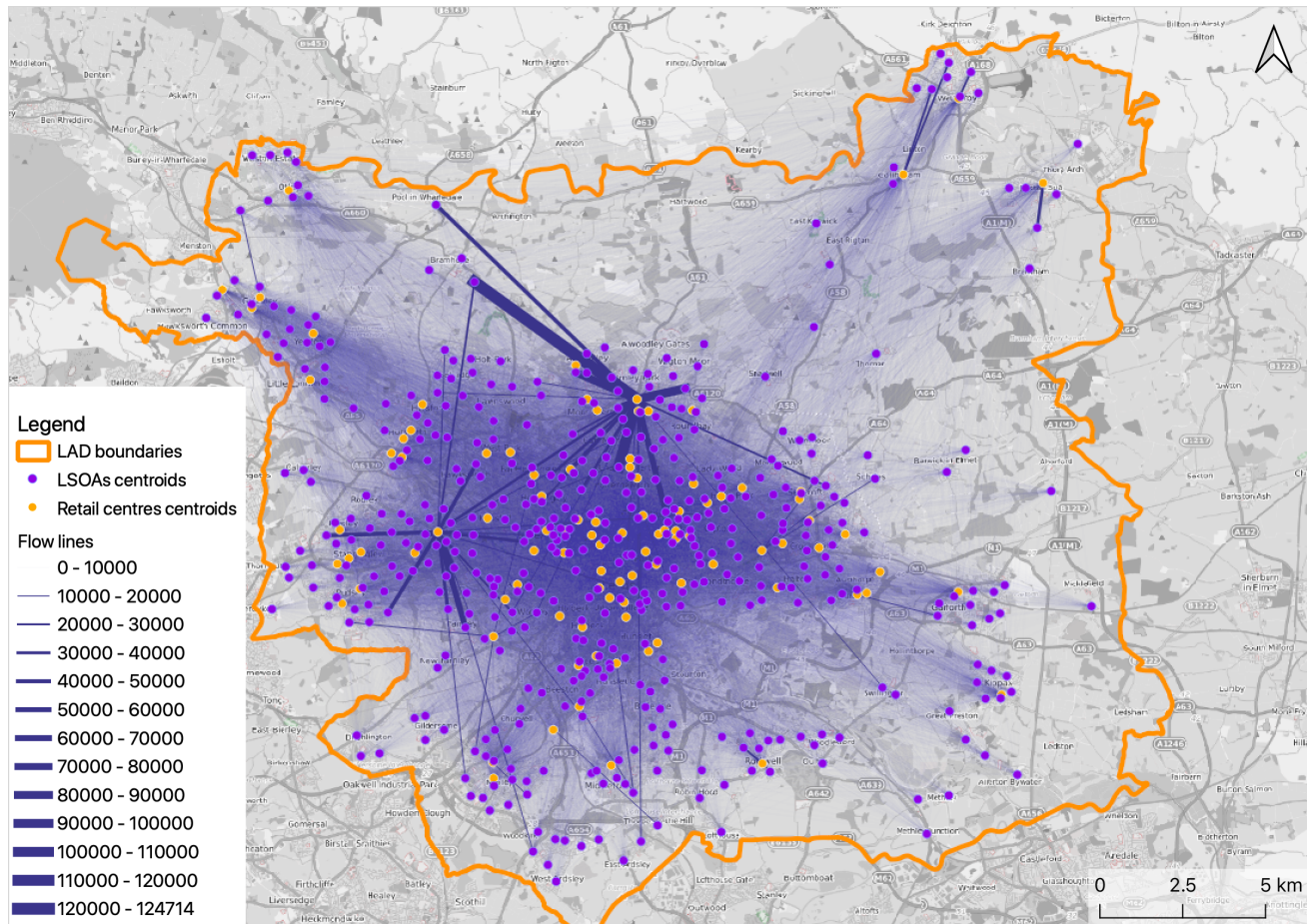


Figure 1: Flows of students from the LSOAs centroids (population weighted) to the retail centres centroids, Leeds LAD

3.2 Discussion

The framework developed in the research could potentially be used as a basis for economic development and policy response to a wide range of economic and social impacts.

Open methodological questions:

- At which level to run the analysis: national scale or regional/LAD
- Dimension of attractiveness values
- Calibration of the model

Acknowledgments

This research was possible thanks to the Wave 1 of The UKRI Strategic Priorities Fund under the EPSRC Grant EP/W006022/1, particularly the “Ecosystem of Digital Twins” theme within that grant & The Alan Turing Institute.

References

- Birkin, M. and Clarke M. (2019). ‘Applied spatial modelling in the twenty-first century: the Wilson legacy. Looking back and looking forward’, *Interdisciplinary Science Reviews*, 44(3-4), 286-300
- CDRC (2021). *Retail centre boundaries*. Available at: <https://data.cdrc.ac.uk/dataset/retail-centre-boundaries>.
- Davies, A., Dolega, L., Arribas-Bel, D. (2018). ‘Buy online collect in-store: exploring grocery click&collect using a national case study’, *IJRDM* 47(3), 278-291.
- HCLG Committee (2019). ‘High streets and town centres in 2030, Eleventh Report of Session 2017–19’. Available at: <https://publications.parliament.uk/pa/cm201719/cmselect/cmcomloc/1010/full-report.html>
- Lovelace, R., Birkin, M., Cross, P., Clarke, M. (2016). ‘From Big Noise to Big Data: Toward the Verification of Large Data sets for Understanding Regional Retail Flows’, *Geographical Analysis* 48, 59–81
- Portas, M. (2011). ‘The Portas Review, An independent review into the future of our high streets’. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/6292/2081646.pdf
- Simini, F., M. C. Gonzalez, A. Maritan, A.L. Barabasi (2012). ‘A Universal Model for Mobility and Migration Patterns’, *Nature*, 484, 96–100.
- Waddington T., Clarke G., Clarke M.C. et al. (2019). ‘Accounting for temporal demand variations in retail location models’, *Geographical Analysis*, 51(4), 426-447
- Wilson, A.G. (1971). ‘A Family of Spatial Interaction Models, and Associated Developments’. *Environment and Planning*, 3, 1-32
- Wrigley, N., Lambiri, D. (2015). ‘British High Streets: from Crisis to Recovery? A Comprehensive Review of the Evidence’. *ESRC/University of Southampton*. Available at: <https://thegreatbritishhighstreet.co.uk/pdf/GBHS-British-HighStreets-Crisis-to-Recovery.pdf>

Special Session SS10.1

5 November 9:00-11:00 am (GMT)

Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance

Proponents

Joana Barros, Chen Zhong, Yang Yue

Does high-density living facilitate non-motorized travel for every resident? Examining the moderating effect of individual daily time-space arrangement in Guangzhou, China

Junwen Lu¹; Suhong Zhou²

¹School of Geography and Urban Planning, Sun Yat-sen University
No. 135, Xingangxi Road, Haizhu District, Guangzhou, 510275, China
+86 13570965427, lujunwen@mail2.sysu.edu.cn

² School of Geography and Urban Planning, Sun Yat-sen University
No. 135, Xingangxi Road, Haizhu District, Guangzhou, 510275, China
+86 13825044799, eeszsh@mail.sysu.edu.cn (correspondent author)

Keywords: daily activity, time-space arrangement, living density, non-motorized travel, moderating effect

1. Introduction

Promoting non-motorized travel has been one of the most important goals of sustainable and healthy urban planning and governance. Many factors affect residents' non-motorized travel rate, such as built environment, community security, quality of open space, individual social status, and so on (Board, 2012; Joh et al., 2009; Targa and Clifton, 2005). Among them, living density, as an important dimension of the built environment, has been widely discussed and has been employed as one of the indicators of sustainable development in a few practices. For example, the World Health Organization recommends increasing compactness of the built environment to prevent obesity and obesity-related chronic diseases (Organization, 2018). It's generally believed that living density is positively associated with non-motorized travel, because high-density living brings compact neighborhoods which provide walkable streets and diverse functions within walking distance (Rodríguez and Joo, 2004; Stefansdottir et al., 2019). Increasing living density seems to be an ideal approach to building non-motorized-travel-friendly cities.

However, the relationship between living density and non-motorized travel was proved to be context-sensitive in recent studies which focused on high-density regions. For example, Yin et al. (2020) examined the non-linear association between population and travel behavior in Chinese cities, and their result proved that living density facilitates non-motorized travel within a certain range and once the density exceeded a certain threshold, its marginal effect diminished sharply. Scholars have provided multiple explanations for high-density living's possible negative effect on non-motorized

travel, including reducing green space provision, bringing poor air quality, leading to stressful lifestyles, lacking healthy food and so on (Day, 2016; Lewis, 2018; Meyer et al., 2012). These explanations make a lot of sense by emphasizing the environmental influence of high-density living and should be well-considered in planning practice. But still, most of the research, though realized the significant effects of individual heterogeneity, potentially admitted that individual activity patterns are homogeneous, which may have impeded the understanding of high-density living. High living density is accompanied by heterogeneous lifestyles and spawns diverse daily activity patterns (Qizhi et al., 2016; Xu et al., 2016). It has been proved that individual daily activity pattern determines people's specific travel behavior to a great extent, sometimes even share greater utility than that of the built environment, in high-density cities (Cao and Fan, 2012). Therefore, examining how living density affects non-motorized travel of residents with different daily activity patterns is of great significance in promoting a better understanding of the effects of high-density living and individual-based policy-making.

The time geography theory provides a great framework to examine how individuals arrange their daily travels and activities in space and time. By introducing the concepts of *anchor point*, *path*, *prism*, *flexibility*, and *constraint*, it enables us to conceptualize individuals' daily activity and conclude their daily time-space arrangement patterns (Ellegård, 2018). As Miller et al. (2018) have found, most people project their daily activities based on home and workplace anchor points and make time-space arrangements under spatial and temporal constraints. Some of the time-space arrangement patterns lead to sustainable mobility and some can not (Figure 1). Therefore, the study sets out to conclude residents' daily time-space arrangement patterns under the time geography framework and examine its moderating effect on the relationship between living density and non-motorized travel. It's hoped that the results can inspire targeted policy-makings in the future.

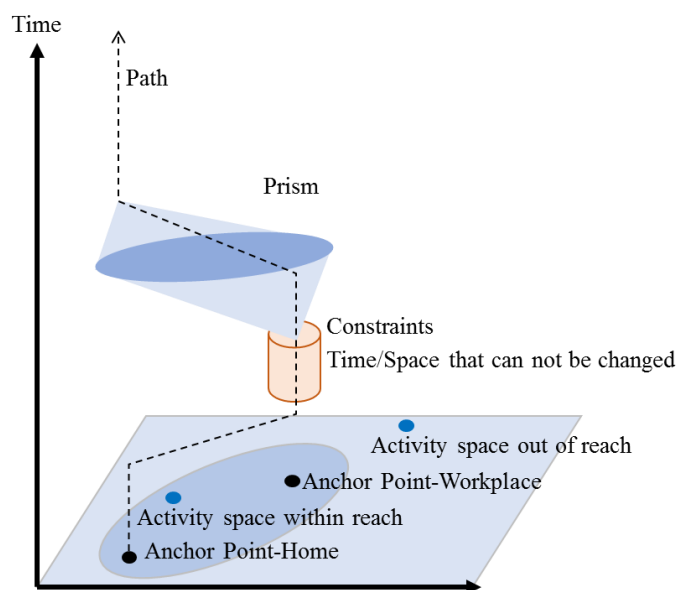


Figure 1: Individual daily time-space arrangement based on home and workplace anchor points

2. Data and method

The study employs one-month mobile phone data of approximately 1.8 million users of October 2020 to examine and conclude their daily time-space arrangement patterns. The data was provided by a major mobile communication supplier in China which holds about 30 percent market share in Guangzhou. It recorded each users' basic attributes and position changes throughout the day. Users' homes and workplaces were identified as the longest stayed locations during sleeping hours (from 10 p.m. to 5 a.m.) and working hours (from 10 a.m. to 5 p.m.). Residents' non-motorized travels were identified according to their speed. The public green and open space were derived from the land-use survey in 2017.

The principal component analysis and k-means cluster analysis were used to conclude time-space arrangement patterns. A linear regression model with moderating variables was used to examine the influence of living density on non-motorized travel the moderating effects.

3. Results and conclusions

Six time-space arrangement patterns were identified basing on time periods a resident first leave home, first arrive workplace, last leave workplace, last arrive home, and the entropy index of the periods on workday and weekend respectively. It's found that living density is positively correlated with time-space arrangement diversity. And despite the living density facilitates non-motorized travel in general, the effect is significantly moderated by time-space arrangement patterns when controlling for basic individual attributes and accessibility of public green and open space.

The results explained the context-sensitive effect of living density on non-motorized travel at individual level by introducing the time-geography framework and examining the moderating effect of individual daily time-space arrangement. It has also identified the groups that are disadvantaged at daily time-space arrangement which prevents them from getting the benefits of high-density living. These groups should be the targets of future policy-making.

Acknowledgments

The research is funded by JPI Urban Europe-NSFC Project “Sustainable Mobility and Equality in mega-city Regions: patterns, mechanisms and governance”.

References

- Board, N. (2012). *Environmental justice, social factors, and gender-related issues in transportation 2012*. Transportation Research Record: Journal of the Transportation Research Board.
- Cao, X. & Fan, Y. (2012). ‘Exploring the influences of density on travel behavior using propensity score matching’ *Environment & Planning B Planning & Design*, 39 (3), pp. 459-470.
- Day, K. (2016). ‘Built environmental correlates of physical activity in China: A review’ *Preventive Medicine Reports*, 3 pp. 303-316. DOI: <https://doi.org/10.1016/j.pmedr.2016.03.007>
Available at: <https://www.sciencedirect.com/science/article/pii/S2211335516300109>.
- Ellegård, K. (2018). *Thinking time geography: Concepts, methods and applications*. Routledge.
- Joh, K., Boarnet, M. G. & Nguyen, M. T. (2009). ‘Interactions among race or ethnicity, attitude, and crime: Analyzing walking trips in the south bay area’, *Transportation Research Board Meeting*, 2009.
- Lewis, S. (2018). ‘Neighborhood density and travel mode: New survey findings for high densities’ *International Journal of Sustainable Development & World Ecology*, 25 (2), pp. 152-165. DOI: 10.1080/13504509.2017.1321052
Available at: <https://doi.org/10.1080/13504509.2017.1321052>.
- Meyer, K. A., et al. (2012). ‘Sleep duration and bmi in a sample of young adults’ *Obesity*, 20 (6), pp. 1279-1287.
- Miller, H. J., Song, Y. & Tribby, C. P. (2018). ‘Green, healthy time-geography: Using time-geographic concepts for sustainable mobility research’, *Time geography in the global context*: Routledge, pp. 75-95.
- Organization, W. H. (2018). *The global action plan on physical activity 2018 - 2030: More active people for a healthier world*. Geneva, Switzerland: World Health Organization.
- Qizhi, M., Ying, L. & Kang, W. (2016). ‘Spatio-temporal changes of population density and urbanization pattern in china (2000-2010)’ *China City Planning Review*, 25 (4).
- Rodríguez, D. A. & Joo, J. (2004). ‘The relationship between non-motorized mode choice and the local physical environment’ *Transportation Research Part D: Transport and Environment*, 9 (2), pp. 151-173.
- Stefansdottir, H., Næss, P. & Ihlebæk, C. M. (2019). ‘Built environment, non-motorized travel and overall physical activity’ *Travel behaviour and society*, 16 pp. 201-213.
- Targa, F. & Clifton, K. J. (2005). ‘Built environment and nonmotorized travel: Evidence from baltimore city using the nhts’ *Journal of Transportation & Statistics*, 8 (3), pp. 55-70.
- Xu, Y., et al. (2016). ‘Another tale of two cities: Understanding human activity space using actively tracked cellphone location data’ *Annals of the American Association of Geographers*, 106 (2), pp. 489-502.

Yin, C., Cao, J. & Sun, B. (2020). 'Examining non-linear associations between population density and waist-hip ratio: An application of gradient boosting decision trees' *Cities*, 107 (3), p. 102899.

How activity pattern associates with income status? Evidence from transit smart card data and AI methods

Qi-Li Gao^{1,2}; Chen Zhong³; Yang Yue^{1,4}

^{1,2} Shenzhen Key Laboratory of Spatial Smart Sensing, Shenzhen University

Shenzhen, 518060, China

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University

Hong Kong

qili.gao@outlook.com (correspondent author)

³ The Bartlett Centre for Advanced Spatial Analysis (CASA), University College London (UCL)

Gower Street, London WC1E 6BT, UK

c.zhong@ucl.ac.uk

⁴ Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University

Shenzhen, 518060, China

yeuyang@szu.edu.cn

Keywords: social inequality, socioeconomic status, activity pattern, smart card data, AI.

1. Introduction

Obtaining socioeconomic status (SES) is of great significance in social research, commercial intelligence, urban policy as well as transportation management (Ding, Huang, Zhao, & Fu, 2019). In particular, the disparities in activity patterns among socioeconomic groups have been considered as an important aspect of social inequality, such as social exclusion (Schönfelder & Axhausen, 2003). Many studies have documented that human activity patterns are highly associated with personal socioeconomic factors (Goulet-Langlois, Koutsopoulos, & Zhao, 2016; Lu & Pas, 1999; Zhang, Sari Aslam, Lai, & Cheng, 2020). Despite the great contributions of these studies, the evidence is still insufficient in several aspects. One of the main reasons is the difficulty of simultaneously obtaining activity data and socioeconomic data of a large number of people. Some works relied on detailed survey-based activity data, the samples were usually too small to be representative enough (Xie, Xiong, & Li, 2016). The availability of human mobility data makes it possible to track large-scale individuals' activity behaviours, the lack of socioeconomic information in such datasets poses challenges in building the relationship between SES and activity patterns (Ghosh & Ghosh, 2017). Although there has been an increase in combining human trajectory data and travel surveys, these studies mainly focused on demographic attributes, such as gender and age, the economic features especially income are insufficient due to personal sensitivity about economic level (Wu et al., 2019). Another disadvantage is the loss of interpret-ability of the prediction results derived from AI-based methods

(Zhang et al., 2020). Therefore, how and to what extent income status influences human activity behaviours have not been well addressed.

To tackle the above mentioned challenges, this research combines smart card data and travel survey data to build the link between income level and human activities. This study involves 1) the extraction of multi-dimensional activity features to capture the spatio-temporal patterns of individuals. 2) the estimation of socioeconomic attributes to represent the contexts of activity places. 3) the prediction of machine learning models to reveal the association between income level and human activity behaviours. The methods of explainable AI are used for the interpretation of features. The framework is applied to a case study of public transit users in Shenzhen, China to validate its effectiveness.

2. Dataset

2.1 Smart card data

The primary dataset in this study is six days of smart card records of public transit users observed in November 2016. The data consists of all trips by bus and subway. Subway records contain anonymous user ID, the origin and destination station, as well the tap-in and tap-out time of each trip. Bus records contain boarding information, including boarding time, bus number and the bus line information. OD trajectories were inferred from smart card records based on spatio-temporal regularities.

2.2 Travel survey data

The Shenzhen travel survey was conducted in the same period with the smart card data. A total of 68,029 households were interviewed face to face about the characteristics of households and travel-related questions. Household annual income was graded into five levels (level 1=(0, 100k), level 2=(100k, 200k], level 3=(200k, 300k], level 4=(300k, 500k], level 5=(500k, +∞), currency: RMB).

3. Methodology

3.1 Activity pattern characterization

The overview of the methodology framework is illustrated in Figure 1. We concluded our features of characterizing human activity patterns into six categories: activity intensity, activity extensity, activity diversity, travel efficiency, spatial location and temporal rhythm. These features quantify characteristics of travel trajectories from different dimensions and capture different perspectives of activity patterns. All the activity features were aggregated at the station scale by residential place.

3.2 Socioeconomic context estimation

We first defined 1 km buffer as the catchment area of a transit station, and proposed a population-weighted approach to calculate income levels at transit stations. Land use diversity, population density and transit accessibility, donated by density of bus stops and availability of subway in the station catchment area were calculated as the socioeconomic attributes of stations.

3.3 Machine learning models for regression

This study used linear regression (LR) model as baseline, and chose three widely-used models to perform regression task, including support vector machine (SVM), random forests (RF) and XGBoost. Three kinds of methods were adopted to measure feature importance: mean decrease impurity, permutation importance and Shapley Additive exPlanations (SHAP) value.

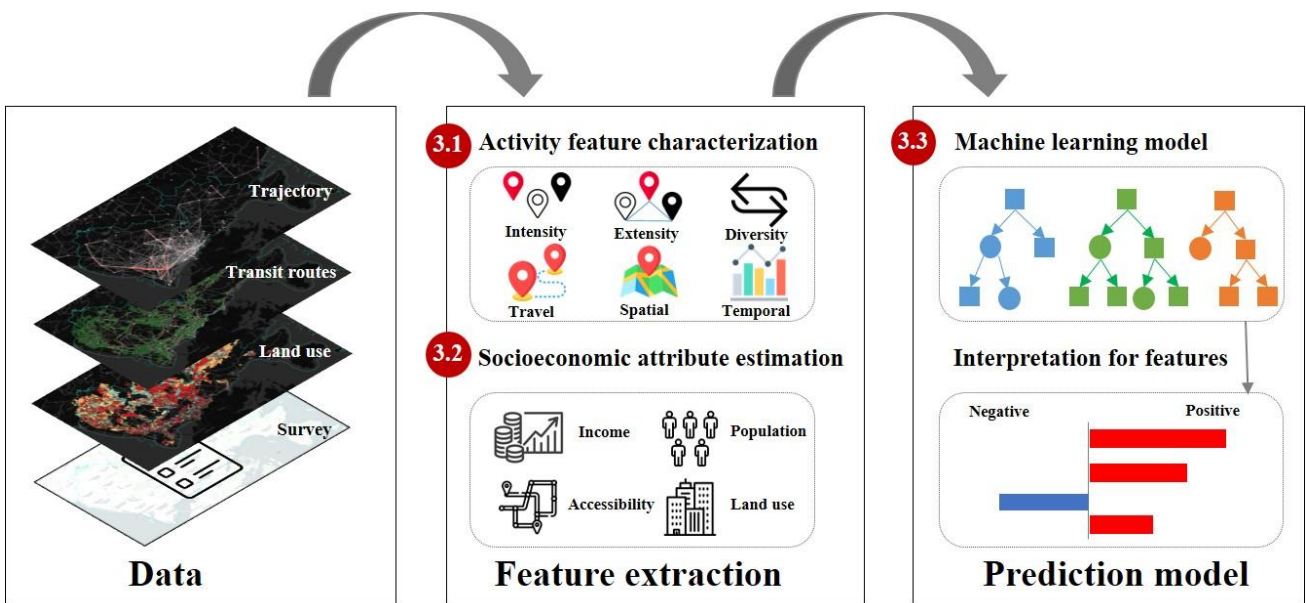


Figure 1: A overview of the methodology framework

4. Results and Discussion

4.1 Prediction results

The regression results of various combinations of features and models are illustrated in Table 1. When only activity features were considered, we yielded the lowest R^2 of 0.413 by LR model and the highest R^2 of 0.677 by RF. When we explored the association between income level and socioeconomic contexts at residence, all the models yielded better performances compared with activity features. The results indicate that income level is more related to residential socioeconomic status than other activity patterns. From the perspective of social segregation, residential segregation is more obvious than activity-related differentiation. When taking both activity features and socioeconomic features, we achieved the highest R^2 of 0.871 and the lowest MSE of 0.015 by RF. The improvement of model

performance suggests that although income level primarily influence the choice of residential place, there is some impacts on other activity patterns, which should not be ignored.

Table 1: Regression model and results

Feature dimension	Model	R^2	MSE
Activity features	LR	0.413	0.069
	SVM	0.408	0.070
	RF	0.677	0.038
	XGBoost	0.671	0.039
Socioeconomic features	LR	0.263	0.095
	SVM	0.259	0.095
	RF	0.840	0.021
	XGBoost	0.759	0.031
Activity features + Socioeconomic features	LR	0.475	0.061
	SVM	0.474	0.061
	RF	0.871	0.015
	XGBoost	0.839	0.019

4.2 Feature importance

RF model performed the best in regression results, and thus was selected to interpret feature importance and the association between income variable and activity patterns. As Figure 2 shows, although feature importance varies slightly between different evaluation methods, the general trend of importance rank is: spatial location > socioeconomic context > activity extensity > activity intensity > travel efficiency > activity diversity > temporal rhythm. The results confirm that income status has the greatest impact on spatial locations and residential socioeconomic contexts. The SHAP value shows that wealthy people prefer the locations with diverse urban functions, better transport accessibility and low population density.

5. Conclusion

This study explored the connection between income status and various activity patterns by combining human mobility data and travel survey data. The results imply that income level is highly associated with human activity behaviours. However, spatial location and socioeconomic contexts (population density, land use diversity and transport accessibility) play the most important role in determining the prediction performance, suggesting that residential segregation by income is more significant than other activity-related differentiation.

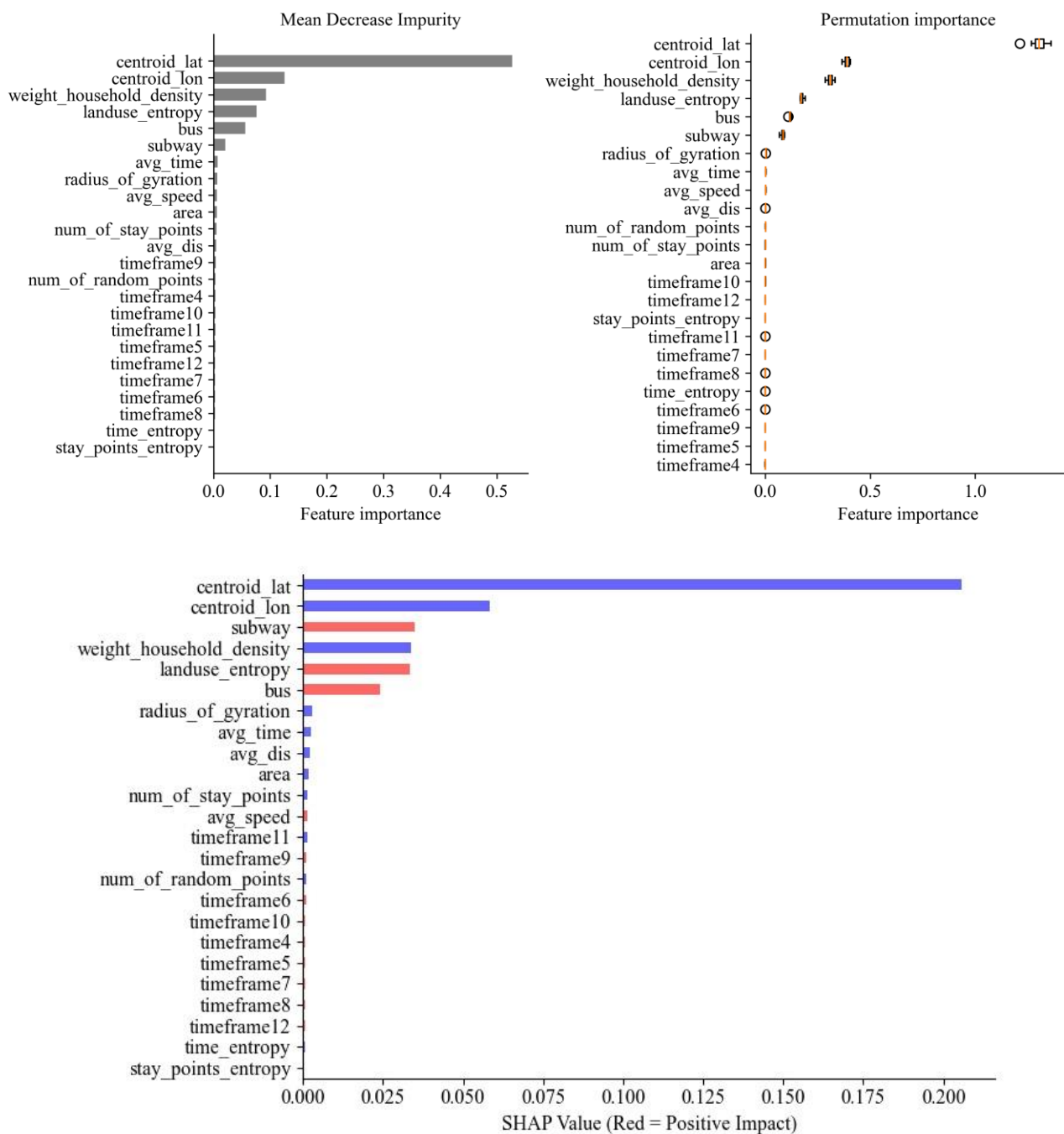


Figure 2. Feature importance

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 42001390, 71961137003) and the Postdoctoral Science Foundation of China (Grant No. 2021M692163) and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 949670), and from ESRC under JPI Urban Europe/NSFC (Grant No. ES/T000287/1).

References

- Ghosh, S., & Ghosh, S. K. (2017). Modeling of human movement behavioural knowledge from GPS traces for categorizing mobile users. Paper presented at the Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia.
- Goulet-Langlois, G., Koutsopoulos, H. N., & Zhao, J. (2016). Inferring patterns in the multi-week activity sequences of public transport users. *Transportation research part C: emerging technologies*, 64, 1-16.
- Lu, X., & Pas, E. I. (1999). Socio-demographics, activity participation and travel behavior. *Transportation Research Part A: Policy and Practice*, 33(1), 1-18.
- Schönfelder, S., & Axhausen, K. W. (2003). Activity spaces: measures of social exclusion? *Transport Policy*, 10(4), 273-286.
- Wu, L., Yang, L., Huang, Z., Wang, Y., Chai, Y., Peng, X., & Liu, Y. (2019). Inferring demographics from human trajectories and geographical context. *Computers, Environment and Urban Systems*, 77, 101368.
- Xie, K., Xiong, H., & Li, C. (2016). The correlation between human mobility and socio-demographic in megacity. Paper presented at the 2016 IEEE International Smart Cities Conference (ISC2).
- Zhang, Y., Sari Aslam, N., Lai, J., & Cheng, T. (2020). You are how you travel: A multi-task learning framework for Geodemographic inference using transit smart card data. *Computers, Environment and Urban Systems*, 83, 101517.

The Role of Mobility on Daily Spatial Quality Exposure Equality: Empirical evidence from Guangzhou, China

Zhou Suhong^{1,2}; Wang Linsen^{1,2}; Song Jiangyu^{1,2}; Lu Junwen^{1,2}; Zheng Zhong³

¹ School of Geography and Planning, Sun Yat-sen University

Guangzhou, Guangdong, 510000, China

eeszsh@mail.sysu.edu.cn (correspondent author)

² Guangdong Provincial Engineering Research Center for Public Security and Disaster

Guangzhou, Guangdong, 510000, China

³ Center for Territorial Spatial Planning and Real Estate Studies, Beijing Normal University

Zhuhai, Guangdong, 519000, China

Keywords: mobility, spatial quality exposure, equality, Guangzhou

Environmental problems such as air and noise pollution, disorder in the built environment, and poor ambient space quality have emerged and seriously affected people's quality of life. It has caused a series of unfairness problems, especially for low-income groups, who usually suffer from worse environmental exposure because of their lower socioeconomic status. The previous literature mainly focused on the differences in people's exposure to the physical and social environment around their residential areas, but lack to consider the spatial environment that people perceive in daily activities, such as safety, liveliness, beauty and mess, which can more comprehensively reflect the quality of space. In China, people are affected by the housing system and the sharp rise in housing prices, different income groups distribute to residential environments with distinct space quality due to gaps in their affordability. Therefore, the differentiation of socioeconomic status often leads to the differentiation of space quality around residence (SQAR) first. And people mainly revolve around residence because of spatiotemporal constraints in daily activities, so people's income level affect their level of exposure to ambient space quality (EASQ) severely. However, it is still possible for people to break through this restriction and access higher-quality spaces, which will depend on their spatiotemporal constraints in daily activities. Human mobility is the most intuitive representation of spatiotemporal constraints, reflecting the active degree of people moving in the urban space. Mobility will contribute to breaking through the restrictions of people's environment around residence, and increase their chances of accessing higher-quality spaces. There is a lot of evidences that mobility will make different sociodemographic groups' environmental exposure levels closer which is summarized as neighborhood effect averaging problem (NEAP).

To explore the above issues, this paper summarizes the following conceptual framework (Figure 1). It is mainly composed of three theoretical hypotheses: different income groups enjoy different level of EASQ (H1); different income groups distribute to environments with different SQAR due to the gap in ability to pay housing price, and then spatiotemporal constraints in daily activities lead to their different level of EASQ (H2); human mobility will help people break through the restrictions of their residence, thereby reducing the EASQ gap between different income groups (H3). In order to verify this framework, the conditional process analysis model can be used. Specifically, total effect of income level on EASQ is deconstructed into direct effect and indirect effect, and SQAR plays a role of mediation factor; the effect of direct and indirect way is moderated by mobility. In addition, the model also incorporates variables such as gender and age to control the influence of other factors that cause differences in activity patterns on the results of the experiment. The experiment will be organized as following process: Firstly, street-view image data of Guangzhou from Baidu Map Co., and online survey data about public perception of spatial environment are used to measure SQAR and EASQ with imaging analysis based on deep learning method, which include space safety, space liveliness, space beauty and space mess; secondly, we identify the user's residence by the mobile phone signaling data, and compare it with the house rent data to match spatial information, and then measure the income level of the user by the level of housing price around their residence; last but not least, SQAR, EASQ and mobility index of all income groups are calculated to compare the difference, and the conditional process analysis model is used to validates the above conceptual framework.

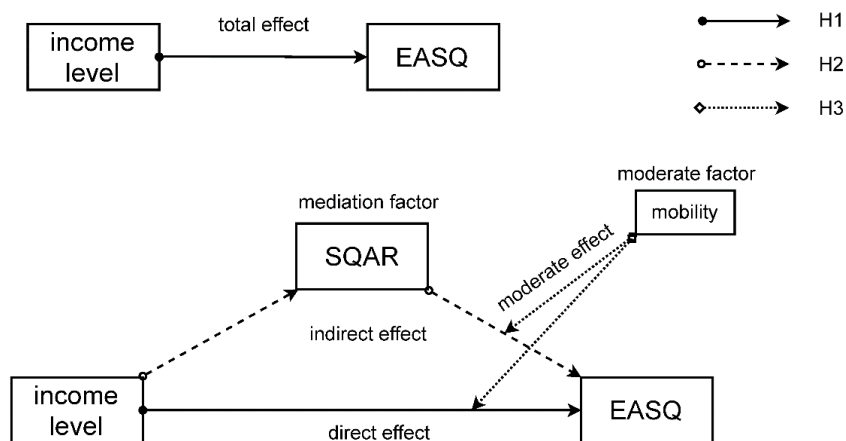


Fig.1 conceptual framework

The calculation results of the conditional process analysis model confirm that different income groups exposure to distinct ambient space quality and human mobility can alleviate that situation (Figure2). In support of H1, people’s income level significantly predicts their EASQ controlling with variables such as gender and age, $B = 0.098$, $SE_B < 0.000$, $p < 0.000$. In support of H2 and H3, it is also found that the effect of income level on EASQ is significantly mediated by SQAR, and moderated by

mobility, the index of moderated mediation is -0.002, BootstrapSE =0.0001, BootstrapCI [-0.0022, -0.0018]. And as shown in Figure 2, all of the total effect, the direct and indirect effects will significantly weaken as the mobility rises. it means that people truly break through restriction of residence environment, and access higher-quality spaces, promoted by higher active mobility.

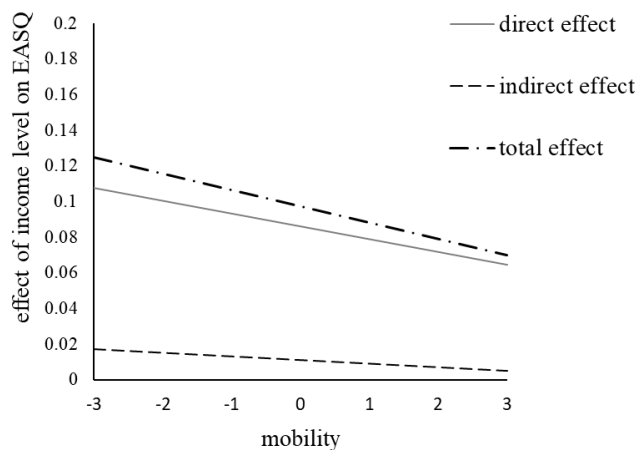


Fig.2 a visual representation of the conditional direct, indirect and total effect of income level on EASQ, and their change with mobility

A congenerous regional simulation model for land use-population-economy changes in the Greater Bay Area

Mingxiao Li¹; Wei Gao¹; Wei Tu^{1,*}; Bichen Fang¹

¹Guangdong Key Laboratory of Urban Informatics, Guangdong Laboratory of Artificial Intelligence and Digital Economy(SZ), Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services, and Research Institute of Smart Cities, Shenzhen University
Shenzhen 518060, China
+86 13480951947, tuwei@szu.edu.cn (correspondent author)

Keywords: regional simulation, congenerous simulation model, land use-population-economics change, cellular automata

Regional scenario simulation models are crucial and reproducible tools for analysing both the causes and consequences of future development patterns (Arsanjani et al., 2013; Chen et al., 2019; Liu et al., 2017). As important indexes of evaluating development trends, scenario simulations of land use, population, and economic development are of great significance to support future development policy decisions (Mustafa et al., 2018; Schwaab et al., 2018). As the Cellular automata (CA) model can generate rich patterns and effectively represent the development processes, a growing body of literature has applied CA models in regional development studies (Guzman et al., 2020; Tong and Feng, 2020). However, most of these models can only simulate the dynamic of one scenario. While, the development process of land use, population, and economy vary concurrently and affect each other (National Research Council, 2005). Thus, a congenerous simulation model will be more effective for portraying future development patterns.

In this study, we present a congenerous regional simulation model (CRS) for land use-population-economics scenarios simulating. The model first applied basic CA models to simulate the scenarios of land use, population and economy separately to capture the coarse development patterns. Based on the coarse simulation results, the simulation results of each other two scenarios will be used as part of the driving factors to revise the simulation result of that scenario, to portray the congenerous impact of the scenarios of land use, population, and economy. The same step is iteratively performed until the changes of each scenario converge to convergence. When performing simulation, the current status of land use, population distribution, and economic development pattern will be used as inputs to be passed into each step of trained models. We then obtain the integrated land use-population-economics congenerous simulation results. The illustration of our proposed congenerous simulation model in land use is shown in Figure 1.

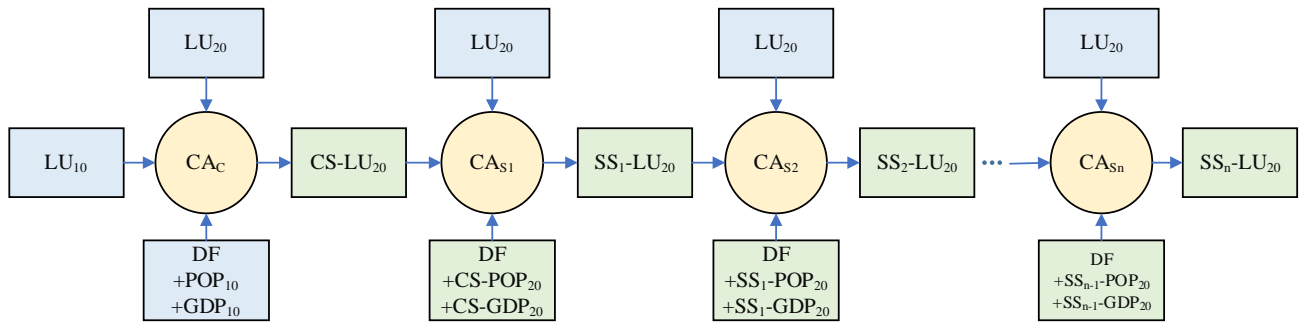


Figure 1: Illustration of congenerous regional simulation model in land use. (original data marked in blue, CA model marked in yellow, and simulated data marked in green).

By using the real-world data in Guangdong-Hong Kong-Macao Greater Bay Area from 2010 to 2020, we verified the effectiveness of our proposed congenerous regional simulation model. The results were shown in Table 1. The results indicated that our proposed model can capture the development process of land use, population and economy well, and outperformed baseline methods (Patch-generating Simulation Model, PLUS (Yao et al., 2017) and Future Land Use Simulation Model, FLUS (Liu et al., 2017)). The model was further applied to obtain the simulation of land use, population, and GDP in 2030 (as shown in Figure 2), which can support the applications such as policy decision and major infrastructure siting optimization.

Table 1: Comparison of simulation performance with baselines.

Scenario	Method	Accuracy	Kappa	MAPE	RMSE
Land Use	CRS	83.61%	0.737	--	--
	PLUS	83.36%	0.733	--	--
	FLUS	78.86%	0.661	--	--
Population	CRS	--	--	14.58%	28.637
	PLUS	--	--	17.75%	39.499
	FLUS	--	--	19.95%	52.113
	WorldPOP	--	--	26.43%	59.413
GDP	CRS	--	--	19.59%	480.242
	PLUS	--	--	22.02%	516.057
	FLUS	--	--	43.43%	1196.795

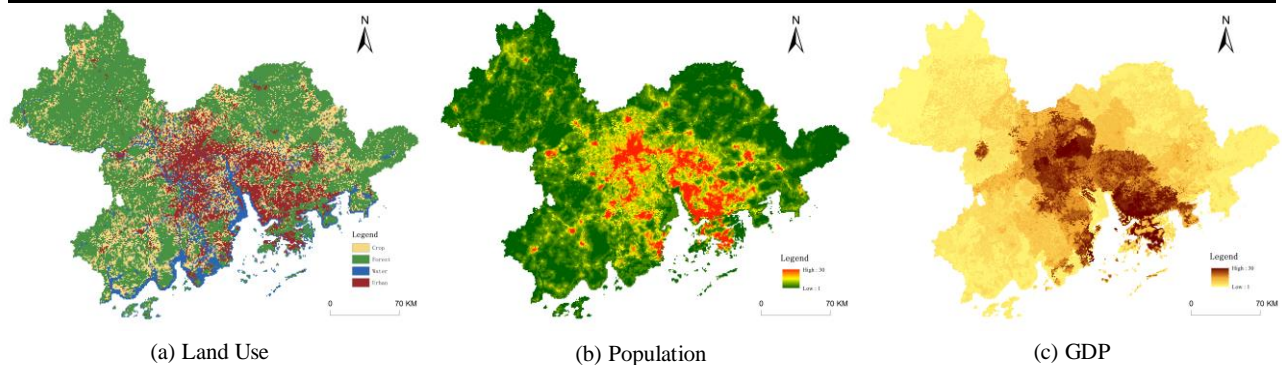


Figure 2: Simulation results of land use (a), population (b), and GDP (c).

Acknowledgments

This research was supported by the National Key Research and Development Program [2019YFB2103104]; the National Natural Science Foundation of China: [71961137003, 42071360, 4210010338]; the China Postdoctoral Science Foundation: [2021M692164]; and the Guangdong Province Basic and Applied Basic Research Fund Project: [2020A151511166].

References

- Arsanjani, J.J., Helbich, M., Kainz, W., Bolorani, A.D., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation* 21, 265–275.
- Chen, Y., Li, X., Liu, X., Huang, H., Ma, S., 2019. Simulating urban growth boundaries using a patch-based cellular automaton with economic and ecological constraints. *International Journal of Geographical Information Science* 33, 55–80.
- Guzman, L.A., Escobar, F., Peña, J., Cardona, R., 2020. A cellular automata-based land-use model as an integrated spatial decision support system for urban planning in developing cities: The case of the Bogotá region. *Land use policy* 92, 104445.
- Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., Li, S., Wang, S., Pei, F., 2017. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning* 168, 94–116.
- Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., Teller, J., 2018. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Computers, Environment and Urban Systems* 67, 147–156.
- National Research Council, 2005. *Population, Land Use, and Environment: Research Directions*, The National Academies Collection: Reports funded by National Institutes of Health. National Academies Press (US), Washington (DC).
- Schwaab, J., Deb, K., Goodman, E., Lautenbach, S., van Strien, M.J., Grêt-Regamey, A., 2018. Improving the performance of genetic algorithms for land-use allocation problems. *International Journal of Geographical Information Science* 32, 907–930.
- Tong, X., Feng, Y., 2020. A review of assessment methods for cellular automata models of land-use change and urban growth. *International Journal of Geographical Information Science* 34, 866–898.
- Yao, Y., Liu, X., Zhang, D., Liang, Z., Zhang, Y., 2017. Simulation of Urban Expansion and Farmland Loss in China by Integrating Cellular Automata and Random Forest. arXiv preprint arXiv:1705.05651.

Spatial preferences of logistics development

Apeksha Tare¹; Merten Nefs²; Eric Koomen¹; Erik Verhoef¹

¹Department of Spatial Economics, Vrije Universiteit Amsterdam

De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

+31 20 5986090, a.tare@vu.nl (correspondent author)

² Deltametropolis Association

Museumpark 25, 3000 AP Rotterdam, The Netherlands

Keywords: logistics, land use, spatial modelling, accessibility

Abstract

This study aims to empirically estimate the relative impact of various accessibility, location and policy factors on logistics development as a land-use change process. A distinction is made between four different types of logistics development with the purpose of identifying any differences in the relative importance of spatial drivers for different types of logistics firms. Logistics land-use change data is derived from a historical dataset depicting logistics growth in the East - Southeast transport corridor of the Netherlands between 1980 – 2020 and a discrete choice modelling approach is employed.

1. Introduction

Owing to the strong growth of the logistics sector in recent decades, there has been an increase in demand for logistics real estate. The rapidly expanding spatial footprint of logistic complexes has led to logistics sprawl that has raised concerns regarding deteriorating quality of life in the hinterland (Aljohani and Thompson, 2016). This highlights the need for a better understanding of the spatial factors affecting location dynamics of logistics firms that will benefit stakeholders involved in the planning process, including policymakers and real estate developers. In this regard, past studies have largely focused on describing the spatial patterns of distribution centre locations and logistics sprawl, but very few have empirically identified the spatial drivers underlying these patterns (Bowen, 2008; Heitz et al., 2018; Onstein et al., 2019; van den Heuvel et al., 2013; Verhetsel et al., 2015; Woudsma et al., 2008). Although accessibility is believed to be a major factor, hardly any studies have quantified the effects of multiple aspects of accessibility on different types of logistics development, and the results are usually dependent on the regional context.

2. Region and policy context

The study focuses on the East - Southeast freight transport corridor of the Netherlands that stretches across four provinces, namely, North Brabant, Zuid-Holland, Gelderland and Limburg. This region, which lies between the sea port of Rotterdam and the German and Belgian hinterlands, is home to many logistics companies that serve key markets within 500 km in Northwest Europe, reaching a population of approximately 150 million. As a result, this region is also critical for the national spatio-economic policy agenda. One of the earliest policies in this regard was the Mainport policy of 1980s which sought to strengthen the position of the port of Rotterdam via the construction of port and hinterland infrastructures. More recently, the spatial-economic policy shifted focus, from financing heavy infrastructure in the entire hinterland corridor to stimulating economic 'Top Sectors' in specific locations. Logistics was one of these identified 'Top Sectors'. In this study, we seek to estimate the impact of this 'Top Sector' policy on logistics development.

3. Data and methodology

For the analysis we use open access geodata of logistic buildings in the East - Southeast corridor (Nefs, 2021) between 1980 and 2020. The dataset was compiled using various available sources, such as the Dutch building administration including construction year (BAG), a Dutch business estates database (Ibis) and company microdata (LISA). All buildings larger than 500 m² within a business estate in the corridor, and marked as either a company of transportation and warehousing (including e-commerce), or wholesale and import-export, were selected. Transportation and warehousing companies larger than 40,000 m² are labelled XXL distribution. The XL retail category was selected by retail company code in a business estate (excluding e-commerce). The resulting data includes ca. 10,000 buildings, ca. 4,000 of which are larger than 2,500 m². The period of analysis used for this study is 1996-2019. The vector data from these years is rasterised to 100m in order to model logistics development as a discrete land-use change process, in which each 100m grid cell represents a unit of observation. Multinomial logistic regression is applied in which the dependent variable has five categories which represent no change (the reference) and change to each of the four logistics categories described earlier. The explanatory variables employed in the model represent various location, accessibility and policy factors identified from prior theoretical and empirical knowledge on logistics location choice.

4. Results

Table 1 presents the results of multinomial logistic regression for estimating the effects of various spatial drivers on the development of four different types of logistics services. There is a positive effect of proximity to highway exits on the likelihood of development of all types of logistics. Highway accessibility has been shown to be an important factor for logistics location in previous studies as well (Bowen, 2008; Verhetsel et al., 2015). Proximity to urban area and customer/employee base also contributes positively to logistics development. However, proximity to train stations does not seem to be important for logistics development with the exception of XL retail. Presence within the highly urbanised Randstad region also lowers the likelihood of logistics growth perhaps due to lower availability of land. Higher land prices also contribute negatively to logistics growth. There seems to be a clear distinction between factors influencing the growth of XL retail type of logistics and other types since XL retail centres are expected to be closer to central city locations (as can also be inferred from a non-negative impact of the presence of urban amenities indicated by the urban attractiveness index). Presence of multi-modal terminals seems to positively impact the development of all logistics with the exception of retail. Finally, the presence within a ‘Top sector’ policy region increases the likelihood of development of transport and warehousing as well as XXL distribution centres. This is further verified when the same analysis is performed for both pre- and post-policy periods separately (results not shown here). There is a clear increase in the size and significance of the effect during the post-policy period (2006-2019) as compared to the pre-policy period (1996-2005).

	Dependent variable: Land-use change to logistics (1996-2019)			
	Transport & logistics	XXL distribution	Trade, import & export	XL retail
<i>Location</i>				
Hedonic land price, residential (2007 Euros)	-0.00270*** (0.000481)	-0.0132*** (0.00128)	-0.00189*** (0.000372)	-0.000273 (0.000385)
Urban attractiveness index	-12.35*** (1.688)	-16.55*** (4.777)	-5.390*** (0.876)	1.732* (0.887)
Ln (distance to urban area)	-0.212*** (0.00846)	-0.152*** (0.0175)	-0.249*** (0.00794)	-0.257*** (0.0167)
Within Randstad	-1.450*** (0.120)	-0.970*** (0.231)	-0.0365 (0.116)	-0.498** (0.245)
<i>Accessibility</i>				
Ln (distance to nearest highway access/exit)	-0.181*** (0.0100)	-0.273*** (0.0135)	-0.232*** (0.00896)	-0.310*** (0.0166)
Ln (distance to nearest train station)	0.00151 (0.0484)	0.0467 (0.0947)	-0.0851** (0.0425)	-0.322*** (0.0823)
Ln (travel time to nearest 100,000 inhabitants)	-2.177*** (0.108)	-2.469*** (0.187)	-2.051*** (0.0977)	-1.353*** (0.183)
Ln (distance to nearest multi-modal node)	-0.315*** (0.0436)	-0.726*** (0.0565)	-0.190*** (0.0401)	0.0299 (0.0886)
<i>Spatial Policy</i>				
Within <i>Logistics top sector</i> region	1.257*** (0.0856)	0.722*** (0.136)	0.147 (0.104)	-0.143 (0.235)
Constant	1.875*** (0.364)	3.640*** (0.575)	1.163*** (0.334)	-3.007*** (0.652)
Observations	808,188	808,188	808,188	808,188
Pseudo R ²	0.106	0.106	0.106	0.106

Table 1: Multinomial logit estimates of logistics land-use change from 1996-2019. Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

5. Conclusion

This study analyses the relative importance of various spatial drivers for four different types of logistics development in the Netherlands. Higher land prices discourage logistics growth in general. Local (highway) accessibility and regional accessibility are both important factors for promoting all types of logistics development. Proximity to multi-modal transport nodes is also critical for warehousing and distribution type of logistics, but not for XL retail centres. XL retail centres seem to have distinct location preferences as compared to other logistics types. The spatio-economic ‘Top sector’ policy of stimulating logistics growth in certain locations seems to have the desired positive effect on transport, warehousing and XXL distribution type of logistics.

References

- Aljohani, K. and Thompson, R.G. (2016). 'Impacts of logistics sprawl on the urban environment and logistics: Taxonomy and review of literature', *Journal of Transport Geography*, 57, pp. 255-263. DOI: 10.1016/j.jtrangeo.2016.08.009.
- Bowen, J.T. (2008). 'Moving places: the geography of warehousing in the US', *Journal of Transport Geography*, 16(6), pp. 379-387. DOI: 10.1016/j.jtrangeo.2008.03.001.
- Heitz, A., et al. (2018). 'Spatial patterns of logistics facilities in Gothenburg, Sweden', *Journal of Transport Geography*, 88, pp. 102191. DOI: 10.1016/j.jtrangeo.2018.03.005.
- Nefs, M. (2021). *Distribution centers 1980-2020 Geodata. 4TU.ResearchData*. Available at: <https://doi.org/10.4121/14791686>
- Onstein A.T.C., Tavasszy, L. and van Damme, D. (2019). 'Factors determining distribution structure decisions in logistics: a literature review and research agenda', *Transport Reviews*, 39(2), pp. 243-260. DOI: 10.1080/01441647.2018.1459929.
- van den Heuvel, F.P., de Langen, P.W., van Donselaar, K.H. and Fransoo, J.C. (2013). 'Spatial concentration and location dynamics in logistics: the case of a Dutch province', *Journal of Transport Geography*, 28, pp. 39-48. DOI: 10.1016/j.jtrangeo.2012.10.001.
- Verhetsel, A., et al. (2015). 'Location of logistics companies: a stated preference study to disentangle the impact of accessibility', *Journal of Transport Geography*, 42, pp. 110-121. DOI: 10.1016/j.jtrangeo.2014.12.002.
- Woudsma, C., Jensen, F.J., Kanaroglou, P. and Maoh, H. (2008). 'Logistics land use and the city: A spatial-temporal modeling approach', *Transportation Research Part E*, 44(2), pp. 277-297. DOI: 10.1016/j.tre.2007.07.006.

Special Session SS01

5 November 9:00-11:00 am (GMT)

Climate change mitigation strategies at urban and territorial scales: modelling tools and quantitative impact assessment

Proponents

Federico Amato, Antonino Marvuglia, Maider Llaguno, Beniamino Murgante, Federico Martellozzo, Benedetto Manganelli

Decrease of the value of the regulating ecosystem service of CO₂ capture by urban trees due to the increase of soil consumption: a quantification of the phenomenon in Rome

G.BENATI¹; F.MARTELLOZZO²

¹Department of MEMOTEF, Sapienza University of Rome

Via del Castro Larenziano 9, 00161 Rome, Italy

+3934893444588, giulia.benati@uniroma1.it (correspondent author)

² Department of Economics and Management , University of Florence

Via delle Pandette, 32, 50127 Florence, Italy

+39 342 99 77 690, federico.martellozzo@unifi.it

Keywords: Soil Consumption, Urban Ecosystem Services, Carbon Storage, Climate Regulation, Urban Trees

About 55% of world population lives in cities and this percentage is expected to grow until 66% within 2050 (UN, 2018). This increase in population is accompanied, as can be expected, by an increase in the built-up area of buildings and infrastructure, and therefore an increase in land consumption (Munafò et al., 2020). Not only that: in the last decades in Italy, the sealed soil has been growing at a much higher rate than population dynamics (Munafò et al., 2020) and, therefore, without being justified by it. Scientific evidence has moreover underlined how the soil consumption is one of the primary causes of the deterioration of the value of ecosystem services (Munafò et al., 2020) and of a consequent decrease in quality of human life, influencing these elements both at global and local scales. Therefore, one of the main challenges for public administrations is to plan urban spaces forecasting an ever-growing demographic pressure and soil consumption, while aiming to meet sufficient standards of environmental sustainability and quality of life for citizens.

The topic is in fact central to the European Agenda, which asks member states to reduce net soil consumption to zero by 2050, to decrease its growth rate in order to make it proportional to the population growth, and to align its trend with the SDGs of sustainable development of cities, ensuring

a satisfactory quality of life in city areas by 2030 (Goal 11) (UN, 2015). To this end, the European Commission has introduced several strategies to protect and enhance urban nature from degradation due to land use change from natural to artificial, emphasizing two main concepts: Green Infrastructures and Nature Based Solutions (Haase D. et al., 2017), both of which aim to enhance and protect urban green spaces and the ecosystem services they provide. In particular, trees in cities not only contribute to making the urban environment more livable for citizens but provide a wide range of regulating ecosystem services that affect importantly urban climate by mitigating the effects of human activities, such as air purification, temperature regulation, etc., and are therefore key elements in built-up areas. Among the various ESs provided by urban tree biomass, a significant body of scientific literature has highlighted its important role in compensating CO₂ emissions through carbon sequestration and storage, that positively affects urban air quality (Barò et al,2014), mitigating the effects of air pollution and protecting human health.

This paper aims to quantify the ecosystem service value of CO₂ retention provided by urban tree biomass and to evaluate the impact of land consumption on this service, taking as a case study the Municipality of Rome.

This area corresponds to the territory classified as Urban Core in the OECD system: a consistent and shared categorization of all the most densely populated European cities, that allows a potential comparison. With 2.779.973 inhabitants and an extension of 1287.36 km², Rome is the most populous and largest municipality in Italy and the third largest in the European Union, as well as the one with the largest amount of green areas (Source: Municipality of Rome). Rome is also the territory where soil consumption is higher than in all other Italian cities.

This research intends to elaborate and test, by means of a relevant case study, an operational framework that allows implementation and replicability in different scenarios, spatially and temporally. To this end as starting data we drew on different geographical information layers of the Copernicus project, developed by the European Environmental Agency and available throughout Europe. Furthermore, the

result was expressed in a unit of measurement that is easily communicable and can be compared with other phenomena, translating the loss of ES in monetary terms, as annual euros lost.

We therefore measured how much, in Rome and for each Municipality, soil consumption affects the Ecosystem Service of carbon sequestration and storage performed by urban trees, implementing the following methodology. I) First of all, we wanted to give an estimate of the quantity of carbon dioxide sequestered per tree on average for each Municipality, a quantity that depends on the number and species of trees. To this purpose, the census of trees in Rome was used (Source: Municipality of Rome). Among all tree species, those most relevant for carbon dioxide sequestration, measured in kg/year, have been selected. II) As the international CO₂ market estimates a value per ton, it is possible to calculate the unit economic value per tree: the value in euros of CO₂ kg absorbed per year per tree of each species. From this estimate and the number of trees per species per Municipality, the sought-after quantity has been derived: a unitary economic value “Average Tree”, meaning the annual absorption per average tree per Municipality. III) In order to obtain the value of this ES for the entire area of Rome, it was necessary to derive the extension of the area covered by trees for all Municipalities. We utilized the High-Resolution Layer of the European project Copernicus, Tree Cover Density, available for several years. The year 2012 was taken as a reference and calculated for each municipality the value of the ES in question for that year through appropriate processing. In doing so, we obtained the first result: the value, expressed in euros per year, of the ecosystem service of CO₂ retention provided by trees for each municipality. IV) We then wanted to estimate the area covered by trees lost due to land consumption between 2012 and 2018. To reach this goal, we used another shapefile, also developed within the European project Copernicus, representing the change of land use and land cover in urban areas: Urban Atlas Change 2012-2018. Within this, the classes of land cover change from natural to artificial were selected, and superimposed with the TCD layer, the area covered by trees lost between 2012 and 2018 for each municipality due to land consumption was found. Multiplying the area found by the Average Tree value obtained in the second point, through the necessary calculations

related to the area occupied by the arboreal crowns, we found the loss of the ES of CO₂ retention by trees in Rome, for each Municipality, due to land consumption between 2012 and 2018, measured in terms of euro per year. V) Finally, the same procedure was carried out in a forecasting perspective for 2030, utilizing a SLEUTH model for the increase in soil consumption (Martellozzo F., Amato F., Murgante B., 2017), a simulation that resulted in a further quantitatively substantial loss of ecosystem services. The analysis has shown that soil consumption has a quantitatively relevant impact on the loss of Urban Ecosystem Services and that, if not properly regulated, this loss will increase in the future. This is why it is of utmost importance to ensure integrated urban planning that incorporates European guidelines not only with regard to CO₂ emissions, but also on the factors that regulate their compensation, such as soil consumption, presence and distribution of urban green areas and the quality and quantity of tree species.

References

Baró, F., Chaparro, L., Gómez-Baggethun, E., Langemeyer, J., Nowak, D.J. and Terradas, J. (2014). Contribution of Ecosystem Services to Air Quality and Climate Change Mitigation Policies: The Case of Urban Forests in Barcelona, Spain. *AMBIO*, 43(4), pp.466–479

Haase D., Kabisch s., Haase a., Andersson E., Banzhaf E., Baró F., Brenck M., Fischer L.K., Frantzeskaki N., Kabisch N., Krellenberg K., Kremer P., Kronenberg J., Larondelle N., Mathey J., Pauleit S., Ring I., Rink D., Schwarz N., Wolff M., (2017). Greening cities – To be socially inclusive? About the alleged paradox of society and ecology in cities, *Habitat International*, Volume 64, Pages 41-48

Martellozzo, F., Amato, F., Murgante, B. and Clarke, K.C. (2018). Modelling the impact of urban growth on agriculture and natural land in Italy to 2030. *Applied Geography*, (91), pp.156–167

Munafò, M. (a cura di), (2020). Consumo di suolo, dinamiche territoriali e servizi ecosistemici. Edizione 2020. Report SNPA 15/20

United Nations, (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. A/RES/70/1

United Nations, Department of Economic and Social Affairs, Population Division (2018). *The World's Cities in 2018—Data Booklet (ST/ESA/ SER.A/417)*

Synergies and trade-offs in ecosystem services' provision: identifying spatial bundling in Sardinia.

Sabrina LAI

Department of Civil and Environmental Engineering, and Architecture – University of Cagliari

Via Marengo 2, 09123, Cagliari, Italy

0039 070 675 5201, sabrinalai@unica.it

Keywords: ecosystem services, multifunctionality, territorial specialization, bundling

As per the definition by the European Commission, a green infrastructure is “a strategically planned network of natural and semi-natural areas [...] designed and managed to deliver a wide range of ecosystem services”. From a natural scientist’s standpoint, this definition is problematic because human-centred, as residual naturalness in Europe is here thought of as something to be managed and planned to maintain ecosystem services (ES), hence the goods and services provided by nature to humans. Moreover, “a wide range” implies that several ES can be delivered simultaneously; while some ES are often jointly provided and intertwined (Bennet et al., 2009), several studies show that some interrelationships are negative and result in trade-offs (Madureira and Andresen, 2014), especially as regards provisioning ES.

Therefore, by building on spatially explicit assessments of seven ES carried out in previous studies (Lai and Leone, 2017; Floris, 2020; Lai et al., 2020; 2021), this piece of work aims at investigating multifunctionality, i.e., an area’s capacity to provide simultaneously multiple ES, to identify areas of territorial specialization on which planning policies can be grounded. Sardinia, an Italian island, is chosen as a case study because its low residential density, low endowment of infrastructure, and persistence of traditional agricultural and farming practices have preserved a good level of naturalness, which is a prerequisite for the delivery of ES. The scale is that of municipalities, which, in Italy, represent the lowest administrative tier in charge of land-use planning, as land-use and land-cover changes are the most important factors affecting ES provision.

The seven ES here considered are as follows: habitat quality as potential nursery for species (QUHAB); global climate regulation (carbon sequestration and storage: CO2SEQ); local climate regulation (mitigation of land surface temperature REGTEMP); agricultural and forestry productivity (PRODAF); ecosystem-based potential recreation (POTRIC); intrinsic value of biodiversity for present and future generations (INTBIO); landscape quality, reflecting cultural identity and sense of place (QUPAES).

The seven ES were assessed and mapped, and their values normalized in the 0÷1 range; next, the mean normalized value was calculated for each Sardinian municipality (n=377). Following Raudsepp-Hearne et al. (2009), Turner et al. (2014), Queiroz et al. (2015), the analysis comprised three steps:

- for each ES, spatial patterns were analysed by assessing spatial autocorrelation through Moran's I index; hotspots and coldspots of municipalities having statistically significant higher or lower values than their surrounding municipalities were identified through the Getis-Ord G_i^* statistics;
- statistically significant (linear) correlations between each pair of ES were assessed through Pearson's correlation coefficient;
- clusters comprising municipalities having a level of in-group similarity higher than their dissimilarities with respect to municipality not belonging to the group were spatially identified by applying first a principal component analysis (PCA), aimed at reducing redundancies due to the present of correlations, and next a cluster analysis through the k-means algorithm.

Except for the autocorrelation analysis, which was performed in GeoDa, version 1.6.7, all the other steps were performed using ArcGIS®ESRI, version 10.7.

Figure 1 provides an overview of the assessment of each ES at the municipal level; the autocorrelation analysis, carried out with the queen contiguity conceptualization, shows evidence of spatial agglomeration ($I > 0$) statistically significant and decreasing up to the fifth level of contiguity for all of the ES but for QUHAB, for which only the first two levels show evidence of autocorrelation. The results of the hotspot and coldspot analysis are shown in Figure 2, left-hand side, and highlight that some parts of the island can be either hotspots or coldspots, depending on the ES at stake.

The Pearson coefficients (here not reported for the sake of space) show that 16 out of the 21 ES pairs are significant linearly correlated; negative correlations always concern PRODAF and CO2SEQ, which therefore compete with other ES.

Since the seven ES are mutually correlated, the territory can be conceived of as a provider of bundles of ES, and the PCA is helpful to reduce redundancies. The PCA shows that approximately 93% of the variance can be explained through four new axes, or around 97% through five axes, which represent as many combinations of the seven ES. Therefore, the k-means algorithm was run seeking out for five clusters, whose spatial distribution is shown in Figure 2, right-hand side, together with the spider diagrams that provide the mean value of each ES in each group.

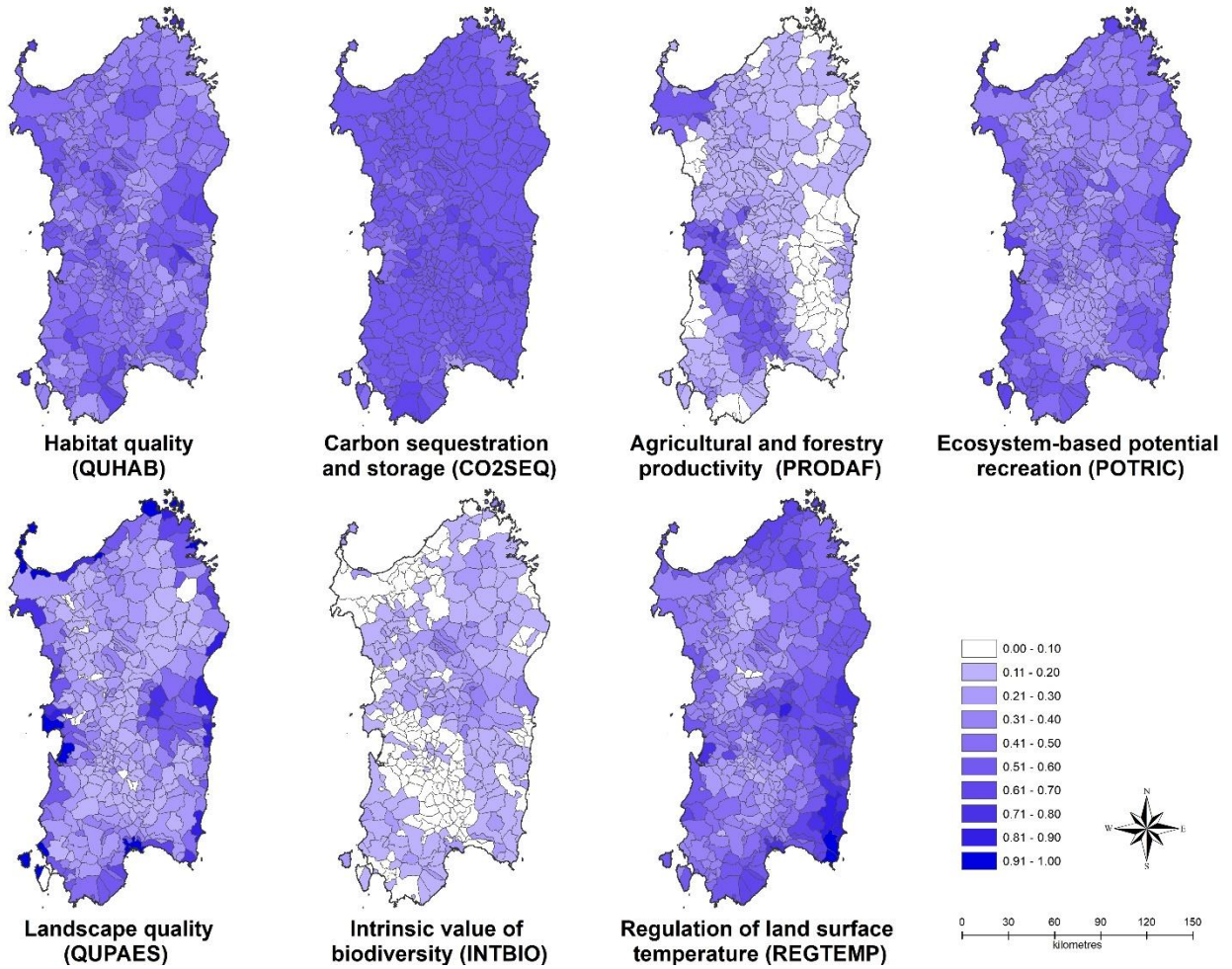


Figure 1: Mean ES normalized values at the municipal level.

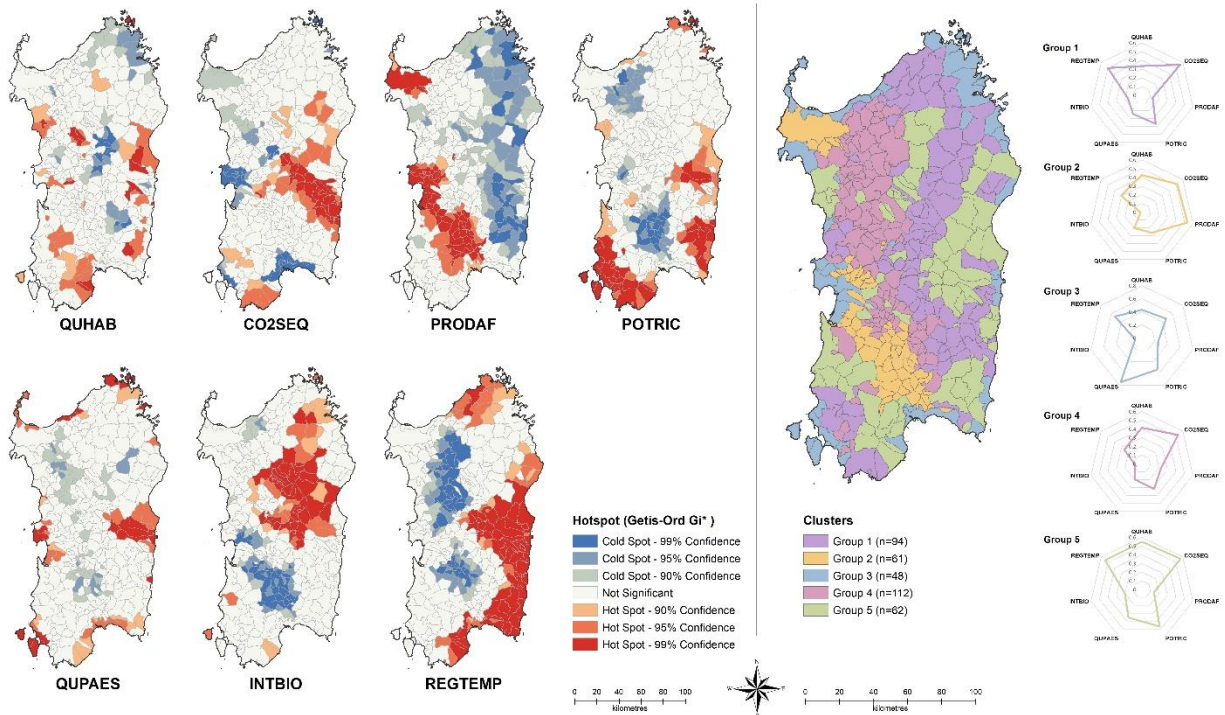


Figure 2: Hotspots and coldspots of mean ES values, municipal level (left), and results from the cluster analysis (right)

As Figure 2 shows, five municipal cluster can be identified; municipalities in each cluster share common features in terms of bundles of ES provided, and they also share some distinctive environmental and socio-economic characteristics. Group 3, where QUPAES and POTRIC dominate, comprises almost exclusively coastal municipalities, whose economies rely on tourism and whose urbanization levels are generally higher than the rest of the island's ones. Group 2, having high values of both PRODAP and CO2SEQ, corresponds to the island's main plains, hosting intensive agriculture and farming that yield comparatively high incomes with respect to Sardinian standards. Groups 1 and 4, both showing high values of CO2SEQ, differ as regards PRODAP (low in 1 and high in 4) and REGTEMP (high in 4 and low in 1); while they both comprise inner and sparsely populated municipalities, they differ as for the morphology, which is gentler in group 4, and the vegetation, which in group 4 is richer in steppe and other herbaceous vegetation and pastures. Finally, group 5 comprises municipalities having high values (in the upper quartile) of all ES except PRODAP; it includes mountain areas, whose landscapes are marked by maquis and forests, the greatest providers of ES in Sardinia.

In conclusion, the analysis of bundles of ES provision at the municipal levels unveils spatial patterns reflecting both ecological and socio-economic patterns. A case can therefore be made for the existence of a territorial specialization (Queiroz et al., 2015) that should be accounted for within land use plans, whose actions can either enhance or degrade the provision of some ES. Awareness should be raised on the fact that actions aiming at enhancing some ES (first and foremost, agricultural productivity) are often detrimental to the maintenance of other ES, such as regulation and cultural ones (Martín-López et al., 2012), and such trade-offs need careful ex-ante assessments in plan-making processes. Future research directions point to the inclusion of a larger number of ES, through which the sensitivity and stability of the clusters could be better assessed, and the inclusion of socio-economic control variables.

Acknowledgments

This study has been carried out in the framework of the Research projects “Paesaggi rurali della Sardegna: pianificazione di infrastrutture verdi e blu e di reti territoriali complesse” (funded by the Autonomous Region of Sardinia, 2017 call for “Basic research projects” as per regional law 7/2007), and “Investigating the relationships between knowledge-building and design and decision-making in spatial planning with geodesign” (funded by Fondazione di Sardegna, 2018 call).

References

Bennett, E.M., Peterson, G.D. and Gordon, L.J. (2009). ‘Understanding relationships among multiple ecosystem services’, *Ecology Letters*, 12, pp. 1394–1404.

- Floris, M. (2020). *I servizi ecosistemici nella pianificazione spaziale come strumenti interpretativi per la definizione di tassonomie territoriali innovative*. PhD University of Cagliari.
- Lai, S. and Leone, F. (2017). 'A methodological approach to identify a multifunctional green infrastructure at the regional scale. A case study from Sardinia, Italy', *Urbanistica Informazioni*, 272 s.i., pp. 836–840.
- Lai, S., Leone, F. and Zoppi, C. (2020). 'Spatial distribution of surface temperature and land cover: A study concerning Sardinia, Italy', *Sustainability*, 12, art. no. 3186.
- Lai, S., Isola, F., Leone, F. and Zoppi, C. (2021). 'Assessing the potential of green infrastructure to mitigate hydro-geological hazard', *TeMA - Journal of Land Use, Mobility and Environment*, s.i. 1/2021, pp. 109–133.
- Madureira, H. and Andresen, T. (2014). 'Planning for multifunctional urban green infrastructures: Promises and challenges', *Urban Design International*, 19, pp. 38–49.
- Martín-López, B., Iniesta-Arandia, I., García-Llorente, M., Palomo, I., Casado-Arzuaga, I., García Del Amo, G., Gómez-Baggethun, E., Oteros-Rozas, E., Palacios-Agundez, I., Willaarts, B., González, J.A., Santos-Martín, F., Onaindia, M., López-Santiago, C. and Montes, C. (2012). 'Uncovering ecosystem service bundles through social preferences', *Plos One*, 7(6), art. no. e38970.
- Queiroz, C., Meacham, M., Richter, K., Norström, A.V., Andersson, E., Norberg J. and Peterson, G. (2015). 'Mapping bundles of ecosystem services reveals distinct types of multifunctionality within a Swedish landscape', *AMBIO*, 44, pp. 89–101.
- Raudsepp-Hearne, C., Peterson, G.D. and Bennett, E.M. (2010). 'Ecosystem service bundles for analyzing tradeoffs in diverse landscapes', *Proceedings of the National Academy of Sciences*, 107(11), pp. 5242–5247.
- Turner, K.G., Odgaard, M.V., Bøcher, P.K., Dalgaard, T., Svenning, J.-C. (2014). 'Bundling ecosystem services in Denmark: Trade-offs and synergies in a cultural landscape', *Landscape and Urban Planning*, 125, pp. 89–104.

The Urban Carbon Budget (UCB) Model: a high resolution spatio-temporal model of CO₂ emissions, dispersion and sequestration in cities

Marlène BOURA¹; Geoffrey CARUSO^{1,2}

¹University of Luxembourg

11 Porte des Sciences, 4366, Luxembourg

marlene.boura@uni.lu (correspondent author)

²LISER

11 Porte des Sciences, 4366, Luxembourg

geoffrey.caruso@uni.lu

Keywords: carbon balance, carbon sequestration, CO₂ emissions, modelling, urban areas

Submission to SS01

1. Background

The ongoing climate change is a global phenomenon induced by human activities (IPCC, 2021). It results from the accumulation of greenhouse gases (GHG) in the atmosphere, mainly since the industrial revolution (IPCC, 2014, 2019, 2021). Carbon dioxide (CO₂) is one of the major GHG, as highlighted by the Kyoto Protocol (United Nations, 1998). It is also the most important GHG in the atmosphere in terms of proportion (IPCC, 2014). Before the pre-industrial era, the global average concentration of CO₂ was about 280 ppm (Lindsay, 2020). The concentration has increased by 48% between 1750 and February 2021 to reach 415.88 ppm in February 2021 (NOAA, 2021).

Anthropogenic CO₂ emissions are mainly released into the atmosphere through the combustion of fossil fuels to produce energy (IPCC, 2014). Since the 1990s, the global demand for energy and energy-related CO₂ emissions have increased steadily (IEA, 2021). With the process of urbanisation, human

settlements are agglomerating but still represent a tiny proportion of the emerged land (Liu *et al.*, 2014). Yet the IEA (2008) projects that cities will be responsible for about three quarters of global CO₂ emissions by 2030. It has therefore become clear that cities must drastically reduce their emissions. More and more cities and metropolitan areas are now taking action around the world (see C40 Cities, Covenant of Mayors or Under2 Coalition).

Previous large-scale studies focussed on regional to global scales (Hutyra *et al.*, 2014) such that spatial and temporal resolutions were rarely meaningful to local or urban areas (Super, 2018). Therefore, the urban carbon cycle must be properly assessed as it impacts regional and global carbon cycles (Hutyra *et al.*, 2014).

We propose a method to obtain disaggregated estimates for CO₂ emissions, sequestration and balance. The method is built on three types of data: land use, sectoral emissions and sequestration parameters. In addition, factors of temporal decomposition are applied. It aims at being easily replicable for a large number of areas (for instance cities) at one or multiple time periods. Estimates are available at a high spatial resolution (1 ha) and high temporal resolutions (year, month and weekday). So far, the alternatives are global and regional transport models, eddy-covariance measurements, urban carbon metabolism, remote sensing or field measurements. Yet they are not appropriate for applications on multiple large urban areas for various reasons.

2. Method

The High Resolution Urban Carbon Balance (UCB) model allows to downscale CO₂ emissions, simulate the dispersion of the emissions and their sequestration. From these outputs, the aggregated and disaggregated urban carbon balance are then assessed using different metrics. We developed estimates of anthropogenic carbon emissions and natural carbon sequestration at a spatial resolution of 1 ha and with annual, monthly and daily temporal resolutions. The carbon balance is then expressed in two ways: in absolute terms (i.e. total amount sequestered over a day) and in relative terms (i.e.

percentage of emissions sequestered over a day). Therefore, the uniqueness of this work comes from multiple aspects:

1. The sectoral CO₂ emissions were downscaled spatially using a high resolution urban land use database for hundreds of urban areas,
2. The seasonality of CO₂ emissions is accounted across country-, sector-, and year-specific time profiles,
3. The urban carbon sequestration capacity was estimated at a high spatial resolution for hundreds of urban areas,
4. The seasonality of the sequestration process is accounted with location-specific monthly factors,
5. The carbon balance estimates allow comparison of a non-spatial model with a spatially explicit model (with basic spatial dispersion settings).

The method can be easily replicated at different locations or time periods. We showcase the UCB model on four cities which represent different degrees of integration of LU. The method is built on three types of data: LU, localised emissions, and factors of decompositions and sequestration parameters.

This approach accounts for (i) the temporal distribution of annual emissions towards monthly and daily emissions, (ii) the temporal distribution of annual sequestration towards monthly and daily sequestration and (iii) the spatial dispersion of emissions. The seasonality in CO₂ emissions and sequestration is taken into consideration by applying a spatial dispersion to one typical day of each month of the year for each urban area. For a weekday, we simulate the spatial dispersion of the emitted CO₂ molecules, from the cells where they are emitted to the surrounding locations. Then, the molecules reaching a cell with a sequestration potential will be removed according to the capacity of the corresponding cell. In this approach, we consider the study area as a closed system: only CO₂

molecules emitted within the system can be sequestered. The budget here corresponds to the difference between the CO₂ emissions after spatial dispersion and removal by vegetation. It is important to note that the budget cannot be negative. The removal corresponds to the amount of anthropogenic emissions the urban forest can actually sequester after spatial dispersion, over a given time period. When expressed in absolute terms, values are in t CO₂ per day. When expressed in relative terms, they correspond to the percentage of emissions which has been sequestered.

3. Conclusion

The High Resolution Urban Carbon Balance (UCB) model allowed to create the first European high resolution CO₂ emissions database for urban areas. The method to do so considers sectoral emissions, applies spatial downscaling based on land use data, and temporal downscaling based on sectors, year and country. The model allowed to create the first high resolution database on sequestration capacity for European urban areas. Additionally, the model proposes to simulate the spatial dispersion of the anthropogenic emissions and their sequestration, for the time being, in a simplistic way. The UCB model can be replicated for other locations or time periods. Moreover, the emission part is not limited to CO₂ emissions and could be used to downscale other gases and air pollutants from a LU dataset and emission or concentration data only.

The resulting carbon balance provide a strong indication on the contribution of each urban areas to the carbon cycles at the regional and global scales.

References

Hutyra, L. R. *et al.* (2014) ‘Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective’, *Earth’s Future*, 2(10), pp. 473–495. doi: 10.1002/2014EF000255.

IEA (2008) *World Energy Outlook 2008*, IEA Publications. Paris.

IEA (2021) *Global Energy Review 2021. Assessing the effects of economic recoveries on global energy*

demand and CO2 emissions in 2021. Paris.

IPCC (2014) ‘Summary for Policymakers and Technical Summary’, in Edenhofer, O. et al. (eds) *Climate Change 2014: Mitigation of Climate Change. Part of the Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, pp. 1–30. doi: 10.1017/CBO9781107415416.005.

IPCC (2019) ‘Summary for Policymakers’, in Masson-Delmotte, V. et al. (eds) *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change*. Cambridge, pp. 1–24. Available at: https://www.ipcc.ch/site/assets/uploads/sites/2/2019/05/SR15_SPM_version_report_LR.pdf.

IPCC (2021) ‘Summary for Policymakers’, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. SPM1–SPM41.

Lindsay, R. (2020) *Climate Change: Atmospheric Carbon Dioxide*, NOAA. Available at: <https://www.climate.gov/news-features/understanding-climate/ climate-change-atmospheric-carbon-dioxide>.

Liu, Z. et al. (2014) ‘How much of the world’s land has been urbanized, really? A hierarchical framework for avoiding confusion’, *Landscape Ecology*, 29(5), pp. 763–771. doi: 10.1007/s10980-014-0034-y.

NOAA (2021) *Trends in Atmospheric Carbon Dioxide. Global Monthly Mean CO2*, Global Monitoring Laboratory.

Super, I. (2018) *Quantification and attribution of urban fossil fuel emissions through atmospheric measurements*. Wageningen University. doi: 10.18174/457839.

United Nations (1998) *Kyoto Protocol to the United Nations Framework -- Convention on Climate*

Change, Review of European Community and International Environmental Law. doi: 10.1111/1467-9388.00150.

European urban cores under pressure: quantifying the congestion of trips into and out from city centres as a function of population size

Estelle MENNICKEN^{1,2}; Geoffrey CARUSO^{1,2}; Philippe GERBER¹; Alex HAGEN-ZANKER³;
Rémi LEMOY⁴

¹Luxembourg Institute of Socio-Economic Research (LISER)
11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg
+352 58 58 55 957, estelle.mennicken@liser.lu (correspondent author)
+352 58 58 55 601, philippe.gerber@liser.lu

²University of Luxembourg
11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg
+325 46 66 44 6625, geoffrey.caruso@uni.lu

³University of Surrey
Department of Civil and Environmental Engineering
Guildford, Surrey, GU2 7XH, UK
+44 (0)1483 682440, a.hagen-zanker@surrey.ac.uk

⁴University of Rouen-Normandie
Département de géographie, aménagement, environnement
1, rue Lavoisier, 76821 Mont Saint Aignan Cedex, France
remi.lemoy@univ-rouen.fr

Keywords: road network, travel times, urban scaling, city size

1. Introduction

The present study is part of a larger endeavour to formally and empirically establish the relationship between urban form and road network characteristics, and more specifically aimed at understanding how intra-urban travel times change with population size and vary across space. Understanding and quantifying traffic congestion at the scale of an entire continent is a serious societal challenge, as it is associated to several social, environmental and economic consequences. The negative effects are mainly the loss of time (hours of delay) inducing productivity and well-being losses, the excess fuel consumption, and the excess emitted CO₂. These unfavourable consequences have to be balanced with agglomeration benefits, such as shorter trips and a reduced car dependency. Economically, recent OECD studies have shown that larger cities in terms of population are generally more productive. A doubling population size corresponds to a 2 to 5% of productivity increase (OECD, 2015).

Specifically to our research interest, scaling relationships between population size of urban areas and their average transport properties (such as the road network length, the total traveled distance, or the congestion) have already been established to find out whether larger cities show disproportionately bigger outputs (e.g. Bettencourt, 2013; Louf, 2014; Chang et al. 2017). However, different studies show contradicting results depending on the cities database and on the measures of urban congestion. The fact that congestion and pollution levels can strongly differ across cities of similar size indicates that they are mostly driven by urban form, transport infrastructure, and more generally by policy choices, either at the metropolitan or at the national level.

The aim of this research piece is to study travel times to and from the main centre of cities to build a novel database across Europe, highlighting the spatial and temporal distribution of congestion on road networks. This radial perspective is useful both methodologically, as it allows for an effective simplification of cities, and empirically as it allows for a reflection upon city planning instruments targeting accessibility inequalities or attempting to counteract congestion. Then, we aim to understand the relation between these identified congestion levels and city size. More specifically, we analyse the data to identify allometric relationships both serving as a generalisation of the internal morphology and the road network structures of cities, and as an indicator of their over- or under-expansion, given their size.

2. Data and methods

2.1 Spatial and temporal data sampling

In this study, urban form is described in accordance with the mono-centric city model (Alonso, 1964; Muth, 1969; Mills, 1972). We use and combine two data sources: the 5 meters resolution land use maps of 303 functional urban areas (FUAs) of the European Urban Atlas (Copernicus, EEA, 2006), and the 1 km² resolution Eurostat population grid (Eurostat, 2006). We extracted the populated urban fabric cells from the land use maps, and randomly sampled at least one cell (k) per 1km² grid (j) to serve as origin or destination of trips, so that we cover the entire city areas. The historical city centers, as the main attractors of trips, are approximated by the location of the city halls. Total city populations N_c are retrieved from the 1km² population grid. The sampling process is illustrated for Namur as an example on figure 1 below.



Figure1: Example of spatial data structure and sample for Namur (BE).

We use the Google Directions API with best guess traffic conditions to model the fastest automobile trips from residential land use cells to the city centers, for every arrival hour between 4 am and 11 am during a typical weekday (Tuesday). These are inbound trips (*i*). We compute reverse outbound trips (*o*) from the city centers to the urban fabric locations, for hourly departure times between 12 pm and 3 am. This procedure resulted in more than 7.5 million simulated trips in total. The structure of the computed travel times is represented for one trip in one city on figure 2 below.

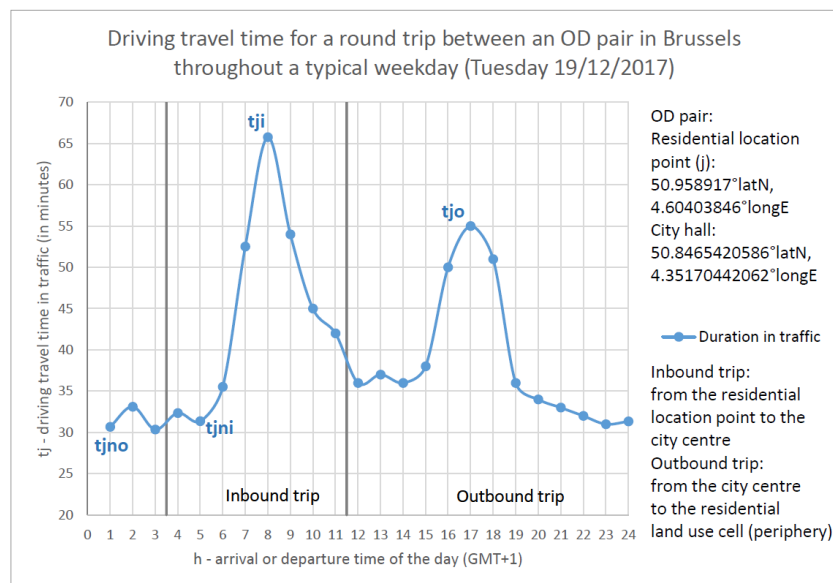


Figure 2: Insight of the travel time database for one trip in Brussels (BE).

2.2 Analytical steps

Our analysis consists in computing four indices in order to study congestion levels in the light of city size, city area, population density, and road network structure. We always relate the maximal trip duration ($t_{ji/o}$) to the minimal trip duration ($t_{jni/o}$), for inbound and outbound trips. The first refers to the peak traffic time and the latter to the free flow time for the corresponding directional trip. We define this difference ($t_{ji/o} - t_{jn}$) as the excess travel time.

The first absolute index A is the sum of the excess travel time over the entire city area, that is informative for city planning matters. The two following indices include a spatial dimension. The B index is the average excess travel time per trip, on all lengths of trips. It thus accounts for the spatial distribution of urban cells inside cities. The C index the percentage of trip lengthening, it relates the excess travel time to the travel time in free flow. The fourth and last index D is the time lost per capita (assuming the entire population is travelling). It is the sum of the excess travel time multiplied by the population at origin/destination, then divided by the total city population (figure 3).

$$\begin{aligned}
 A_i &= \sum_j^J t_{ji} - t_{jn} \quad (\forall j \in c, \text{in min}) & B_i &= \frac{1}{J} \sum_j^J t_{ji} - t_{jn} \quad (\forall j \in c, \text{in min}) \\
 A_o &= \sum_j^J t_{jo} - t_{jn} \quad (\forall j \in c, \text{in min}) & B_o &= \frac{1}{J} \sum_j^J t_{jo} - t_{jn} \quad (\forall j \in c, \text{in min}) \\
 C_i &= \frac{1}{J} \sum_j^J \frac{t_{ji} - t_{jn}}{t_{jn}} * 100 \quad (\forall j \in c, \text{in \%}) & D_i &= \frac{\sum_j^J (t_{ji} - t_{jn}) N_j}{N_e} \quad (\forall j \in c, \text{in min/capita}) \\
 C_o &= \frac{1}{J} \sum_j^J \frac{t_{jo} - t_{jn}}{t_{jn}} * 100 \quad (\forall j \in c, \text{in \%}) & D_o &= \frac{\sum_j^J (t_{jo} - t_{jn}) N_j}{N_e} \quad (\forall j \in c, \text{in min/capita})
 \end{aligned}$$

Figure 3: Equations of the four congestion indices

Finally, we relate the congestion indices to the total city population in order to reveal potential scaling relationships. We study the variation of our urban quantity, i.e. the congestion, against city population in the mathematical form of a power law.

3. Results and conclusion

The travel times in traffic conditions allow us to observe the magnitude of congestion throughout the day. By identifying the most frequent inbound and outbound traffic peaks for each city, we reveal a

North-South spatial heterogeneity. Then we provide a congestion database with four aggregate indices for 303 cities in Europe and establish a city ranking. Average time losses tend to be smaller for return trips to the periphery than for trips to the city centre, showing that traffic is more concentrated at one specific time in the morning towards the centre, while it is more distributed in the evening back to the periphery due to activities after work. The organisation of roads such as one-way streets and traffic lights is designed to faster leave an urban centre than to enter it.

On the city ranking, we globally observe an important influence of city population on the indices. Even though the A index scales super-linearly with the population, indicating a proportionally bigger aggregate time loss, the D index of time loss per capita shows a sub-linear scaling exponent (see figure 4). This suggests that bigger cities have a better spatial organisation of activities in relation to their road network structure.

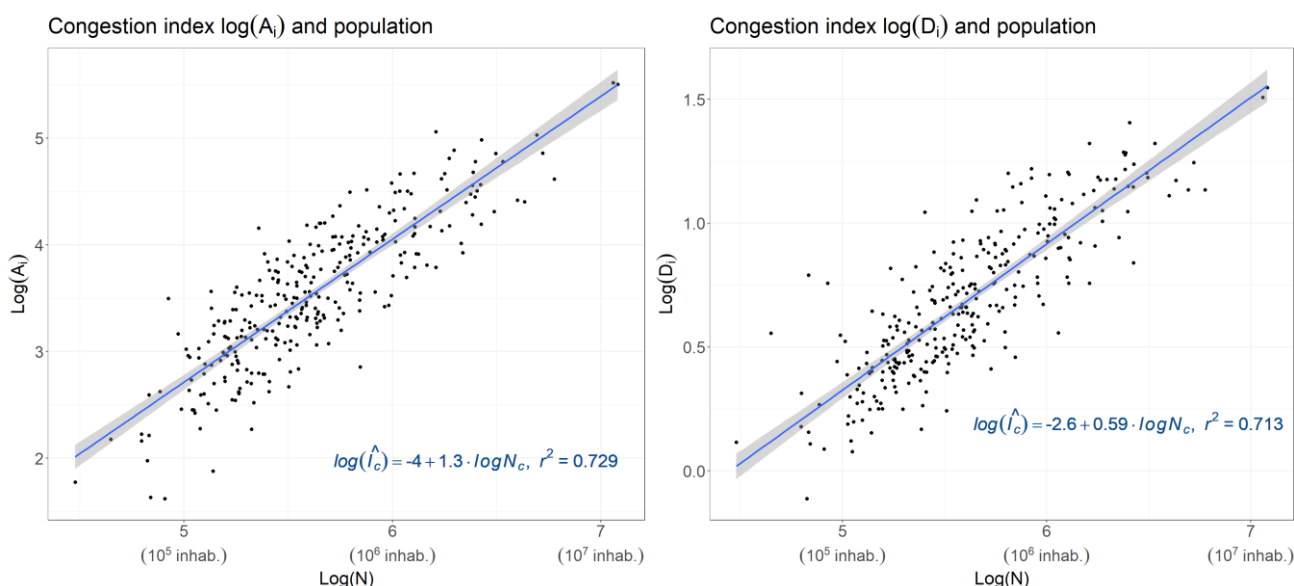


Figure 4: Influence of city size on congestion indices

Then, we see that city size is not the only explanatory factor. Cities from the urbanisation corridor from North England to North Italy are more congested than expected (based on their population), as suggested by the positive regression residuals. We also discuss outliers, such as cities that have relatively low congestion scores over all indices, yet with an average population size (Ostrava or Zaragoza for example), or cities with low values of total or average congestion but high relative trip lengthening (high C index for Belfast for example), or on the contrary high scores overall except for one index. Individual case studies can be examined more in depth, based on specific interests.

References

Alonso, W. (1964). *Location and Land Use*. Harvard University Press.

OECD (2015). *The Metropolitan Century: Understanding urbanization and its consequences*. OECD Publishing, Paris.

Bettencourt, L. M. (2013). The origins of scaling in cities. *Science*, 340(6139), 1438-1441.

Chang, Y. S., Lee, Y. J., and Choi, S. S. B. (2017). Is there more traffic congestion in larger cities?-scaling analysis of the 101 largest us urban centers. *Transport Policy*, 59:54-63.

Copernicus, EEA (2006). *Urban Atlas*. Available at: <http://land.copernicus.eu/local/urban-atlas/urban-atlas-2006>. (Accessed 2019-11-04).

Eurostat (2006). Geostat population grids. Available at: <http://ec.europa.eu/eurostat/fr/web/gisco/geodata/reference-data/population-distribution-demography/geostat>. (Accessed 2019-11-04).

Louf, R. and Barthelemy, M. (2014). Scaling: lost in the smog. *Environment and Planning B: Planning and Design*, 41(5):767-769.

Mills, E. (1972) *Studies in the Structure of the Urban Economy*. The Johns Hopkins Press.

Muth, R.F. (1969). *Cities and housing : the spatial pattern of urban residential land use*. University of Chicago press.

Centrality and city size effects on NO₂ ground and tropospheric concentrations within European cities

Yufei WEI¹; Geoffrey CARUSO^{1,2}; Rémi LEMOY³

¹University of Luxembourg

11 Porte des Sciences, 4366, Luxembourg

(+352) 46 66 44 9347, yufei.wei@uni.lu (correspondent author)

²Luxembourg Institute of Socio-Economic Research

11 Porte des Sciences, 4366, Luxembourg

(+352) 46 66 44 6625, geoffrey.caruso@uni.lu

³University of Rouen

7 Rue Thomas Becket, 76130, France

(+33) 02 35 14 6000, remi.lemoy@univ-rouen.fr

Keywords: NO₂, population, centrality, Sentinel-5P, scaling

1. Background and objectives

As urban population grow worldwide (UN, 2019), exposure to nitrogen dioxide (NO₂) is an increasing concern. In Europe, NO₂ caused 71,000 premature deaths in 41 countries in 2013 (EEA, 2016).

The concentration of NO₂ in a city can vary evidently (Restrepo, 2021), and hence suitable for marking the distribution of air pollution within cities. Some individual studies indicate NO₂ concentrations in city centres are higher than that in the periphery (Kirby et al., 1998). Comparative studies however are lacking that would consider a large set of cities to derive a more general quantification of centrality effects, independent of local landscape heterogeneity. Theory about how air pollution profiles relates to internal urban structure is also limited, but Schindler et al. (2017) show traffic-induced NO₂ levels decrease with the distance to the city centre. Empirical enquiries are therefore needed to quantify distance gradients and NO₂ levels at the periphery and centre, especially for different population size. Surprisingly, only a few studies (e.g. Bechle et al., 2011) indicate how NO₂ aggregates over a few cities within the top ranks of some specific world regions or globally, relate to population size. From a systematic review we estimated a 10-fold increase in population leads to a 1.45-fold increase in average concentrations (Wei et al., 2019). Nevertheless, there is still no quantitative estimates of the

effects of city size on NO₂ concentrations across cities of an entire continent, and no idea of whether population size impacts the distribution of pollution within cities, especially along centrality.

We aim to derive quantitative relations of NO₂ concentrations with both city size (expressed in population terms) and distance to the centre across major European cities. We compare estimates for the two types of data that are usually considered in the literature, respectively ground-level NO₂ from monitoring stations and remotely sensed tropospheric NO₂.

Eventually we aim to inform urban planners how to best organise the population across and within cities for reducing hazardous air pollutants. Following the studies of Lemoy and Caruso (2020), we link intra-urban structures with city size (scaling) effects to promote integrated urban science (Batty, 2013) towards urban sustainability.

2. Data

We include 378 Functional Urban Areas (FUAs) of 33 European countries. We use annual mean NO₂ surface concentrations in 2018 from monitoring stations and daily tropospheric NO₂ vertical columns from 25.10.2018 to 24.10.2019 by Sentinel-5P. We use FUA population data in 2015 to assess population size. Centrality effects are computed from the Euclidean distance from each station/cell to the city hall.

3. Methods

Based on the previous meta-analysis where the relation between the annual mean NO₂ surface concentrations and population size was found to be of the log-log form (Wei et al., 2019), we regress the logarithmic annual mean NO₂ surface concentrations and the logarithmic annual mean tropospheric NO₂ columns against the logarithmic distance to city centres and the logarithmic population size respectively and together.

We use dummy variables showing the fixed effects from monitoring stations of different background when regressing logarithmic annual mean NO₂ surface concentrations. Inspired by previous studies (Nicholas Hewitt, 1991; Schindler et al., 2017), we include the logarithmic minimum annual mean

tropospheric NO₂ vertical column of each FUA as a background factor when regressing logarithmic annual mean tropospheric NO₂ vertical columns.

4. Results

4.1 Regressions of Logarithmic Annual Mean NO₂ Surface Concentrations

As expected, logarithmic distance from monitoring stations to the city centre negatively impacts the logarithmic annual mean NO₂ surface concentrations, and larger cities have higher levels of logarithmic annual mean NO₂ surface concentrations.

The full specification of the model consisting of logarithmic distance, logarithmic population, and the dummy variables for background explain 54% of the logarithmic annual mean NO₂ surface concentrations. The monitoring stations background is the most influential factor, followed by logarithmic population size. Logarithmic distance is the least important factor.

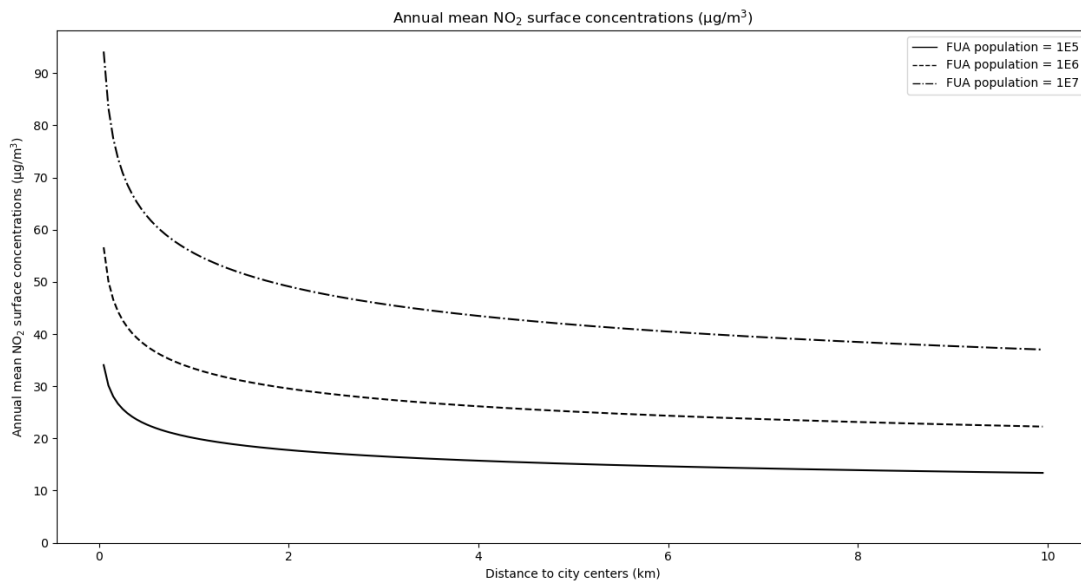


Figure 1: Effects plot of annual mean NO₂ surface concentrations

Figure 1 is the effects plot of our estimates. We see that when the distance to the centre increases from 100m to 1km, for a city with 100-thousand inhabitants, its annual mean NO₂ surface concentration decreases by 10 µg/m³; for a city with 1 million inhabitants, its annual mean NO₂ surface concentration decreases by 17 µg/m³; for a city with 10-million inhabitants, its annual mean NO₂ surface concentration decreases by 28 µg/m³. When the distance to city centres increase from 1km to 10km,

for a city with 100-thousand inhabitants, its annual mean NO₂ surface concentration decreases by 7 μg/m³; for a city with 1 million inhabitants, its annual mean NO₂ surface concentration decreases by 11 μg/m³; for a city with 10-million inhabitants, its annual mean NO₂ surface concentration decreases by 19 μg/m³. The gradient decays sharply near the centre. Then it decreases with the distance to the city centre, and it tends to become an asymptote at around 9km from the city centre.

4.2 Regressions of Logarithmic Annual Mean Tropospheric NO₂ Vertical Columns

Similar results are obtained for tropospheric NO₂. We find logarithmic distance to the city centre negatively impacts the logarithmic annual mean tropospheric NO₂ vertical columns, and that larger cities have higher levels of logarithmic annual mean tropospheric NO₂ vertical columns.

Logarithmic distance and logarithmic population together explain nearly 30% of annual mean tropospheric NO₂ columns concentrations. Adding the logarithm of the minimum column NO₂ for each FUA improves explanatory ability evidently (reaching 87%). It is also the most important factor in deciding logarithmic annual mean tropospheric NO₂, followed by logarithmic population size. Logarithmic distance is again the least important factor.

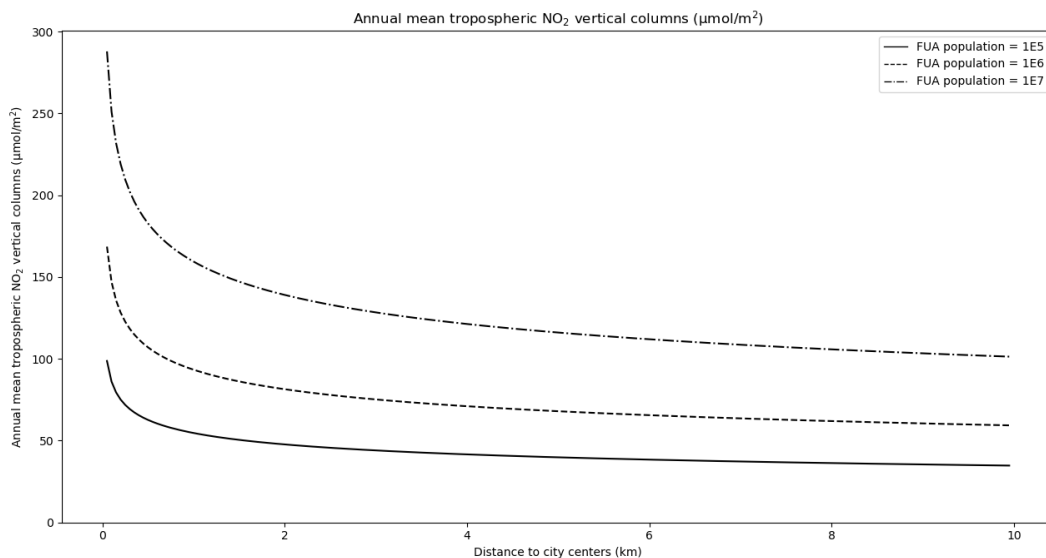


Figure 2: Effects plot of annual mean tropospheric NO₂ vertical columns

Figure 2 is the effects plot of our estimates. We show that when the distance to the city centre increases from 100m to 1km, for a city with 100-thousand inhabitants, its annual mean tropospheric NO₂ column

decreases by 31 $\mu\text{mol}/\text{m}^2$; for a city with 1 million inhabitants, its annual mean tropospheric NO₂ column decreases by 54 $\mu\text{mol}/\text{m}^2$; for a city with 10-million inhabitants, its annual mean tropospheric NO₂ column decreases by 92 $\mu\text{mol}/\text{m}^2$. When the distance to city centres increase from 1km to 10km, for a city with 100-thousand inhabitants, its annual mean tropospheric NO₂ column decreases by 20 $\mu\text{mol}/\text{m}^2$; for a city with 1 million inhabitants, its annual mean tropospheric NO₂ column decreases by 34 $\mu\text{mol}/\text{m}^2$; for a city with 10-million inhabitants, its annual mean tropospheric NO₂ column decreases by 58 $\mu\text{mol}/\text{m}^2$. The gradient decays sharply near the centre. Then it decreases with the distance to the city centre, and it tends to become an asymptote around 9km from the city centre likewise.

Acknowledgments

This research belongs to the SCALE-IT-UP project funded by the FNR CORE program. Mader M., Kilgarriff P., Lemoy R. provided the coordinates of the city hall.

References

- [1] Batty, M. (2013). *The new science of cities*. Massachusetts: MIT Press.
- [2] Bechle, M. J., Millet, D.B. and Marshall, J.D. (2011). 'Effects of income and urban form on urban NO₂: Global evidence from satellites', *Environmental Science & Technology*, 45(11), pp. 4914-4919.
- [3] EEA (2016). *Air quality in Europe - 2016 report*. Copenhagen: European Environment Agency. Available at: <https://www.eea.europa.eu/publications/air-quality-in-europe-2016> (Accessed; 8 August 2021).
- [4] Kirby, C., Greig, A. and Drye, T. (1998). 'Temporal and spatial variations in nitrogen dioxide concentrations across an urban landscape: Cambridge, UK', in Sokhi, R.S. (ed.) *Urban air quality: Monitoring and modelling*. Dordrecht: Springer, pp. 65-82.
- [5] Lemoy, R. and Caruso, G. (2020). 'Evidence for the homothetic scaling of urban forms', *Environment and Planning B: Urban Analytics and City Science*, 47(5), pp. 870-888.
- [6] Nicholas Hewitt, C. (1991). 'Spatial variations in nitrogen dioxide concentrations in an urban area', *Atmospheric Environment. Part B. Urban Atmosphere*, 25(3), pp. 429-434.
- [7] Restrepo, C.E. (2021). 'Nitrogen dioxide, greenhouse gas emissions and transportation in urban areas: Lessons from the Covid-19 pandemic', *Frontiers in Environmental Science*, 9, pp. 204.
- [8] Schindler, M., Caruso, G. and Picard, P. (2017). 'Equilibrium and first-best city with endogenous exposure to local air pollution from traffic', *Regional Science and Urban Economics*, 62, pp. 12-23.
- [9] UN (2019). *World urbanization prospects: The 2018 revision*. New York: United Nations. Available at: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf> (Accessed; 8 August 2021).
- [10] Wei, Y., Lemoy, R. and Caruso, G. (2019). 'The potential scaling of urban heat island and nitrogen dioxide with urban population: A systematic review' [PowerPoint presentation]. *ECTQG2019*. Available at: http://www.ectqg.eu/assets/download/ECTQG2019_Programme-and-book-of-abstracts.pdf?submit=Download+the+final+programme+and+book+of+abstracts (Accessed: 8 August 2021).

Parallel Session PSE3

5 November 9:00-11:00 am (GMT)

Networks

The first to be hit: spatio-temporal of analysis first-wave COVID-19 diffusion in two regions of northern Italy

Stefania BERTAZZON¹; Alison HANES¹; Rizwan SHAHID²; Isabelle COULOIGNER¹

¹Department of Geography, University of Calgary
2500 University Dr. NW, T2N 1N4, Canada
+1 403 220 7355, bertazzs@ucalgary.ca

²Primary Health Services, Alberta Health Services,
10101 Southport Road SW, T2W 3N2, Canada

Keywords: Diffusion, Spatial Analysis, Space Time Clusters, Lombardy, Veneto

1. Introduction

18 months ago, the WHO declared COVID-19 a pandemic; yet there is still a lot that we have not understood about this disease and, as geographers, about its diffusion in space and time. Indeed, a troubling aspect of this disease is its geographic diversity at scales that range from global to continental, national, and regional. Why did outbreaks occur in *some places* but not in others? Why was the transmission slower in *some places* than others? Is there something about certain *places* that can inhibit the virus transmission? Identifying local diffusion patterns, as well as transmission-inhibiting vs. -enhancing factors can help our fight with the current health emergency, as well as assist us in preparing for potential future epidemic outbreaks, allowing us to define effective local preparedness strategies, which, as we know, tend to be more effective than country-wide one-size-fits-all strategies (Gibertoni et al., 2021; Rosenkrantz et al., 2021).

In this study we analyze the first wave of COVID-19 diffusion in portions of the country that has unwillingly become the first in Europe and remains, to date, one of the hardest hit by mortality: Italy. When the first wave hit it, Italy was effectively a tabula rasa, serving as a case study, almost a testbed, demonstrating the epidemic transmission before containment measures were put in place. From a geographic perspective, the Italian example is even more compelling because of the dramatic geographic diversity of its outcomes. The virus spread from two distinct foci in the north and north-east, with hardly any significant clusters in central and southern Italy. Lombardy alone, the wealthiest Italian region, home to Milan and to 17% of Italy's population, has accounted for 48% of cases and 64% of deaths, on average, on each day between February 24th and April 20th, 2020. While local health policies undoubtedly played a major role, the catastrophic outcome in wealthy Lombardy remains

inexplicable in comparison with notoriously more disadvantaged southern regions (e.g., Campania, home to Naples and to 10% of the country's population, accounted for 1.6% of cases and of 17% of deaths, on average, over the same period).

We developed tools and employed advanced methods to analyze the first wave of COVID-19 diffusion in the two northern regions that were first hit: Lombardy and Veneto. Multivariate spatio-temporal analysis was applied to assess a pool of factors that may be associated with the transmission, with distinct local patterns.

2. Data and Methods

Spatio-temporal data has been publicly available since the early days of the pandemic, featuring several variables at daily intervals at the regional level (20 regions) as well as daily *New Cases* for the 107 provinces. In the spring of 2021, the Statistic Institute (ISTAT) published daily mortality rates for each of the 7,904 municipalities (comuni) from 2015. From these data, we calculated 5-year averages and 2020 excess mortality for each municipality over each of the first 14 weeks of the first epidemic wave (February – June 2020). Despite some limitations, this database provides an indication of deaths that can be directly attributed to COVID-19. We further acquired and validated a travel-time distance matrix across municipalities (ISTAT), as well as locations of major hospitals, and demographic data for the 2 regions of interest. We have collected additional variables, including population and demographics, while the collection of environmental data, including pollution levels, along with data on economic and industrial activities, completed by commuting patterns, is currently underway.

The diffusion pattern in the two regions was assessed by using clustering and spatial autocorrelation tools, and specifically by constructing a bivariate Moran's I, that compares temporally lagged data (by weekly intervals). This index is refined by the use of a neighbourhood specification where distances are weighted by ISTAT travel-time distances.

Spatial regression models (SAR-lag and GWR), estimated for each region, model the spatio-temporal diffusion pattern based on a set of independent variables that include lagged mortality rates, demographic variables (relevant age groups), proximity to hospitals, along with mobility indicators, such as proximity to international travel hubs (ports and airports) and border crossings, tourist arrivals, migrant workers, and levels of economic activity.

2. Results and Discussion

Even though epidemic outbreaks hit both regions at the same time, the outcomes were very different, with Lombardy taking a much higher toll, and displaying a very different spatial and spatio-temporal

pattern. Figure 1 illustrates the pattern of *New Cases*, assessed by local Moran's I, over the first 40 days of the epidemic (10-day intervals), at the provincial level, for the whole Italy. The only significant clusters, until then, were centered on the Lombardy provinces.

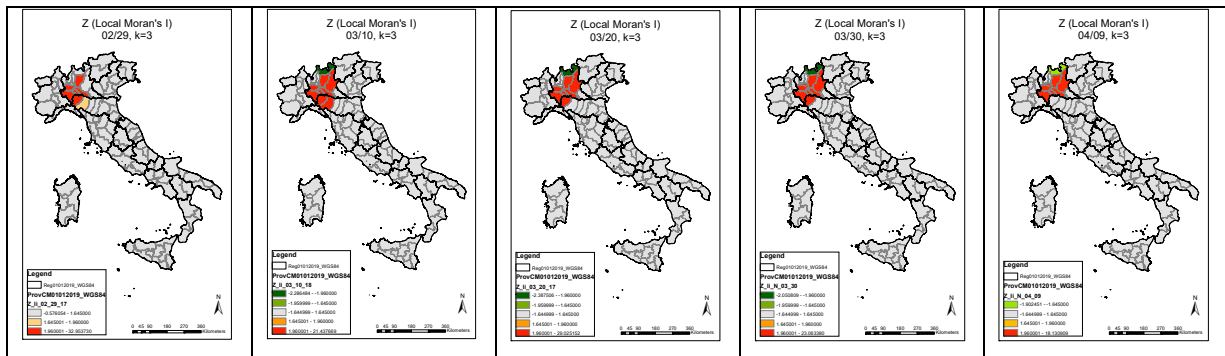


Figure 1: Local Moran's I of *New Cases* in Italy through the first 40 days of epidemic.

Figure 2 illustrates the spatio-temporal of excess mortality (compared to the 2015-2019 average) in the 2 regions, Lombardy and Veneto, throughout the first pandemic wave (February to June 2020).

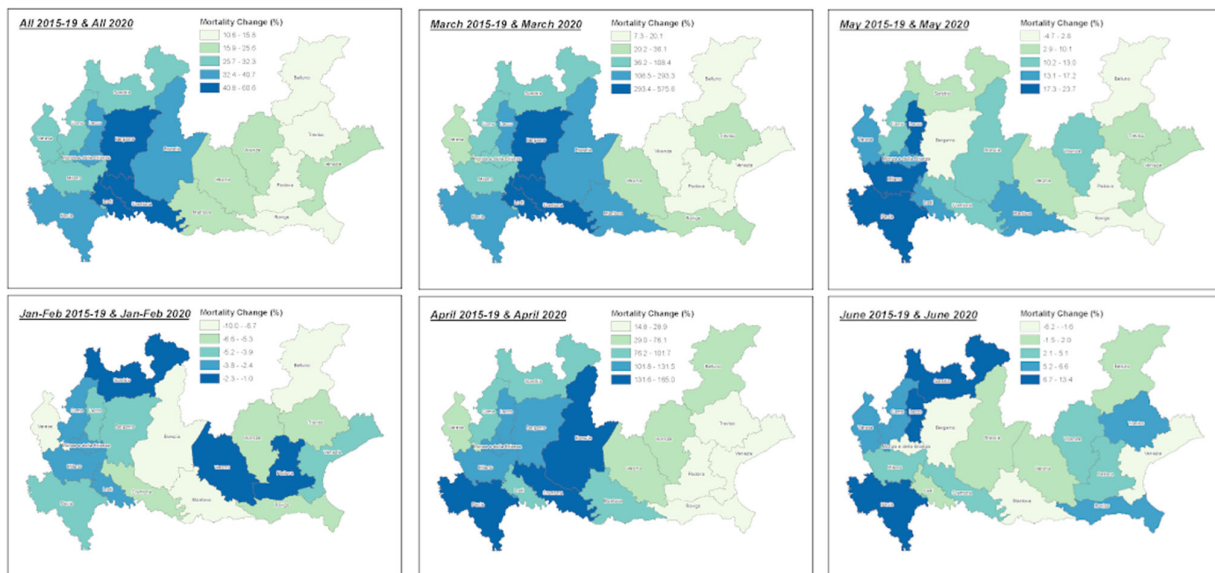


Figure 2: Excess mortality in Lombardy and Veneto in the first epidemic wave.

Figure 3 shows, as an example, a spatio-temporal analysis of excess mortality in Veneto, using a temporally lagged bivariate Moran's I at the municipal level, at weekly intervals.

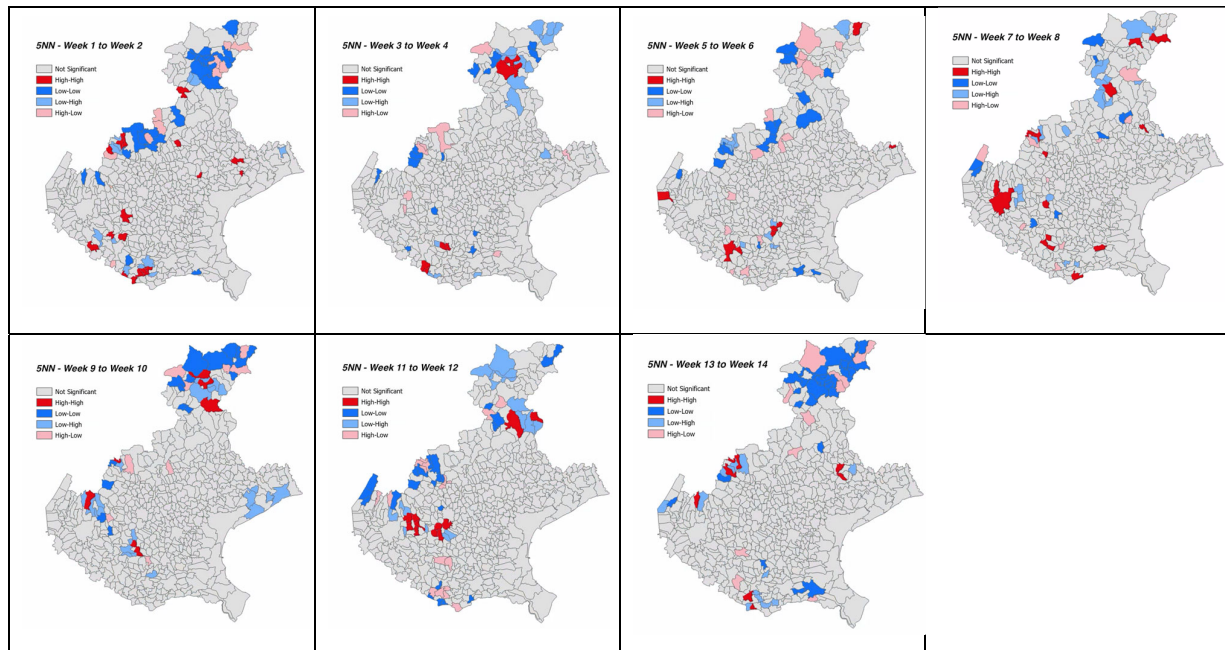


Figure 3: Spatio-temporal analysis of excess mortality in Veneto during the first epidemic wave.

The more rapid implementation of drastic, emergency health policies in Veneto appears to have played a major role in containing the clusters before their spread. As well, while Lombardy, and especially Milan, enjoys some of the best hospitals and health facilities in the country, the geographic concentration of health facilities within the capital city may have enhanced the virus diffusion in the outer rings, in other provinces, and in rural areas, where access to health care facilities was comparatively more difficult, especially for the elderly and more fragile population. Conversely, the geographical pattern of widespread small-sized, family-owned industry, common in Veneto and the eastern part of Lombardy, appears associated with a tightly connected social tissue, and intense movements across municipalities, especially in the low-lying plain and valleys. This productive landscape, common to eastern Lombardy and most of Veneto, appears associated with the worst outcomes in the Lombard provinces of Bergamo, Brescia, and Lodi, as well as in fringes of south-bordering Emilia Romagna, yet much less so in Veneto, potentially in connection with the above health policies and geographically widespread health care facilities (WHO, 2020). Notably, Milan is the main economic hub and financial capital in the country, with intense international communication, whereas Veneto tops all regions for tourist arrivals and, with its eastern location and seaports experiences privileged communication with eastern Europe and the Middle East. Seasonality of tourist flows may have played a role in the epidemic diffusion during the first wave.

3. Conclusion and further work

This work contributes to the body of knowledge on epidemic diffusion, by employing novel and advanced spatial analysis and quantitative geography for the understanding of first-wave COVID-19 in the Italian regions that were the first to be hit in Europe. The findings of this research point to factors, measured over space, and assessed with spatial statistical methods, and to their varying local association with mortality over time. These results, drawn from an empirical situation of relatively low preparedness, and different levels and speeds of response, both at the health care and at the political level, provide important insights to increase preparedness levels, at the geographic level, for potential developments of the current pandemic and the possibility of potential future ones. This work has some limitations: it relies on excess mortality estimates, which are only a proxy for COVID-19 mortality, and a portion of the work is still underway. Future lines of investigations include: refining the multivariate predictive models; extending this work to a multiscale framework, by using provincial-level incidence data and multivariate regional-level data; extending the scale by investigating and comparing a wider (if not complete) set of regions and territories. A further line of investigation will be the comparative study of other countries, at comparable geographical detail.

Acknowledgements

The authors wish to thank the members of the team that drafted the research proposal “Place matters: a geospatial analysis of first-wave COVID-19 diffusion” in the spring of 2020 for their contribution to discussions and ideas.

References

- Gibertoni, D., Adja, K.Y.C., Golinelli, D., Reno, C., Regazzi, L., Lenzi, J., Sanmarchi, F. and Fantini, M.P., 2021. Patterns of COVID-19 related excess mortality in the municipalities of Northern Italy during the first wave of the pandemic. *Health & place*, 67, p.102508.
- Rosenkrantz, L., Schuurman, N., Bell, N. and Amram, O., 2021. The need for GIScience in mapping COVID-19. *Health & Place*, 67, p.102389.
- WHO, Regional Office for Europe, 2020, An unprecedented challenge: Italy’s first response to COVID-19.

Spatially-explicit modeling of the SARS-CoV-2 transmission between population groups in Jerusalem

Ido Klein¹; Golan Ben-Dor¹; Aleksey Ogulenko¹; Avner Saadoun²; Itzhak Benenson¹

¹Porter School of the Environment and Earth Sciences, Tel Aviv University

PO Box 39040 Ramat Aviv Tel Aviv, 69978, Israel

+972509799900, idshklein@gmail.com (correspondent author)

²Strategy policy and planning division, Jerusalem municipality

Safra square 4, Jerusalem, 94141, Israel

Keywords: COVID-19, MATSim, Spatially explicit epidemic modeling

1. Introduction

COVID-19 pandemic reached Jerusalem in March 2020 and, from the very beginning, spread over the city's major population groups - Ultra-Orthodox Jews (UO, 25%), General Jewish population (GJ, 45%), and Arab population (30%), at different rates. Namely, the fraction of infected in the UO group was always the highest. After the first month of the pandemic, the public opinion turned out to be unambiguous: UO's disobedience to the epidemic restrictions is the main reason for their high infection rate and threatens public health nationwide. Rare opponents of this view argued that Jerusalem subpopulations are essentially (though not absolutely) segregated in space in all major aspects of their life - residence, activity centers, daily mobility routes - and this segregation can hamper the cross-infection between subpopulations.

We aim to understand the role of the UO subpopulation in the COVID-19 dynamics in Jerusalem. Such analysis demands spatially explicit data on COVID-19 dynamics for each subpopulation. Regrettably, the quality of the data for the Arab population is insufficient. We, therefore, focus on cross-infection between two groups of the Jewish population.

2. JEPISIM, the Dynamic Model of COVID-19 Pandemic in Jerusalem

We investigate COVID-19 dynamics in Jerusalem with the EPISIM (Muller et al., 2020) - extension of the MATSim, an agent-based spatially explicit urban traffic simulation environment that is based on direct simulation of human travel behavior and adaptation to transport (Horni et al, 2016). While MATSim focuses on agents' trips, EPISIM accounts for contacts during the trips and at the facilities where agents perform their daily activities. EPISIM includes three agent-focused sub-models of (1) Contact, (2) Infection, and (3) Progression - development of COVID-19 disease.

EPISIM requires a calibrated and validated model of the city transportation activities. We employ the JMATSIm, a MATSim model for Jerusalem Metropolitan Area developed for assessment of the current status and the future of the Jerusalem transportation network.

3. JEPISIM validation

We validated JEPISIM based on the age-specific data on Daily Confirmed Cases (DCC) in 190 Jerusalem statistical areas, supplied by the Jerusalem municipality. The COVID-19 dynamics by subpopulations are presented in Figure 1. As mentioned, the Arab subpopulation was excluded from the further analysis.

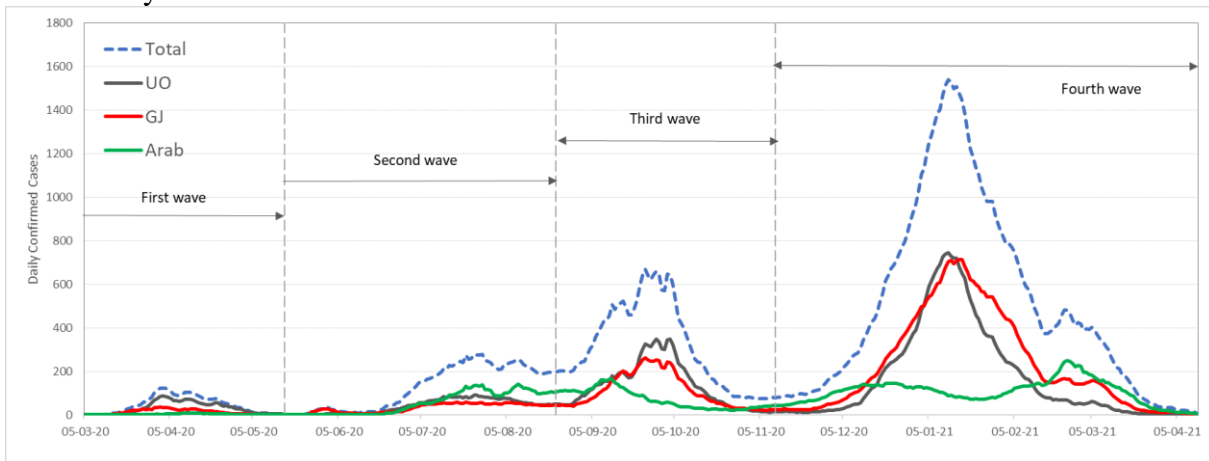


Figure 1. Daily Confirmed Cases (DCC) in Jerusalem, by subpopulation

4. JEPISIM Model Scenarios

4.1. General view of COVID-19 dynamics in Jerusalem and education activities during the second and third waves

We focused on the second and third epidemic waves for three reasons. First, Jerusalem faced the second wave having resources for infection testing and tracing mobilized. Second, at that moment, the public was already fully aware of the COVID-19 risks. Third, the transition to the third wave was essentially affected by the renewal of educational activities. The latter is especially important in the Jerusalem case, where about 40% of the population is younger than 18. Figure 2 presents the general outline of the second and third epidemic waves.

We start modeling the second wave from June 1st by “infecting” 160 randomly chosen agents during the first two weeks. We reflect the schedule of the education system and compliance with the established restrictions by varying the coefficient f_x – a fraction of residual education activities performed by subpopulation X. Figure 3 presents our assumptions on the f_x dynamics. UO subpopulation incompliance with the country-level restrictions is reflected by $f_{UO} = 0.7$ between June 21^d and July 7th.

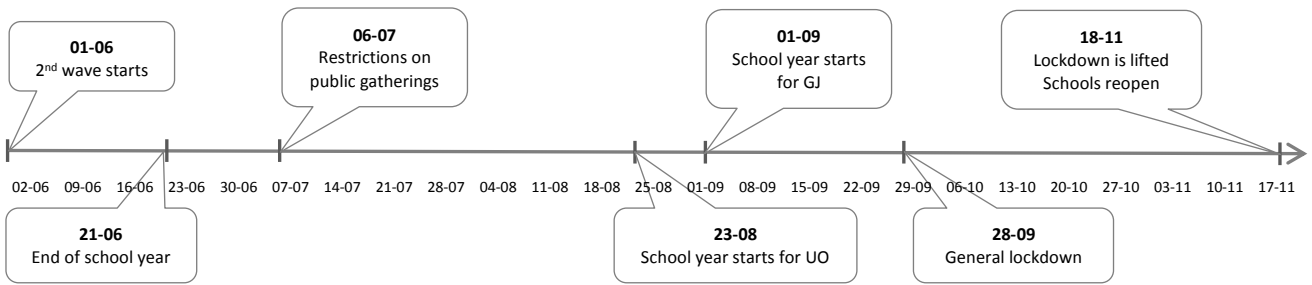


Fig. 2. Timeline of the second and third epidemic waves in Jerusalem.

With chosen f_{GJ} and f_{UO} , JEPISIM fits well the empirical data on COVID-19 dynamics during the second and third epidemic waves for both subpopulations. Below we call it the baseline scenario.

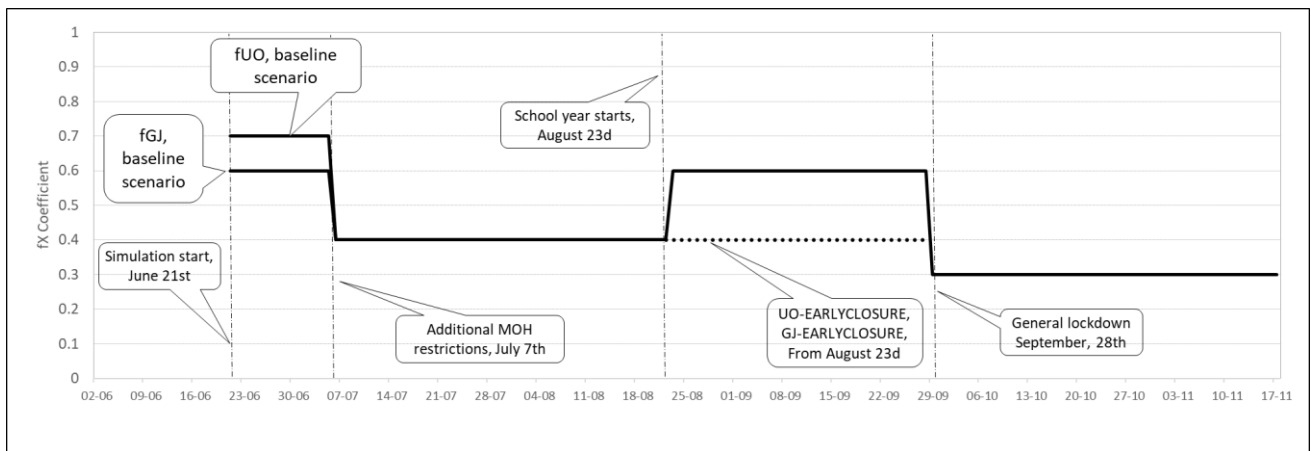


Fig. 3. f_x dynamics for the baseline scenario and deviations of f_x in alternative scenarios

4.2 Alternative Scenarios

To investigate the role of each subpopulation in the COVID-19 epidemic dynamics, we compare several scenarios:

1. GJ-ONLY-INFECTION: Only GJ agents can infect and be infected.
2. UO-ONLY-INFECTION: Only UO agents can infect and be infected.
3. UO-EARLYCLOSURE: UO education system remains closed ($f_{UO} = 0.4$) until September 28th, f_{GJ} follows the baseline scenario.
4. GJ-EARLYCLOSURE: GJ education system remains closed ($f_{GJ} = 0.4$) until September 28th, f_{UO} follows the baseline scenario
5. UO-SMALLFAMILY: Intensity of contacts within the UO households is reduced to match the average number of infected in a UO household with the one in a GJ household (typically, the UO family has twice as much children comparing to the GJ family (Malach and Cahaner, 2020)). f_{GJ} and f_{UO} follow the baseline scenario.

For each scenario, we present within- and between-groups infection rates over the period of the second and third epidemic waves.

5. Results

Table 1 presents the total numbers of infected and within- and between-group infection rates for the baseline scenario. The majority of the COVID-19 cases are induced by a within-group contagion. However, 14% of the UO-members are infected by the GJ-members, and 21% of the GJ-members are infected by the UO-members.

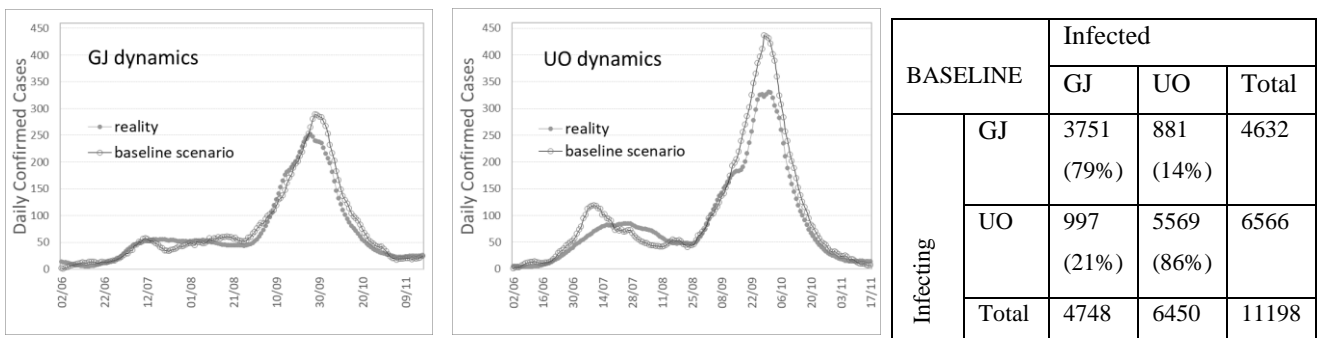


Table 1. The baseline scenario COVID-19 dynamics and within and between-subgroups infection rates

Figure 4 presents the infection dynamics for the GJ-ONLY-INFECTION and UO-ONLY-INFECTION scenarios compared to the baseline scenario. Evidently, the independent outbreak would not develop within the GJ subpopulation, while for the UO subpopulation, the dynamics during the fourth wave would have to repeat itself, albeit on a smaller scale.

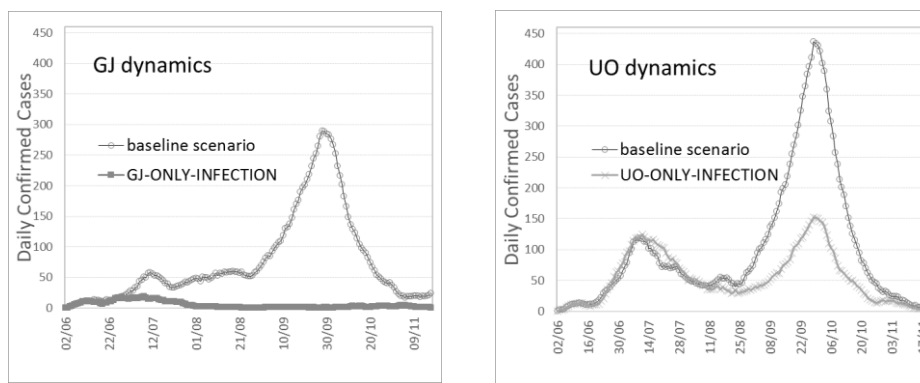


Fig. 4. The COVID-19 dynamics in the GJ-ONLY-INFECTION and UO-ONLY-INFECTION scenarios

Table 2 presents the outcomes of the UO-EARLYCLOSURE scenario. Earlier closure of the UO education system during the fourth wave halves the DCC for the UO subpopulation and outbreak fades.

For the GJ, the reduction factor is 1.5. Cross-infection between groups also changes: transmission from UO to GJ is lower, while transmission in opposite direction is higher.

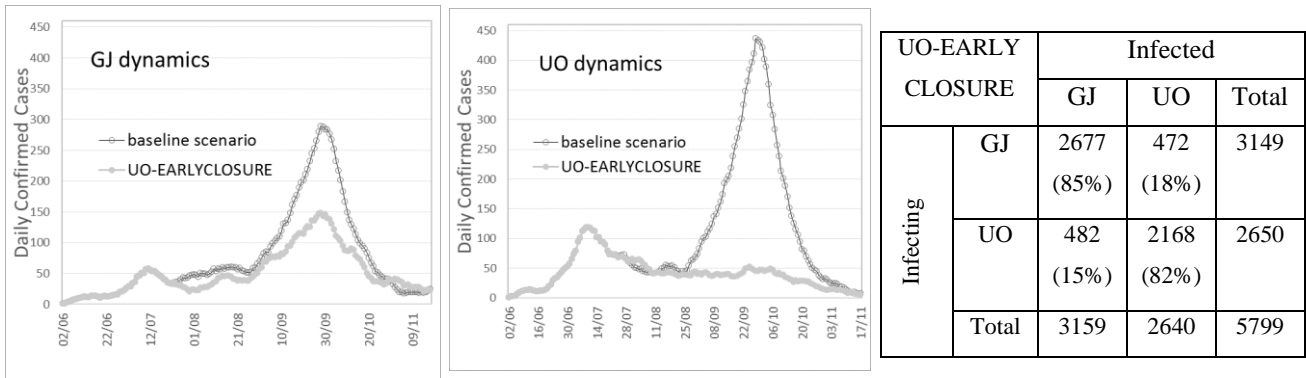


Table 2. The COVID-19 dynamics in the UO-EARLYCLOSURE scenario and within and between-subgroups infection rates

Dynamics in the GJ-EARLYCLOSURE scenario is symmetric to the previous one (Table 3), although the number of the infected UO-members is essentially lower than in the baseline scenario

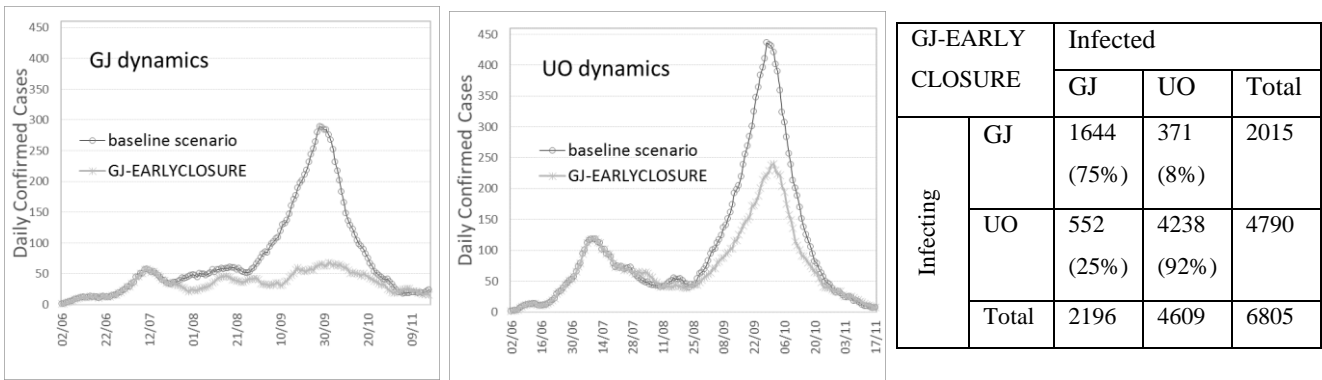


Table 3. The COVID-19 dynamics in the GJ-EARLYCLOSURE scenario and within and between-subgroups infection rates

Table 4 presents the outcomes of the UO-SMALLFAMILY scenario. Here, a lower infection rate within the UO households would have reduced the number of cases for UO to one-third of its baseline value, while the reduction factor for GJ subpopulation is 2.5. The transmission rate from GJ to UO would remain the same, while the opposite effect would weaken.

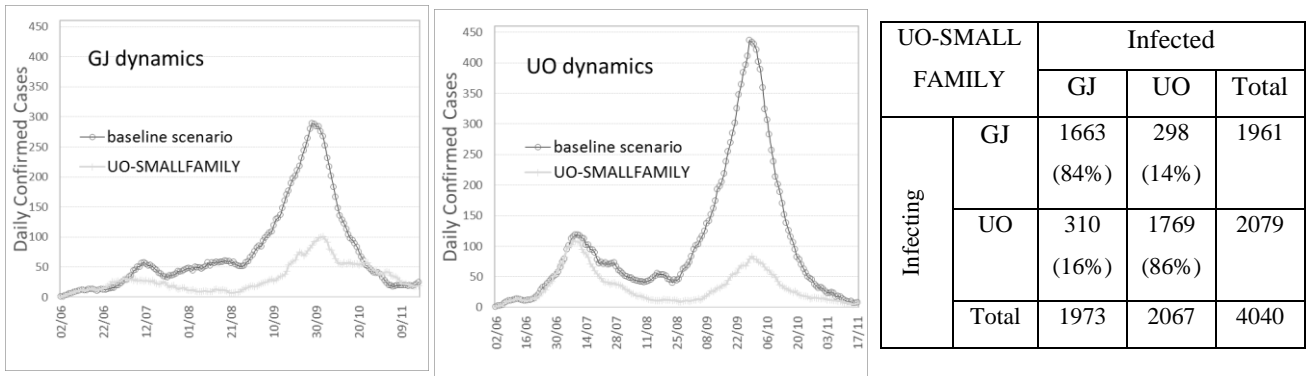


Table 4. The COVID-19 dynamics in the UO-SMALLFAMILY scenario and within and between-subgroups infection rates

6. Conclusion

JEPISIM model adequately represents the COVID-19 pandemic dynamics in Jerusalem during the second and third epidemic waves for two Jewish subpopulations. Due to the segregation, the major driver of the epidemic lies inside each of them. Yet, UO is responsible for, roughly, 25% of infections among the GJ. Without this factor, the outbreak within the GJ subpopulation would have died down. At the baseline contact level between two subpopulations, better adherence to the restrictions in the education system would halve the infection rate. The major driver of the epidemic within the UO subpopulation is family transmission: if only its rate had dropped to the GJ level, the total number of cases would reduce by one-third. Overall, public opinion on the UO's crucial role in the infection spreading seems to be a half-truth. Indeed, we demonstrate that, seemingly, the UO subpopulation disobedience to MOH restrictions added significantly to the disease transmission among the rest of the Jerusalem population. On the other hand, we also demonstrate that the background reason for the faster infection spreading within the UO subpopulation is an inherently larger size of their households.

Acknowledgments

This research was supported by the Research Grant 0603417421 of Israeli Smart Transportation Research Centre (ISTRC). The data are supplied by the Jerusalem Municipality and Transport Master Plan Team.

References

- Horni, A., Nagel, K. and Axhausen, K.W., 2016. *Introducing MATSim* (pp. 3-7). Ubiquity Press.
- Malach, G, Cahaner, L, 2020. *Statistical report on Ultra-Orthodox society in Israel 2020*. The Israeli Democracy Institute.
- Müller, S.A., Balmer, M., Neumann, A. and Nagel, K., 2020. Mobility traces and spreading of COVID-19. *MedRxiv*.

A multivariate modelization to grasp bicycling spatial heterogeneity within French fonctionnal urban areas, the example of Toulouse.

Léo TAILLANDIER¹

1 PhD student, University of Paris 1 Panthéon-Sorbonne
5, cours des Humanités, 93322 Aubervilliers cedex, France
+33 (0)6 46 72 75 83, taillandierleo@gmail.com

Keywords : bicycling, mobility, suburbs, multivariate modelization, French fonctionnal urban area

1.Introduction

For twenty years, bicycling, as a daily mode of transport, has been evolving in a very heterogeneous way within French fonctionnal urban areas. Cyclists has been more and more numerous in the centre parts but not in the inner suburbs where bicycling have been decreasing, and in the outer suburbs where its decrease is very substantial¹. In this presentation, I bring to light factors explaining spatial heterogeneity of bicycling within French urban areas with the example of Toulouse. It resonates with a growing corpus of studies which try to enlighten characteristics of spatial context that favour bicycling with model-based methodologies. These studies explored the different spatial components that could be twinned with a higher bicycling modal share. Thus, the « bikeability » of spaces results not only from bicycling policies but also from features of the natural environmental features (weather, slopes..), and also the built environmental features (road network characteristics, public transport network...). However, in Europe, almost all of these studies take place in the central part of the urban areas where bicycling has yet a significant part of the modal share (mainly Scandinavian and Dutch cities). Few of them actually question why there is a lower bicycling rate in the suburbs. The aim of my work is to enlighten the causes of the spatial heterogeneity of bicycling in French urban areas, in particular the growing gap between the centre and the suburbs. By questioning this gap, my second aim is to wonder if factors, that promote or impede bicycling in other contexts, are the same in France as in other European countries, the same among french fonctionnal urban areas, and the same between the suburbs and the centre of fonctionnal urban areas.

¹ In this study the components of the French fonctionnal urban area called « centre », « inner suburb » and « outer suburb » refer to the foreign classification of the French statistical institute (INSEE) that divided French urban areas in three parts according to morphological criteria (center and inner suburbs) and according to a proportion of daily work commuters in a municipality (outer suburbs).

2.Hypothesis and methodology

My hypothesis stand on the principle that the decision to ride a bike for daily trips results from a combination of factors that (1) can't be summarized by the diptych « length of trips » and « population density », (2) are distributed heterogeneously in urban areas (composition effects), and (3) are not as strongly in correlation with bicycling depending on the parts of the urban area. Thus, I built a multivariate model with data mainly coming from the Toulousian travel survey of 2013 (the « Enquêtes Ménages Déplacements ») that inventories trips over one whole day of 9000 individuals. Toulouse is an urban area of 1.300.000 inhabitants, that is characterised by a recent and fast demographic growth twined with urban sprawl. Like in most of the great French urban areas, bicycling modal share is quite low (about 3%), it has increased in the centre part of the urban area since the beginning of the 2000's but decreased in the inner suburbs and decreased sharply in the outer suburbs. In my multivariate model, factors influencing bicycling are embodied by around twenty explanatory variables (figure 1). The variable of interest is binary : the individual that makes at least one trip by bicycle among all the trips of the day surveyed is coded 1. Some of the explanatory variables depict built and natural environmental characters surrounding individuals' residence. I tested different sizes of individuals' surrounding spaces to question what is the proper perimeter to grasp the influence of the different spatial factors. Furthermore, as control variables, I took into account the weight of factors influencing bicycling that result from the socio-demographic features and from individuals' mobility practices. I stratified this model depending on the parts of the urban area to grasp spatial variations of effects of factors. Because of the spatial limitations of the survey, the model stratification divides the sample in two groups, one of them aggregates those who reside in the central part of the urban area (that is to say the municipality of Toulouse), another one aggregates inner suburbs residents and a part of the outer suburb residents.

3.Results :

The stratified model corroborates the fact that there is a significant gap between the proportion of cyclists in the Toulousian suburbs and in the town-centre of Toulouse. However, a map of these proportions with a thinner spatial division² shows more complex spatial schemes (map 1). There is a clear bicycling decrease from the centre to the outer suburbs but this decrease is rather gradual. There is an intermediate zone in the western part of the inner suburbs of Toulouse where there are between 2 and 8 % of cyclists. This zone doesn't match with the inner suburb borders and several centre neighbourhoods are closer to the inner suburb than the hypercentre in consideration of its

2 This spatial grid is a specific subdivision used in the local mobility survey, this is the thinnest grid that can be used to guarantee solidity of the simplest statistical measures. It is equivalent to a group of municipality in the outer suburbs, a municipality in the inner suburbs and a neighbourhood in the centre town.

proportion of cyclists. In a same way, the very low rate of cyclists in the eastern part of the inner suburbs sounds like the rates of the great majority of the outer suburbs.

This spatial description is mostly enlightened by the results given by the model. It reinforces the idea that long trips and possession of a car are the stronger factors which are correlated with a weak probability to be a cyclist. The gradual decrease of share of cyclists from the centre to the outer suburbs matches with the increase of the share of car owners and with the increase of the mean length trips. The spatial inequalities of bicycling services and infrastructures distribution, like cycling paths or bike-share, is less determining than the two factors quoted above but widen the gap between the centre and the other parts of the urban area because they are still very concentrated in the hypercentre of Toulouse.

The spatial heterogeneity of bicycle share within the inner suburb can be explained by other influencing factors. The lower share of cyclists in the eastern part of the inner suburb, than in the west, stems from the over-representation of elderly people. An individual aged of more than 65 years old has two twice less chances to be cyclists than those aged between 18 and 65 and three times less chances than people under 18. This opposition between eastern and western inner suburbs ensues also from the fact that the first is more hilly than the second. Individuals who have a very hilly neighbourhood have around 60 % less chance to be a cyclist than those who reside in a flat neighbourhood. Competition between public transports and cycling adds to the heterogeneity of bike share, the model shows that individuals living at less than 200 meters from a transport public station have three times less chance to be a cyclist compared to those that live at more than 500 meters from these stations. Thus, some neighbourhoods in the south of the municipality of Toulouse or Blagnac in the North-West have a lower proportion of cyclists compared to their nearby neighbourhoods partly because they are well served by tram, subway or urban train lines.

Beside these composition effects, there are few factors influencing bicycling that are specific to the centre or the inner suburbs. The more surprising difference is the relation between bicycling and intermodality. In the center town those who use at least two different transport mode in one travel are very less likely to be cyclists. This difference stems partly because the inhabitants of the centre live very close to a transport station and walk to join it, but this intercation doesn't explain entirely this difference.

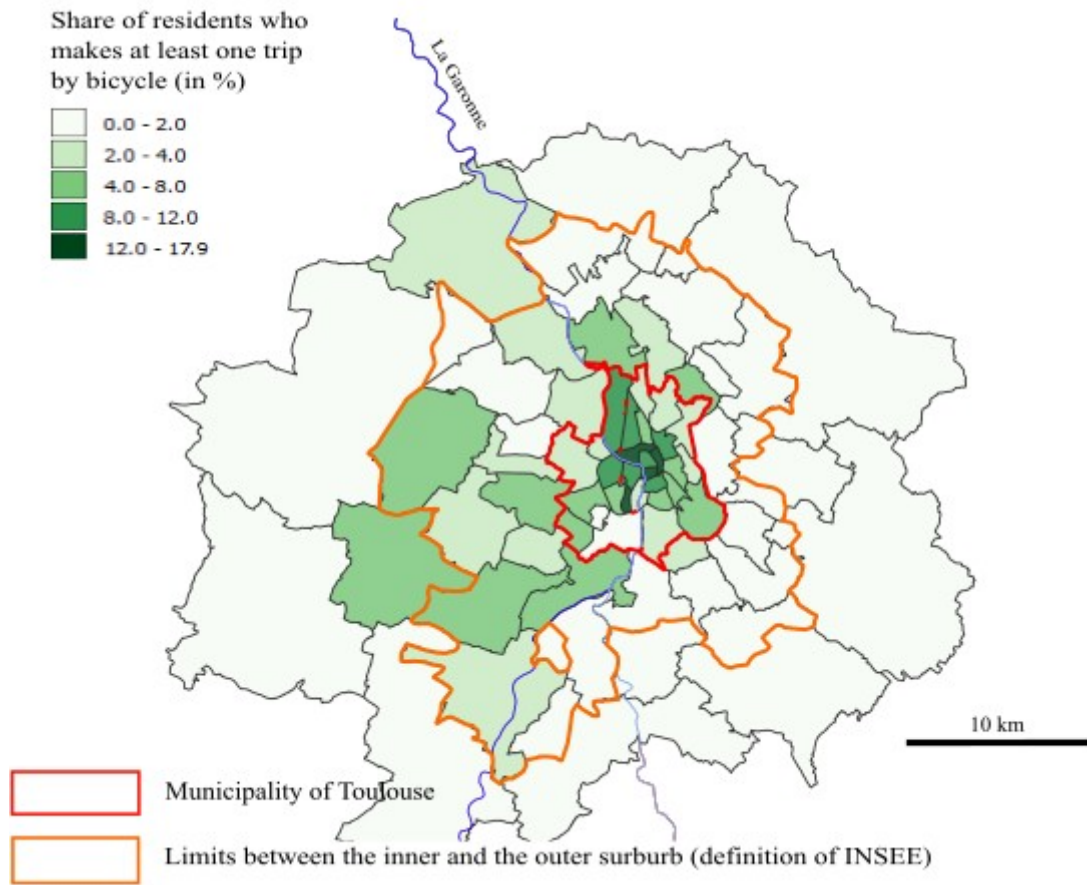
By crossing map analysis and multivariate modelization, this work brings many elements to understand why bicycling share is so heterogeneous, within and between the different parts of the Toulousian urban area. Now, these results can be deepen in different directions. Firstly, our modelization graspes a great part of the causes of the spatial phenomenons but some variables could be added or sharpened in a different way to enlighten other explanations. Secondly, we need to

compare our conclusions about Toulouse with other French urban area to understand if there is a common system between the French urban area or there are different elements, in particular in cities where bike share is higher. Finally, the spatial limitation of the Toulousian survey prevents me from distinguishing inner suburbs and outer suburbs, here again, the study of another fonctionnal urban area, where the local mobility survey covers, can shade light to the potential specificity of the outer suburbs compared to the inner.

Variables	Odds-ratio	95% Confidence intervals	p-value
Number of bicycle parking lots less than 500 meters from home			0.018
(0]	—	—	
(0,2]	1.09	0.75, 1.55	
(2,5]	2.00	1.31, 3.02	
(5,30]	1.23	0.70, 2.09	
Distance in km between home and the nearest self-service bike station terminal			0.031
(0,0.2]	—	—	
(0.2,0.5]	0.79	0.57, 1.10	
(0.5,1]	0.30	0.13, 0.64	
(1,3]	0.54	0.32, 0.91	
(3,5]	0.50	0.25, 0.97	
(5,30]	0.58	0.31, 1.05	
Distance in km between home and nearest public transport station			<0.001
(0,0.2]	—	—	
(0.2,0.5]	1.45	0.87, 2.51	
(0.5,1]	3.19	1.85, 5.70	
(1,2]	2.88	1.60, 5.38	
(2,4]	2.41	1.19, 4.97	
(4,30]	2.65	1.18, 6.02	
Way length limited to 30 km / h or less in the individual neighbourhood (m / m²)			0.12
(0]	—	—	
(0,0.4]	1.28	0.85, 1.95	
(0.4,1.5]	1.66	1.00, 2.75	
Elevation categories (coefficient of variation calculated on the differences between altimetric points)			<0.001
(0,0.01]	—	—	
(0.01,0.03]	1.12	0.85, 1.49	
(0.03,0.05]	0.94	0.66, 1.33	
(0.05,0.2]	0.39	0.26, 0.58	
Density of cycle lanes (meters per meters²)			0.062
(0]	—	—	
(0,0.03]	1.19	0.69, 2.05	
(0.03,0.04]	1.85	0.95, 3.62	
Number of intersections of the road network per m²			0.041
(0,0.005]	—	—	
(0.005,0.0075]	1.13	0.71, 1.86	
(0.0075,0.009]	1.55	0.95, 2.59	
(0.009,0.16]	0.94	0.49, 1.78	

Age categories			<0.001
5 to 17 years old	—	—	
18 to 24 years old	0.31	0.14, 0.63	
25 to 39	0.52	0.21, 1.24	
40 to 54	0.60	0.24, 1.49	
55 to 64	0.57	0.22, 1.43	
more than 65	0.19	0.07, 0.49	
Sex			<0.001
Men	—	—	
Women	0.57	0.45, 0.73	
Socio-professionnal categories			<0.001
Employees	—	—	
Farmers	1.37	0.07, 7.60	
Craftsmen / traders	0.70	0.28, 1.53	
Managers / liberal professions	1.84	1.27, 2.69	
Workers	1.20	0.80, 1.80	
Pensionners	0.87	0.50, 1.48	
unemployed, inactives	1.43	0.56, 3.18	
schoolchildren	0.40	0.15, 1.07	
College and hih school students	0.85	0.36, 1.94	
Students and apprentices	1.44	0.75, 2.72	
Car ownership			<0.001
no	—	—	
yes	0.36	0.26, 0.50	
Bike parking at the workplace			<0.001
no	—	—	
yes	2.63	1.88, 3.71	
E-bike ownership			<0.001
no	—	—	
yes	2.70	1.56, 4.47	
Intermodality practice			<0.001
no	—	—	
yes	0.27	0.17, 0.43	
Total length of trips made in one day by the person	0.92	0.90, 0.94	<0.001
Total number of trips made by the person in a day	1.37	1.27, 1.47	<0.001
Number of chains of trips made by the person in a day	1.13	0.98, 1.30	0.10
Degree of complexity of the person's travel chains			0.036
Only simple chains	—	—	
At least one complex chain	0.77	0.61, 0.98	

Figure 1 : Set of variables kept in the final model with their degree of significance and the direction of their relationship with the variable of interest.



Map 1 : Bicycling spatial heterogeneity in the toulousian fonctionnal urban area

Partners' non-food-related activities and arrangement of food-related household labor: a case study of coupled adults in Toronto, Canada

Bochu Liu¹; Michael Widener¹ (correspondent author); Lindsey Smith¹

¹Department of Geography and Planning, University of Toronto, St. George

100 St. George Street, Toronto, M5S 3G3, Canada

+1 4169393616, michael.widener@utoronto.ca (correspondent author)

Keywords: food-related housework, gendered labor division, within-household time allocation, time constraints, cohabiting partners.

Abstract

1. Research gaps and purposes

Food-related housework, including meal preparation and after-meal cleaning up, is of great importance to diet-related health due to its close relationship to consumption of the more healthful homemade food (Wolfson and Bleich, 2015). Past work has reported an inclination to substitute home preparation and cooking with quicker options when individuals experience binding time constraints incurred by non-food-related activities (e.g., long hours of work) (Jabs and Devine, 2006). However, the associations between food-related activities and time use of non-food-related activities have been *hitherto* confined to the individual level, which misses the ways a cohabiting partner's time allocation comes into play in the arrangement of food labor at the household level.

Apart from the health effects, the diverse arrangement patterns of food-related housework between coupled partners has implications for gender inequalities. Past studies on household labor division have found that women, in general, take a higher load of food-related tasks than men in terms of duration, participation rate (Taillie, 2018), and perceived responsibilities (Lake *et al.*, 2006). Although these studies have uncovered the unequal division of food-related household labor by gender, how these division patterns are contextualized within coupled partners' non-food daily routines has been scarcely explored.

In addition, past studies have found that interpersonal interactions of partners' activities vary between households residing in high- and low-density areas (Ettema, Schwanen and Timmermans, 2007). Such disparities between urban and suburban neighborhoods were documented on coordination of paid work, childcare, and a coarse definition of household labor (Ettema, Schwanen and Timmermans, 2007; Schwanen, Ettema and Timmermans, 2007). Whether the associations of partners' non-food-

related time use and arrangement of food-related housework are contingent on residential areas remains to be explored.

To address the aforementioned gaps, this paper aims to examine how partners' time allocated to non-food-related activities impacts both the total amount of time spent on food-related housework and the difference in time spent on food-related housework between coupled women and men. We further explore whether the associations between partners' non-food-related time use and food-related housework vary between urban and suburban neighborhoods using a sample of opposite-gender couples living in Toronto, Canada.

2. Data and methods

The data used in this paper was collected in Toronto, Ontario, Canada in March and April 2019 from the Food Activities, Socioeconomics, Time use, and Transportation (FASTT) Study, and consists of a paper survey about socio-demographics and food behaviors and a seven-day time use diary. The FASTT study attempted to collect information about co-residing household members, making it feasible to acquire daily time use logs of coupled adults. The analytical dataset for this paper contains 108 daily time-use logs from 17 pairs of cohabiting partners (by chance, all are female-male partners) living in one urban and two suburban neighborhoods. Five activity categories of work, caregiving, non-food-related housework, food-related housework, and recreation are derived from the open-end questions about "what they were doing" in each of the ten-minute interval throughout a seven-day period. Food-related housework comprises unpacking groceries, meal preparation or cooking, snack preparation, and dishwashing or cleaning up after meal.

This study employs mixed linear regressions with random effects to account for the multi-level structure of the data where daily observations are nested within couples. In this way, the estimated coefficients of daily duration variables can reflect the associations between non-food-related activities and food-related housework at the daily level. The first outcome variable of interest is the total duration of food-related housework, defined as the sum of time spent by coupled partners on a day. The second outcome variable considered is the difference between partnered adults in duration of food-related housework, and is measured by subtracting the man's time spent on food-related housework from that of their partner's because the woman's time spent on food chores is typically greater. Men and women's time spent on work, caregiving, non-food housework, and recreation are the explanatory variables of interest. Interaction terms of those variables and a dummy variable representing the urban residential neighborhood are included to test the potential urban-suburban disparities. Socio-demographic covariates including age, perceived income adequacy, and ethnicity. The models also

controls for presence of cohabiting children, adults, or older adults, which may confound within-household time allocation.

3. Results

Results show that the associations between time spent on non-food activities and arrangement of food-related housework varied by gender and residential neighborhood. In response to men's increased work duration, women from both urban and suburban residential neighborhoods increased their time spent on food-related housework to maintain an unchanged total at the couple level and therefore led to an increased gender difference in food labor between women and men. When women increased their work duration, men from both urban and suburban neighborhoods also increased their time spent on food-related housework, which resulted in a decreased gender gap. However, suburban men's additional time spent on food-related housework was not high enough to maintain the total household duration of food chores, as was seen with urban households.

A similar urban-suburban disparity was found when women carried out more care-related activities outside home. When women from suburban neighborhoods spent additional time on caregiving, the total duration of food-related housework decreased with an unchanged the gender difference in this food labor. In contrast, when suburban men spent higher amount of time on caregiving or non-food-related housework, an increased gender difference between women and men and an unchanged couple-level total were observed. This suggests that men from suburban neighborhoods were irresponsive to their partners' increased duration of caregiving, while women modulated their food-related housework to compensate for men's reduced engagement in food-related housework. This contrast between women and men's response to their partners' increased non-food-related labor was not found among urban couples.

Additionally, urban men's increased duration of recreation was unexpectedly associated with a larger gender difference in time spent on food-related housework while maintaining the status quo of total household duration. This finding suggests that women may even increase their food-related housework as a response to men's increased time spent on discretionary activities.

4. Conclusions and policy implications

The gendered associations of time allocated for non-food-related activities and arrangement of food-related housework suggest that women were more responsive to their partners' increased non-food activities and may still be responsible for food-related chores even with increased non-food-related household responsibilities. It highlights the gender inequality displayed in one's responsiveness to the change of partners' time spent on non-food-related tasks. The urban-suburban discrepancy in these

associations reflects varied strategies coupled adults used to cope with women's increased time constraints. While assigning more food-related housework to one partner worked for urban households to secure a stable amount of time spent on food-related chores at the household level, suburban couples were likely to reduce the total household time spent on the activity to cope with more binding time constraints.

These preliminary findings underlie the need to account for partners' time allocation and residential contexts when studying dynamics of food-related activity arrangement within households and inform urban policy makers about the household perspective when designing effective interventions to promote dietary health and gender equality.

Acknowledgments

Bochu Liu was supported by Maple Leaf Board Scholarship in Food Insecurity honoured by Maple Leaf Centre for Action on Food Security. The Food Activities, Socioeconomics, Time-use, and Transportation (FASTT) Study is supported in part by funding from the Social Sciences and Humanities Research Council. In addition, this research was undertaken, in part, thanks to funding from the Canada Research Chairs Program and the Ontario Early Researcher Award awarded to Michael Widener.

References

- Ettema, D. F., Schwanen, T. and Timmermans, H. J. P. (2007) 'The effect of locational mobility and socio-demographic factors on task and time allocation in households', *Transportation*, 34(1), pp. 89–105.
- Jabs, J. and Devine, C. M. (2006) 'Time scarcity and food choices: An overview', *Appetite*, 47(2), pp. 196–204.
- Lake, A. A. et al. (2006) 'Food shopping and preparation among the 30-somethings: Whose job is it? (The ASH30 study)', *British Food Journal*, 108(6), pp. 475–486.
- Schwanen, T., Ettema, D. F. and Timmermans, H. (2007) 'If you pick up the children, I'll do the groceries: spatial differences in between-partner interactions in out-of-home household activities', *Environment and planning. A*, 39(11), pp. 2754–2773.
- Taillie, L. S. (2018) 'Who's cooking? Trends in US home food preparation by gender, education, and race/ethnicity from 2003 to 2016', *Nutrition Journal*, 17(1), pp. 1–9.
- Wolfson, J. A. and Bleich, S. N. (2015) 'Is cooking at home associated with better diet quality or weight-loss intention?', *Public Health Nutrition*, 18(8), pp. 1397–1406.

Comparing geographies of care in contrasting neighbourhoods in Toronto, Canada

Lindsey G. SMITH¹; Bochu LIU¹; Michael J. WIDENER¹

¹University of Toronto

Department of Geography and Planning, Sidney Smith Hall, 100 St George Street, Toronto ON M5S 3G3, Canada

+1 (416) 978-3375, lindsey.smith@utoronto.ca (correspondent author)

Keywords: travel, care, gender, time use, GPS

1. Introduction

Social, cultural, and spatial contexts can influence the division of labour and time spent caregiving, which we define as provision of care for dependent children or adults. Despite changes to caring and employment patterns over the past 50 years, differences in time use, notably by gender, income, and place, persist (Moyser and Burlock, 2018). These differences can result in feelings of time scarcity (Strazdins *et al.*, 2011), which may have implications for outcomes including mental health and travel behaviour.

Women's increased participation in the labour market has been met with some increases in childcare participation by men, however the division of care remains unequal (Bianchi *et al.*, 2012). Given the dynamic and unpredictable demands of caring, caregiving responsibilities have traditionally been taken on by those with less rigid work patterns and lower wages (Carmichael, Charles and Hulme, 2010) which reinforces gender inequality in the labour market. Place further plays an important role in the negotiation of caring responsibilities. At the macro-level, policies relating to caregiving supports vary by country (Stanfors, Jacobs and Neilson, 2019) while local cultures and household expectations can inform gendered carescapes (Chick, Heilman-Houser and Hunter, 2002; Bianchi and Milkie, 2010). Additionally, the location of destinations such as homes, schools, healthcare facilities, and the availability of transport infrastructure to facilitate travel between them are intrinsically linked to time spent travelling and available for care.

Trips made for care, particularly by those juggling employment responsibilities, are often short and follow complex trip-chaining patterns (Lee, Hickman and Washington, 2007). Relatively 'fixed' trips, such as the school run, can limit spatial and temporal flexibility and ultimately impact employment options and available leisure time (Wheatley, 2013). The car is frequently perceived as a solution to these constraints, in some cases irrespective of distance, congestion, or travel times (Jain, Line and

Lyons, 2011; Westman, Friman and Olsson, 2017). Gendered differences in care and mobility, fixity constraints on carers' time use, and increased car use highlight important equity and sustainability issues surrounding geographies of care. However, many transport datasets do not fully account for care-related travel (De Madariaga, 2016).

1.2 Study aims

We aim to build on existing evidence around geographies of care which has been centred on qualitative and self-reported evidence. We use GPS and time use data to map and describe the complex patterns of travel for care (e.g., mode, distance, trip-chaining), how these fit with time allocated to other activities, and how these compare by neighbourhoods with differing urban form. We intend to narratively consider these findings alongside equity and sustainability issues relating to gender and travel modes, as well as recent changes to working arrangements including working from home options.

2. Methods

Data were used from the Food Activities, Socioeconomics, Time, and Transportation (FASTT) Study. Respondents were invited if they lived in one of three neighbourhoods in Toronto, Canada (Parkdale, Rexdale or West Hill) and were a parent of a child living at home. Neighbourhoods were selected due to their comparable income levels and contrasting retail and transport opportunities (Figure 1). For example, suburban neighbourhoods of Rexdale and West Hill have comparatively lower population density and more car-oriented transport environments compared with the more urban neighbourhood of Parkdale.

Participants completed a questionnaire detailing sociodemographic information and feelings of time pressure and health (n=125). Participants were also invited to download the FASTT smartphone application which captured their location data using GPS for eight days (n=106). Seven days of time use diary data were collected concurrently (n=90) whereby primary and secondary activities, location and/or transport mode, and company were reported for ten-minute intervals for seven consecutive days. Participants were free to describe activities in their own words. After the diaries were returned to the research team, research assistants coded activities into standardised categories which included accompanying children to school and household members to activities.

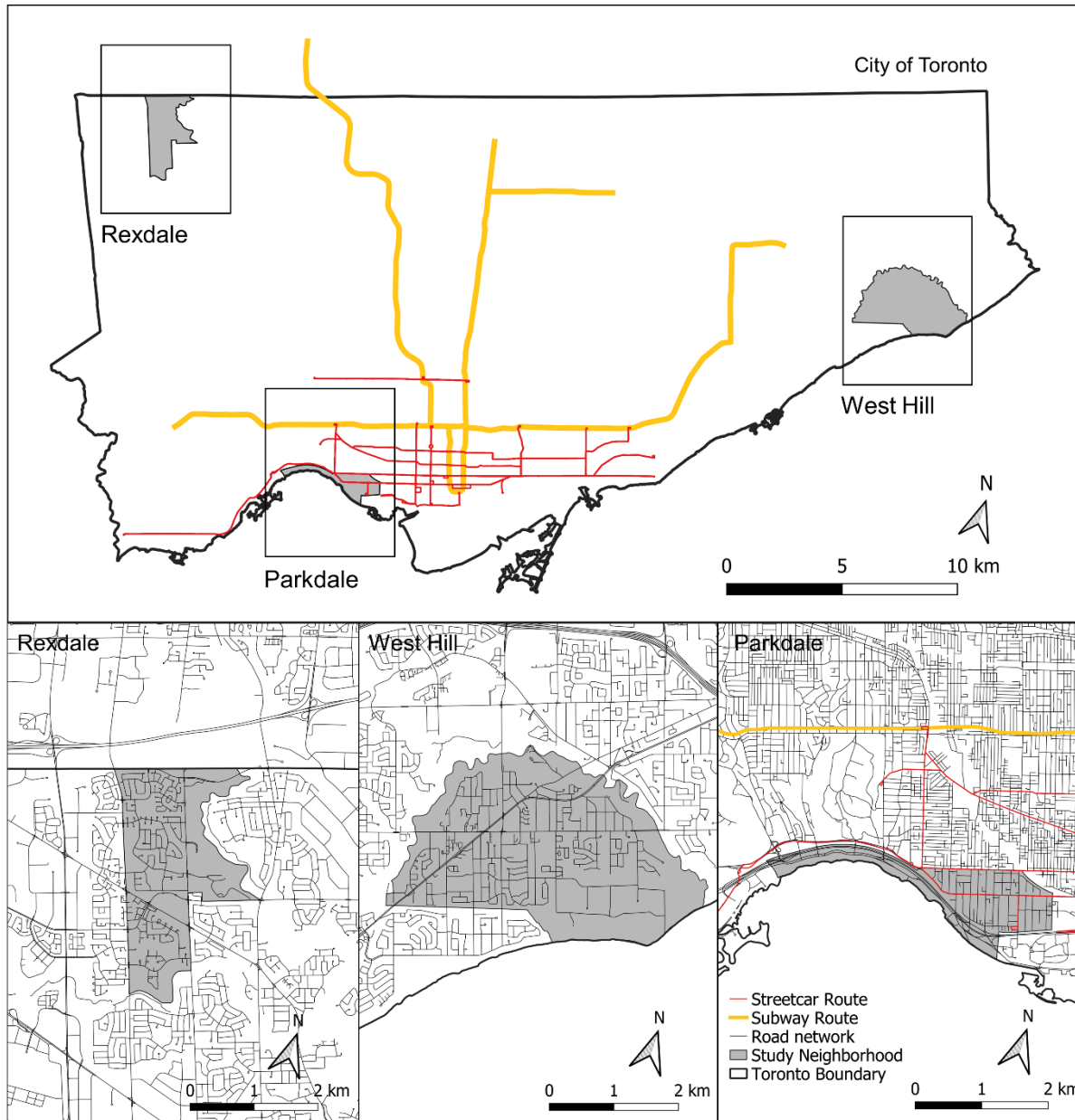


Figure 1. City of Toronto and neighbourhoods of Rexdale, West Hill, and Parkdale (adapted from upcoming publication by Liu et al.)

Participants were included for analysis if they had questionnaire data, three days’ worth of GPS data with a minimum of eight hours, and time use diary data which included travel for care activities (n=54). To extract trips made for care from the data, GPS point data were first converted into linear trips using the Itinerum-tripkit library (TRIP Lab, 2019). GPS point data were then matched to time use diary entries based on timestamp so that all GPS points occurring within a ten-minute interval were matched to the same reported activity. Attributes were queried to select GPS points where the primary activity involved travel for care. Information regarding the trip destination, travel mode, and the nature of the trip; whether a round trip or multi-activity trip were exported. Selected points for each trip were then cross referenced with the linear trips to identify distances travelled.

3. Preliminary findings and study implications

Descriptive analysis and preliminary findings from the study showed that the majority of participants performing trips for care were women, however, there were no clear differences in time spent travelling for care by gender within the sample and between household members. Residents in the more urban neighbourhood of Parkdale appeared to report the most time spent travelling for care and caregiving, however, investigation into relationships between care trips, time spent caring, and health outcomes is ongoing. Analysis also being undertaken to assess differences in distances travelled, transport mode, and whether trips were multipurpose by gender and residential neighbourhood.

An important methodological advance of the study is the high level of detail within the dataset which allows for trips made specifically for care to be captured. The freedom for participants to describe their activities within the time use diaries meant that participants were not constrained by categories. Combined with objective GPS monitoring, a clear picture of how trips for care are integrated into daily routines and coordinated with household members can be provided. Presently, there is a need to better represent trips for care in research. Once complete, this study will provide a step towards more closely reflecting the mobility and specific transport needs of men and women. Examples of spatial patterning and coordination of multipurpose trips between household members will also be highlighted. Adding to the knowledge base of travel patterns and emphasising care as an important topic for transport policy may help to inform more equitable transport design and direct policies which address gender gaps.

Acknowledgments

The Food Activities, Socioeconomics, Time, and Transportation (FASTT) Study is supported in part by funding from the Social Sciences and Humanities Research Council. In addition, this research was undertaken, in part, thanks to funding from the Canada Research Chairs Program and the Ontario Early Researcher Award. We are grateful to the research assistants for their help with data collection and to the study participants.

References

- Bianchi, S. M. *et al.* (2012) 'Housework: Who did, does or will do it, and how much does it matter?', *Social Forces*. doi: 10.1093/sf/sos120.
- Bianchi, S. M. and Milkie, M. A. (2010) 'Work and family research in the first decade of the 21st century', *Journal of Marriage and Family*. doi: 10.1111/j.1741-3737.2010.00726.x.
- Carmichael, F., Charles, S. and Hulme, C. (2010) 'Who will care? Employment participation and

willingness to supply informal care’, *Journal of Health Economics*, 29(1). doi: 10.1016/j.jhealeco.2009.11.003.

Chick, K. A., Heilman-Houser, R. A. and Hunter, M. W. (2002) ‘The impact of child care on gender role development and gender stereotypes’, *Early Childhood Education Journal*, 29(3). doi: 10.1023/A:1014528424032.

Jain, J., Line, T. and Lyons, G. (2011) ‘A troublesome transport challenge? Working round the school run’, *Journal of Transport Geography*, 19(6). doi: 10.1016/j.jtrangeo.2011.04.007.

Lee, Y., Hickman, M. and Washington, S. (2007) ‘Household type and structure, time-use pattern, and trip-chaining behavior’, *Transportation Research Part A: Policy and Practice*, 41(10). doi: 10.1016/j.tra.2007.06.007.

De Madariaga, I. S. (2016) ‘Mobility of care: Introducing new concepts in urban transport’, in *Fair Shared Cities: The Impact of Gender Planning in Europe*. doi: 10.4324/9781315581835-9.

Moyser, M. and Burlock, A. (2018) *Time use: Total work burden, unpaid work, and leisure, Women in Canada: A gender-based statistical report*.

Stanfors, M., Jacobs, J. C. and Neilson, J. (2019) ‘Caregiving time costs and trade-offs: Gender differences in Sweden, the UK, and Canada’, *SSM - Population Health*, 9. doi: 10.1016/j.ssmph.2019.100501.

Strazdins, L. et al. (2011) ‘Time scarcity: another health inequality?’, *ENVIRONMENT AND PLANNING A*, 43(3), pp. 545–559. doi: 10.1068/a4360.

TRIP Lab (2019) *Welcome to itinerum-tripkit’s documentation!* Available at: <https://itinerum-tripkit.readthedocs.io/en/latest/index.html#> (Accessed: 21 June 2021).

Westman, J., Friman, M. and Olsson, L. E. (2017) ‘What drives them to drive?-Parents’ reasons for choosing the car to take their children to school’, *Frontiers in Psychology*, 8(NOV). doi: 10.3389/fpsyg.2017.01970.

Wheatley, D. (2013) ‘Location, vocation, location? Spatial entrapment among women in dual career households’, *Gender, Work and Organization*, 20(6). doi: 10.1111/gwao.12005.

Special Session SS10.2

5 November 11:20 am -1:20 pm (GMT)

Sustainable Mobility and Equality in Mega-city Regions: Patterns, Mechanisms and Governance

Proponents

Joana Barros, Chen Zhong, Yang Yue

Delineating Functional Urban Areas in Chinese Mega City Regions Using Fine-grained Population Data and Cellphone Location Data: A Case of Pearl River Delta

Zifeng CHEN¹; Anthony G. O. YEH²

¹School of Geography and Planning, Sun Yat-sen University
Sun Yat-sen University (East Campus), Panyu, Guangzhou, 510000, China
+86-13544011526, czfurban@gmail.com (correspondent author)

²Department of Urban Planning and Design, The University of Hong Kong
Pokfulam Road, Hong Kong, 999077, Hong Kong SAR
+852-90933794, hdugoy@hkucc.hku.hk (correspondent author)

Keywords: functional urban area, mega city region, commuting, prefecture-level city, Pearl River Delta

A functional urban area (FUA) is composed of nearby urbanized areas that have intensive functional connections or linkages (Hall and Pain, 2006). FUAs as an analytical concept arose in the Western context of massive suburbanization in the post-war 20th century. Such a concept gains more and more popularity in recent decades as cities worldwide have become increasingly networked within broader regions due to the rise in producer services and fast development of transport and communication technologies (Scott, 2019; Yeh and Chen, 2020). Many developed countries have official definitions of FUAs. Scholars have also endeavored to (re)define FUAs to improve or supplement the official definitions.

The delineation of FUAs can serve as a practical solution to the limitations of city definition in China, which have long been recognized by scholars as a fundamental issue in the studies of Chinese urban system (Zhou and Ma, 2005; Chan, 2007). In China, many officially defined cities, usually referred to as “prefecture-level cities”, are de facto city-regions that comprise multiple clusters of urban areas separated and surrounded by rural areas. This is due to a policy called “city administering county” extensively implemented in the early 1980s to expand the territorial power of cities. Thus, in a strict sense, a prefecture-level city in China is not an equivalent of a city in the Western context. It cannot well represent the spatial extent of urban areas as the administrative area of a city could be much larger than urban areas. Nor can it manifest the functional connections between urban areas of different locations which could be functionally separate despite being placed in the same jurisdiction. The limitations of city definitions in China necessitate scholarly efforts to redefine “real” boundaries of

cities through delineation of FUAs in order to avoid biased interpretation of regional spatial interactions and to facilitate meaning city-regional governance.

So far, a few scholarly attempts to delineate FUAs in China have been documented, most of which adopted minor administrative divisions, such as jiedao and townships, to aggregate commuting data. The present study, using the Pearl River Delta (PRD) mega city region as a case study, extends the existing literature by utilizing fine-grained and uniform spatial unit (i.e., 1 km grid cells) to aggregate commuting flows that are derived from cellphone location data of 2 weekdays. Central grid cells are identified on the basis of population density that is derived from fine-grained population data. Contiguous central grid cells are grouped to form the core areas of FUAs. On this basis, we identify periphery grid cells of FUAs on the basis of a 25% commuting threshold. An inter- and intra-FUA connectivity check is further conducted to avoid insufficient or excessive demarcation of FUAs.

A total of 26 FUAs in the PRD mega city region are identified through this approach. The population is unequally distributed among these FUAs and concentrated in the largest two FUAs, namely Greater Shenzhen and Guangzhou-Foshan. Within each FUA, the core area tends to be more densely populated than the periphery area in a significant manner. The spatial extent of FUAs is found largely inconsistent with the administrative boundaries of prefecture-level cities, which is somewhat comparable to the spatial manifestation of the 23 UAs in the San Francisco Bay Area that comprises 9 counties.

The novelty of the proposed method is that the delineation is totally independent of administrative divisions. Like most developing countries, geographic units used in demographic and economic statistics in China (e.g., jiedao) tend to be much larger than people's home neighborhoods or workplaces. In the present study, commuting flows are aggregated in the spatial units of 1 km grid cells to investigate the functional connections among different locations. A 1 km grid cell is smaller than most of the existing administrative divisions. Using 1 km grid cells as the spatial unit is conducive to delineating FUAs more precisely.

The delineation of FUAs can help to address the critical problems associated the city definitions in China. It unravels that prefecture-level cities are inequivalent to FUAs, in the way that some of these prefecture-level cities have multiple FUAs that are functionally separate from each other, while some FUAs may surpass the administrative borders and be shared by two or more prefecture-level cities. Therefore, the boundaries of prefecture-level cities cannot well represent the spatial extent of urban areas and the functional connections between urban areas of different locations. This finding draws attention to the use of reliable spatial units in the studies of regional spatial interactions.

The problems of city definitions in China, as highlighted in the present study through the delineation of FUAs, also suggest that urban planning and governance in China will continually be challenged by jurisdictional fragmentation. While some FUAs are found overflowing the administrative boundaries

and being shared by two or more prefecture-level cities, urban planning and governance are mostly delimited by administrative boundaries, thus failing to cope with problems outside their administrative boundaries. This situation necessitates efforts at the regional scale to enhance inter-government cooperation of different prefecture-level cities to address the problems that happen in their shared FUAs.

The proposed method is applicable not only to China but also to other developing countries since the geographic units used in demographic and economic statistics tend to be very large and commuting data are officially not available in most developing countries.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant Number 71961137003 and 41971408; and Chan To-Haan Endowed Professorship Fund of the University of Hong Kong under Grant Number HKU-207060183.

References

- Chan, K.W., 2007. Misconceptions and complexities in the study of China's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics*, 48(4), pp.383-412.
- Hall, P.G. and Pain, K. eds., 2006. *The Polycentric Metropolis: Learning from Mega-city Regions in Europe*. London and New York: Routledge.
- Scott, A.J., 2019. City-regions reconsidered. *Environment and Planning A: Economy and Space*, 51(3), pp.554-580.
- Yeh, A.G.O. and Chen, Z., 2020. From cities to super mega city regions in China in a new wave of urbanisation and economic transition: Issues and challenges. *Urban Studies*, 57(3), pp.636-654.
- Zhou, Y. and Ma, L.J., 2005. China's urban population statistics: A critical evaluation. *Eurasian Geography and Economics*, 46(4), pp.272-289.

Measuring the Polycentricity based on urban and intercity transportation networks in Greater Bay Area: a cross-scale method in the context of Node-Place model

Haocheng Sun¹; Michael Batty²; Chen Zhong^{1,2}

¹University College London, Centre for Advanced Spatial Analysis (CASA)

Gower Street, London WC1E 6BT, United Kingdom

ucfnhsu@ucl.ac.uk (correspondent author)

² University College London, Centre for Advanced Spatial Analysis (CASA)

90 Tottenham Court Road, London W1T 4TJ, United Kingdom

44-(0)7768 423 656, m.batty@ucl.ac.uk

^{1,2} University College London, Centre for Advanced Spatial Analysis (CASA)

Gower Street, London WC1E 6BT, United Kingdom

c.zhong@ucl.ac.uk

Keywords: GBA, Spatial Structure, Morphological Polycentricity, Functional Polycentricity, Node-Place Model

For a long time, the study of urban spatial structure, especially the relationship between cities, has been regarded as an important research field in urban planning and regional science. The study of urban spatial structure can be traced back to the beginning of the 20th century, which mainly originated from the theory of urban location (Alonso, 1964). As time enters the 21st century, the concept of monocentricity has gradually given way to polycentricity (Green, 2007; Meijers and Burger, 2010; Batty, 2016). There are various signs that people have perceived that polycentricity is happening in the city, but in fact this concept is vague (Burger and Meijers, 2011a; Brezzi and Veneri, 2014; Giffinger and Suitner, 2014). This ambiguity is mainly reflected in the cognition of the centre and spatial structure.

The Guangdong-Hong Kong-Macao Greater Bay Area (GBA) plan was promulgated by the State Council of China in 2015 (Walker and Schafran, 2015). As one of China's three Mega-City Clusters, GBA is considered an important innovation platform for national implementation of resource allocation, coordination, and division of labour (Hui et al., 2020). It is foreseeable that the spatial structure of GBA will gradually move closer to polycentricity in the future. Therefore, exploring the current spatial structure of GBA will provide a strong basis for a series of urban construction measures in the future. This is the primary motivation of this research from the perspective of practical application.

Therefore, this research will use the Node-Place model proposed by Bertolini in 1999 as an entry point to fill the above-mentioned research gaps in this field. This research proposes an extended model defined as Node-Place-Settlement-Container+Mobility (NPSC+M) to evaluate the different dimensions of urban station areas. Among them, the ‘NPSC’ dimensions are used to evaluate the morphological spatial structure, and the ‘M’ dimension is used to evaluate the functional polycentricity. We provide a multi-dimensional and comprehensive multi-index polycentricity evaluation framework based on this extended model, and further analyse the connections of morphological polycentricity and functionality polycentricity.

1. Visualize and analyse the spatial structure of GBA

In order to quantitatively evaluate the morphological polycentricity and functional polycentricity of GBA and its major cities, this paper processes the collected data to construct 19 indicators in five dimensions of NPSC+M. On this basis, the indicators are weighted by the CRITIC method and integrated into the index of each dimension. Finally, through Global Moran’s I, we put forward a quantitative evaluation result of the Centralized-Dispersed degree for each city in GBA, and visualized their morphological and functional spatial structure through the LISA spatial autocorrelation method to evaluate their polycentricity. We found that currently only Shenzhen in GBA meets the definition of morphological polycentricity and functional polycentricity (Figure 1 and 2). Guangzhou is currently in the process of transforming from monocentricity to polycentricity. Compared with the former two, Foshan and Dongguan have weaker urban competitiveness. This results in them still being monocentric or even dispersed. At the regional level, GBA already possesses morphological polycentricity (although only Guangzhou and Shenzhen), but the flow of functions tends to gather in Shenzhen.

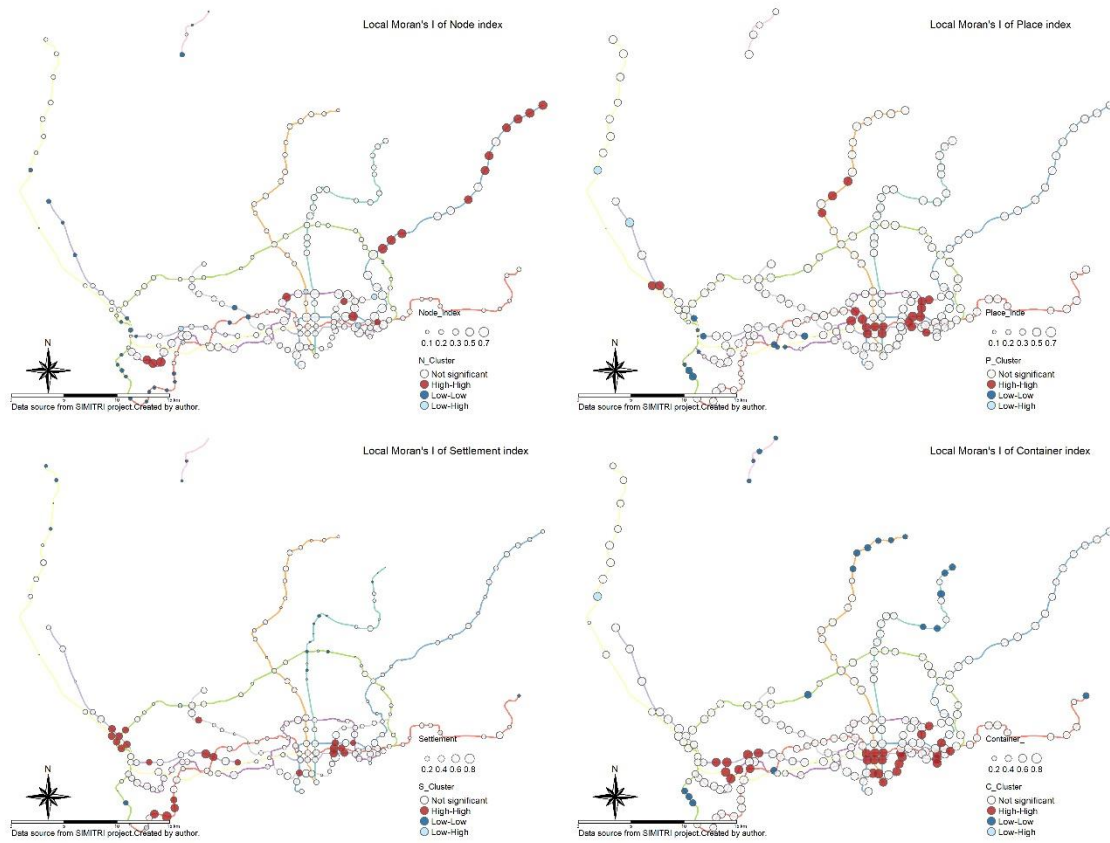


Figure 1: LISA maps of morphological spatial results in Shenzhen, based on the k2 matrix; 1) Node dimensional; 2) Place dimensional; 3) Settlement dimensional; 4) Container dimensional morphological spatial structure

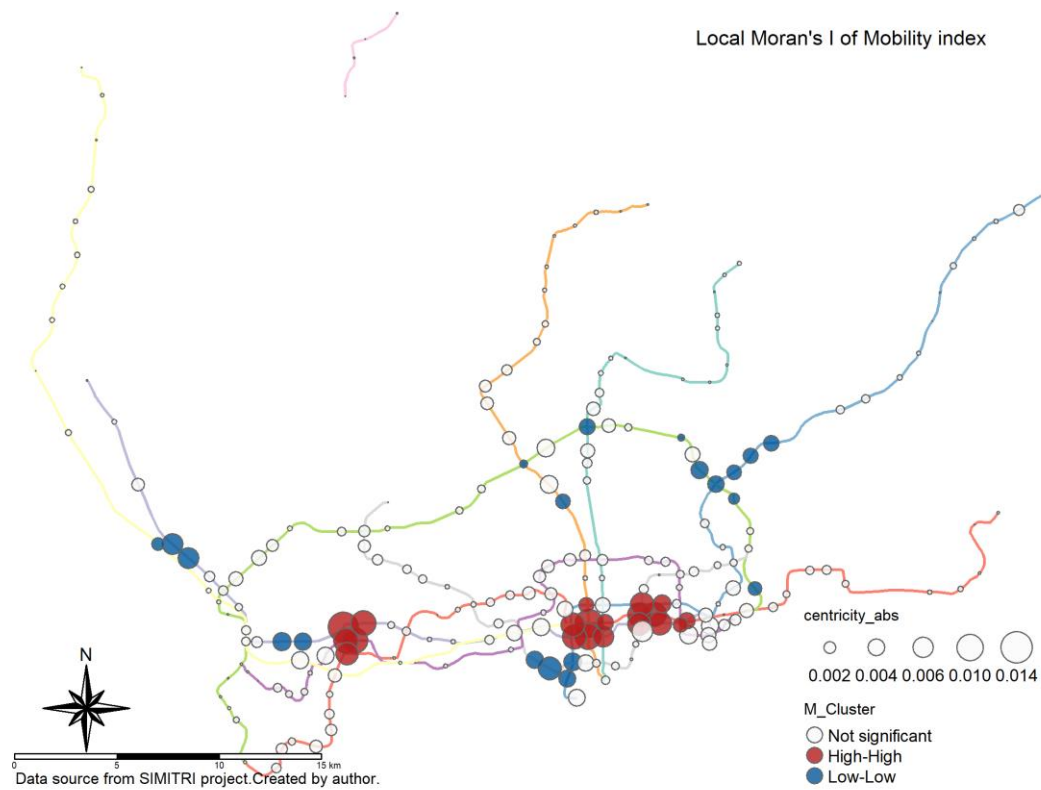
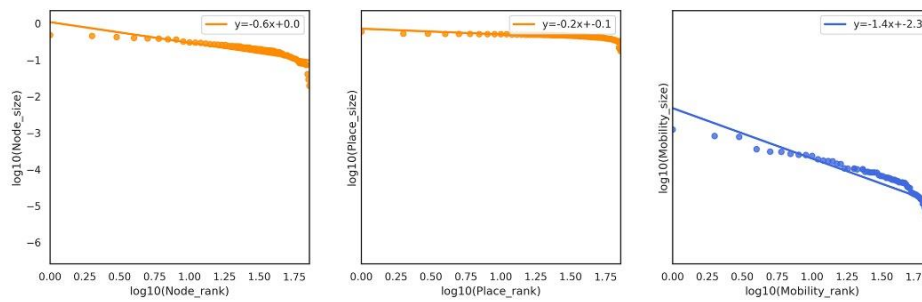
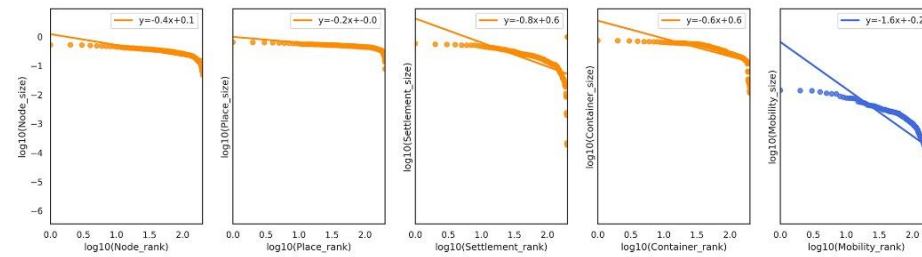


Figure 2: LISA map of Mobility dimensional results in Shenzhen, based on the k2 matrix

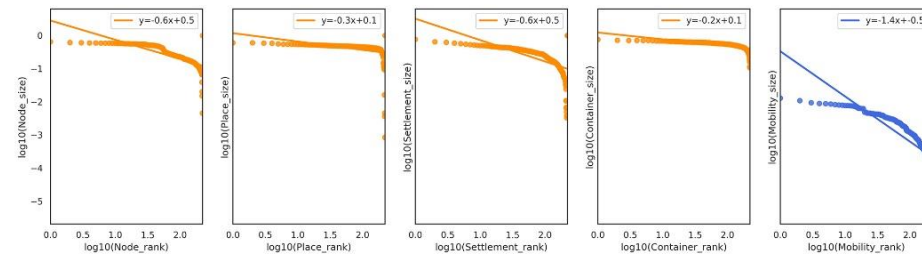
1) Rank-size distribution of Node, Place and Mobility dimensions in GBA



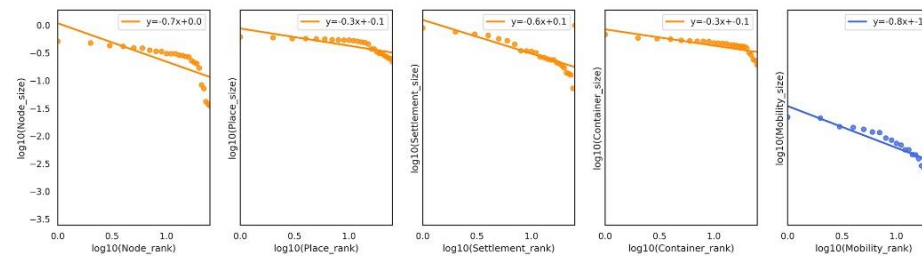
2) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Guangzhou



3) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Shenzhen



4) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Foshan



5) Rank-size distribution of Node, Place, Settlement, Container and Mobility dimensions in Dongguan

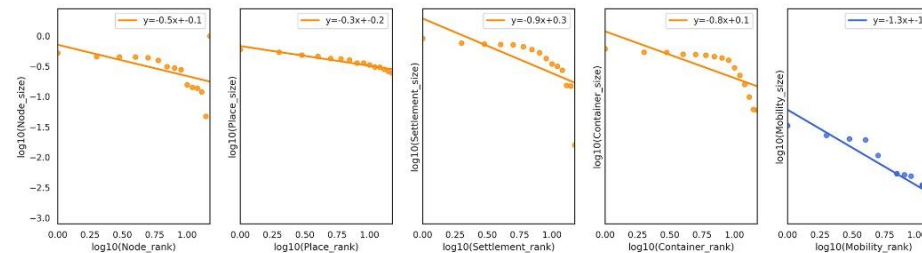
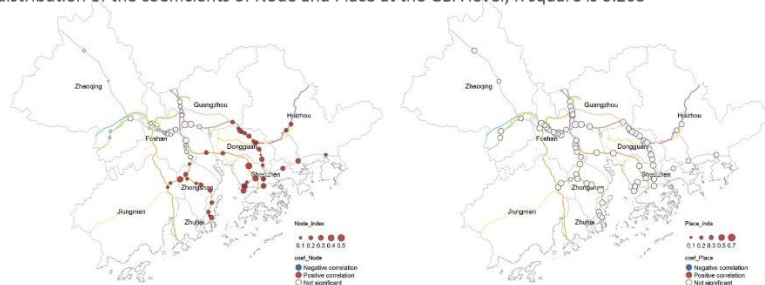
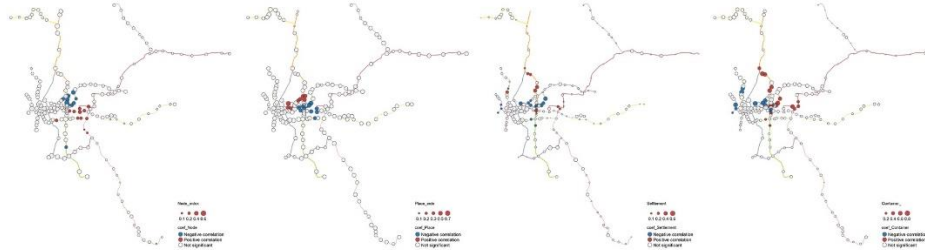


Figure 3: Rank-size distribution of 1) GBA, 2) Guangzhou, 3) Shenzhen, 4) Foshan, 5) Dongguan in NPSC+M dimensions;

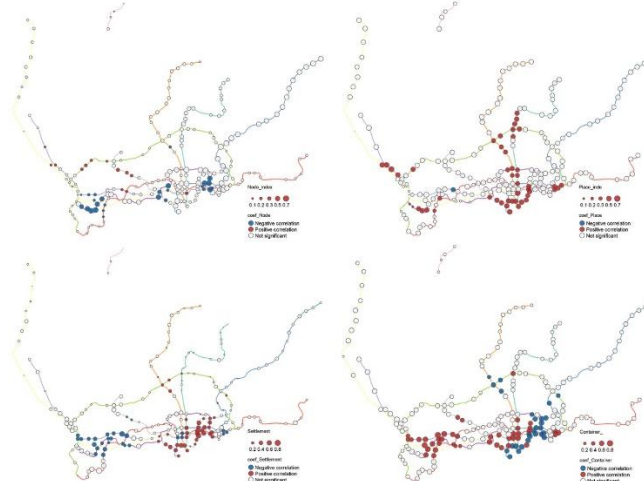
1) The distribution of the coefficients of Node and Place at the GBA level; R square is 0.263



2) The distribution of the coefficients of Node, Place, Settlement and Container at the Guangzhou level; R square is 0.918



3) The distribution of the coefficients of Node, Place, Settlement and Container at the Shenzhen level; R square is 0.924



4) The distribution of the coefficients of Node, Place, Settlement and Container at the Foshan level; R square is 0.467



5) The distribution of the coefficients of Node, Place, Settlement and Container at the Dongguan level; R square is 0.596

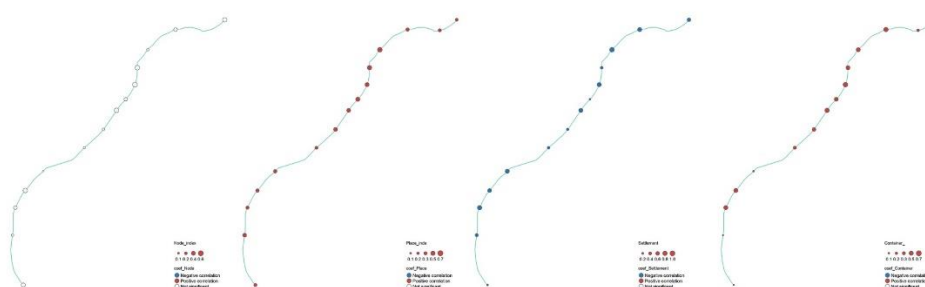


Figure 4: The coefficients of the GWR model are distributed in 1) GBA, 2) Guangzhou, 3) Shenzhen, 4) Foshan, 5) Dongguan. Node, Place, Settlement, Container are used as independent variables, and Mobility is used as dependent variable.;

2. Comparison of morphological polycentricity and functional polycentricity in GBA

On this basis, we further studied the relationship and mismatch between morphological polycentricity and functional polycentricity in GBA. For the former, we chose the Rank-size distribution method which has been confirmed by Meijers (2011) to be feasible. For visualizing the mismatch of morphological polycentricity and functional polycentricity, we tried to use geographically weighted regression (GWR) and confirmed the contribution of this method in this field. we made some conclusions from the results (Figure 3 and 4):

- 1) The development of physical space of GBA is more balanced than the development of functional connection, which is completely contrary to the conclusions obtained in the United Kingdom and the United States (Burger *et al.*, 2011; Arribas-Bel and Sanz-Gracia, 2014).
- 2) Morphological polycentricity and functional polycentricity have been shown to be interdependent. This finding also verifies the conclusion of Meijers (2011);
- 3) For regional sub-central cities (Foshan, Dongguan), Horizontal land use is the key to stimulating functional flow. For regional central cities (Guangzhou, Shenzhen), the flow of functions has a closer relationship with the urban spatial form.
- 4) Focusing on mismatches, the urban centres of Shenzhen and Guangzhou show that morphological elements are not enough to support high- strength functional flow. The main problem in Foshan is that the residential space dominates the vertical spatial form of the city. There is a serious imbalance between the distribution of communities in Dongguan and the flow of population.

In general, the concept of polycentricity is still vague (Möck and Küpper, 2019). Nevertheless, we have boldly attempted to combine the Node-Place model with the polycentricity evaluation and respond to the ambiguity with a fixed spatial unit (station) and multi-scale evaluation method. However, our current method still has many limitations, such as the ambiguity caused by the choice of analysis space unit. We encourage other scholars to look at polycentricity from the perspective of ‘evaluation’ rather than ‘identify’. In this way, not only will polycentricity not become a far-fetched and meaningless concept, but it will also help the city's cognition and reality transformation.

Acknowledgments

The first author would like to express his gratitude to his supervisor, the second author, for his continuous teaching and help. This article benefits from SIMETRI project.

References

- Alonso, W. (1964). 'Location and Land Use'. *Location and Land Use*. Harvard University Press. doi: 10.4159/HARVARD.9780674730854/HTML.
- Arribas-Bel, D. and Sanz-Gracia, F. (2014). 'The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010'. <http://dx.doi.org/10.1080/02723638.2014.940693>. *Routledge*, 35 (7), pp. 980–997. doi: 10.1080/02723638.2014.940693.
- Batty, M. (2016). 'Polynucleated Urban Landscapes': <http://dx.doi.org/10.1080/00420980120035268>. Sage PublicationsSage UK: London, England, 38 (4), pp. 635–655. doi: 10.1080/00420980120035268.
- Brezzi, M. and Veneri, P. (2014). 'Assessing Polycentric Urban Systems in the OECD: Country, Regional and Metropolitan Perspectives'. <http://dx.doi.org/10.1080/09654313.2014.905005>. *Routledge*, 23 (6), pp. 1128–1145. doi: 10.1080/09654313.2014.905005.
- Burger, M. J., de Goei, B., van der Laan, L. and Huisman, F. J. M. (2011). 'Heterogeneous development of metropolitan spatial structure: Evidence from commuting patterns in English and Welsh city-regions, 1981–2001'. *Cities*. Pergamon, 28 (2), pp. 160–170. doi: 10.1016/J.CITIES.2010.11.006.
- Burger, M. and Meijers, E. (2011a). 'Form Follows Function? Linking Morphological and Functional Polycentricity': <http://dx.doi.org/10.1177/0042098011407095>. SAGE PublicationsSage UK: London, England, 49 (5), pp. 1127–1149. doi: 10.1177/0042098011407095.
- Burger, M. and Meijers, E. (2011b). 'Form Follows Function? Linking Morphological and Functional Polycentricity': <http://dx.doi.org/10.1177/0042098011407095>. SAGE PublicationsSage UK: London, England, 49 (5), pp. 1127–1149. doi: 10.1177/0042098011407095.
- Giffinger, R. and Suitner, J. (2014). 'Polycentric Metropolitan Development: From Structural Assessment to Processual Dimensions'. <http://dx.doi.org/10.1080/09654313.2014.905007>. *Routledge*, 23 (6), pp. 1169–1186. doi: 10.1080/09654313.2014.905007.
- Green, N. (2007). 'Functional Polycentricity: A Formal Definition in Terms of Social Network Analysis'. *Urban Studies*. *Routledge*, 44 (11), pp. 2077–2103. doi: 10.1080/00420980701518941.
- Meijers, E. J. and Burger, M. J. (2010). 'Spatial Structure and Productivity in US Metropolitan Areas': <http://dx.doi.org/10.1068/a42151>. SAGE PublicationsSage UK: London, England, 42 (6), pp. 1383–1402. doi: 10.1068/A42151.
- Möck, M. and Küpper, P. (2019). 'Polycentricity at its boundaries: consistent or ambiguous?' <https://doi.org/10.1080/09654313.2019.1666802>. *Routledge*, 28 (4), pp. 830–849. doi: 10.1080/09654313.2019.1666802.

The evaluation of intra-urban developments through urban scaling indicators: empirical evidence from the Greater Bay Area, China

Jinzhou Cao^{1,2}, Wei Tu^{1,2}, Chen Zhong³, Qili Gao^{1,2}, Qingquan Li^{1, 2}

1 Guangdong Key Laboratory of Urban Informatics, Shenzhen Key Laboratory of Spatial Smart Sensing & Research Institute of Smart Cities, Shenzhen University, Shenzhen 518060, China

2 Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518060, China

3 The Bartlett Centre for Advanced Spatial Analysis (CASA), University College London (UCL) Gower Street, London WC1E 6BT, UK

Key Words: Scaling laws; Urban indicators; Urban mobility; Network analysis; The Greater Bay Area

Introduction

The rapid urbanization has led to a rapid population growth in cities. Cities are typical complex systems, within which various urban elements including the population, infrastructures and information intricately interact with each other(Li et al., 2017). The efforts to reach a better understanding of urban systems call for a new perspective to quantitatively describe interdependent relations of these elements rather than examine them individually.

Scaling laws is one of beneficial but simple rule to help understand complex urban systems. The existence of scaling laws has been widely proved in complex systems and has been theoretical and empirical explained by various mechanism models(Batty, 2008). Scaling laws indicate trends

between two explainable indicators in systems of cities in a power-law-like regime $Y \sim P^\beta$, where β is the scaling exponent (Arcaute Elsa et al., 2015). The urban indicators describe the spatial configurations of urban elements, such as population distributions, infrastructural layouts, and levels of socio-economic activities. Urban scaling laws help quantify how urban indicators develop with the increase of urban population and assess the performance of a city.

Although considerable progress has been made in the study of urban scaling laws, however, as rapidly urbanizing cities, another crucial perspective in urban developments lacks understanding. Urban mobility describes residents movements and characterizes the networked interactions of travels between different intra-urban regions (Song et al., 2010). It is also essential for investigating the scaling laws in human mobility to characterize urban mobility patterns (Tang et al., 2016). Meanwhile, compared with the scaling laws in urban elements generally remaining stable for a period of time, human mobility patterns are quite dynamic and the temporal dynamics of scaling processes present key clues in the evolution of the urban systems.

Moreover, the spatial scale is another issue of determining the understanding for urban systems. Most of current studies of urban and mobility scaling laws treat the targeted cities as a whole and study the macroscopic scaling properties over a region or a country, thus lacks judgment on the difference of the micro-scale within the cities. It has been reported that the scaling laws also exists in the intra-urban system composed by districts or communities (Louail et al., 2015), however, an intra-urban scale around a few meters, which benefits more refined urban research and urban planning remains poorly understood.

In this study, we address the following two questions: whether the interaction among urban indicators at intra-urban scale around a few meters obey urban scaling laws; and what the scaling laws of the human mobility characterized by a series of statistical measures derived from the complex

network perspective followed. These two questions highlight the necessity of a comprehensive study to quantify the intra-urban scaling laws in terms of both urban indicators and human mobility at meters level. Considering the typical mega-city region in China, the Guangdong-Hong Kong-Macau Greater Bay Area (GBA) as the study area, we firstly selected urban population, building areas and socioeconomic activities as the urban elements at 500m regular grid level and quantified the scaling relationship between them and compared it with the paradigm derived from urban size. The urban mobility networks were then constructed by aggregating the averaged travel flows of all population recorded every month from 2017 to 2019 within the city and considered the temporal variation in scaling relations between degree, strength and clustering coefficient to evaluate the development of the urban systems.

Study Area and data

The Guangdong-Hong Kong-Macau Greater Bay Area (GBA), which is our study area, is one of the major bay areas in the world. The GBA is located in the southern coastal area of China, comprising the two Special Administrative Regions of Hong Kong and Macao and nine mainland cities in Guangdong Province. It, covering < 1% of China's land area, creates 30% of the whole of mainland China GDP in 2020 with only 5% of total population and is the most active mega-city region in the country. The GBA as a whole is indeed a complex urbanized region with population dynamics, intensive economic activities and innovative clusters.

To derive the quantitative relationship between urban indicators, our analysis synthesizes information from a variety of sources. We use four extensive data, including a vector building data, a road network data, a point of interests (POIs) data, and a mobile phone data. Detailed data descriptions are as follows.

Building data. The building data were crawled from a digital map in China.

Road network data. The data for road networks were downloaded from the OpenStreetMap (OSM) (<http://www.openstreetmap.org/>) in June 2020.

POIs data. We collected POIs data from GaoDe, the largest online map service in China. Here we extracted points of restaurants and shops for our analysis.

Mobile phone data. A granular mobile phone tracking data was acquired from a dominant communication operator in China. The dataset recorded long-period phone users' stay activities and their movements between signaling towers from 2017 to 2019. The positions of phone users had been recorded at the signaling tower level. The data reveal information in two aspects. One is the grid-cell data, such as the aggregated numbers of the user visits for every cell and certain period or the total home locations of users for each cell. The operator also provided another type of data, namely the travel flows between two certain areas per day. A travel-flow record includes origin-grid id, destination-grid id, and population count belongs. Based on these datasets, we construct various population measures and urban mobility networks over time to capture the population and flow dynamics.

Methodology and ongoing results

This study develops a quantitative approach to understanding the scaling laws in both urban elements and human mobility. Our approach involves the following major parts. The first part involves in the urban indicators. We will firstly produce various urban indicators from various data, such as total building area, total road length, to represent urban infrastructural development, and POIs of commercial facilities, to represent urban socio-economic activities.

In this study, we consider the diverse measures of gridded population as alternative simple registered population usually reported in the cities' official statistics to uncover a better measure of the urban population distribution. To do so, we employ three measures, i.e., the number of home locations of phone

users(NH), the number of work locations of phone users(NW), the number of stays of phone users(NS), which is more appropriate proxy for estimating socioeconomic activity. These three measures reflect a multi-dimensional portrait of dynamic urban population within a given region.

We then constructed urban mobility networks with different timings which will be used as the measures of scaling laws in human mobility. Specifically, a set of weighted directed networks was created by aggregating and summing the travel flows of all individuals at time t . We selected one week data from every month from June 2017 to June 2019 as the monthly network. Finally, a total of 36 networks represents monthly properties of urban mobility. We then employed a series of complex network-driven measures, including the node degree, node strength, and clustering coefficient, to characterize the urban mobility networks and define two types of scaling relationship among these measures: the scaling relation between degree and strength and the scaling relation between degree and clustering coefficient.

In the third part, given the detailed spatial distributions of urban indicators and network measures, the urban boundaries will be divided into a set of $500\text{m} \times 500\text{m}$ grid regular grids, then aggregate them into corresponding grid. A fitting procedure will be performed to derive the scaling exponents at last.

Based on these datasets, we will present various empirical observations. First, the results of fitted scaling laws in urban indicators and urban mobility and spatiotemporal evolution will be analyzed. By further exploring city differentiation of scaling laws and spatiotemporal variations of scaling coefficients, the interaction between urban indicators and urban mobility behaviour in different cities can be profoundly revealed. The results of the scaling relations demonstrated a multi-facet portrait of urban mobility networks, which provides crucial implications and policies for data-informed urban planning.

References

- Arcaute Elsa, Hatna Erez, Ferguson Peter, Youn Hyejin, Johansson Anders, & Batty Michael. (2015). Constructing cities, deconstructing scaling laws. *Journal of The Royal Society Interface*, 12(102), 20140745. <https://doi.org/10.1098/rsif.2014.0745>
- Batty, M. (2008). The Size, Scale, and Shape of Cities. *Science*, 319(5864), 769–771. <https://doi.org/10.1126/science.1151419>
- Li, R., Dong, L., Zhang, J., Wang, X., Wang, W.-X., Di, Z., & Stanley, H. E. (2017). Simple spatial scaling rules behind complex cities. *Nature Communications*, 8(1), 1841. <https://doi.org/10.1038/s41467-017-01882-w>
- Louail, T., Lenormand, M., Picornell, M., García Cantú, O., Herranz, R., Frias-Martinez, E., Ramasco, J. J., & Barthelemy, M. (2015). Uncovering the spatial structure of mobility networks. *Nature Communications*, 6, 6007. 83. <https://doi.org/10.1038/ncomms7007>
- Song, C., Koren, T., Wang, P., & Barabási, A.-L. (2010). Modelling the scaling properties of human mobility. *Nature Physics*, 6(10), 818–823. 80. <https://doi.org/10.1038/nphys1760>
- Tang, J., Zhang, S., Zhang, W., Liu, F., Zhang, W., & Wang, Y. (2016). Statistical properties of urban mobility from location-based travel networks. *Physica A: Statistical Mechanics and Its Applications*, 461, 694–707. <https://doi.org/10.1016/j.physa.2016.06.031>

High-speed rail development and access equity: the case of Guangdong province, China

Zhiyuan LI¹ ; Juste RAIMBAULT^{1,*}

¹Center for Advanced Spatial Analysis, University College London
Gower Street, London WC1E 6BT, United Kingdom

*j.raimbault@ucl.ac.uk

Keywords: High-speed Rail, Access Equity, Greater Bay Area

Submitted to Special Session SS10

In recent years, China has witnessed a nationwide and relatively fast development of high-speed rail (HSR) infrastructures, improving considerably accessibility between large cities. At regional scales, this however implies unbalanced development, as major poles were the priority of the 8+8 planning scheme at the national scale. In particular within mega-city regions, transport access equity becomes a central concern for future planning policies. This contribution focuses on the case of the Guangdong-Hong Kong-Macao Greater Bay Area, and more broadly on Guangdong province. Gathering socio-economic data and train timetables, we first build travel-time based and opportunity-based accessibility indices. We then estimate spatial interaction models using mobility data provided by Tencent, to build refined accessibility indices. These are applied to the current HSR network and to planning scenarios. We find that HSR considerably improved accessibility in most cities with average inter-city travel time reduced from 210 min to 168 min. This first development however increased polarisation around main corridors and increased access inequality. The mid-to-long term railway plan significantly flattens access differences and provides in particular a rebalancing between East and West of the region. This evolution can be understood as an example of transport network maturation at multiple spatial and temporal scales, and suggests a need for future research to understand the interplay between governance processes at different scales and the land-use transport interaction system.

A system dynamics model of urban mobility resilience when exposed to fuel-related threats

Daniela V. R. LARA¹; Antônio N. RODRIGUES DA SILVA²

¹Department of Transportation Engineering, São Carlos School of Engineering, University of São Paulo
Av. Trabalhador São-carlense, 400, 13566-590, São Carlos, SP, Brazil
+5516 3373-9595, daniela.lara@usp.br (correspondent author)

²Department of Transportation Engineering, São Carlos School of Engineering, University of São Paulo
Av. Trabalhador São-carlense, 400, 13566-590, São Carlos, SP, Brazil
+5516 3373-9595, anelson@sc.usp.br

Keywords: urban mobility resilience, fuel-related threats, system dynamics, qualitative modeling, causal loop diagram

1. Introduction

Urban mobility systems and their components are subject to disruptive events (threats), which can compromise their essential state of functionality and cause multiple negative impacts (Cantillo et al. 2019; Kim et al. 2018). When exposed to a threat, a resilient urban mobility system should be able to maintain or rapidly return to the desired functions (Meerow et al. 2016), to absorb the first damage, to reduce the impacts from a disturbance (Ribeiro and Pena Jardim Gonçalves 2019), to adapt to change, and to quickly transform systems that limit current or future adaptive capacity (Meerow et al. 2016; Ribeiro and Pena Jardim Gonçalves 2019).

Due to transport's dependence on fossil fuels, possible scenarios of scarcity (peak oil), unavailability, or variation in this resource price may affect the global economy, leading to higher food and goods prices, also impacting the transport sector (Fernandes et al. 2017). In addition, higher transport costs can have negative impacts on transport equity, leaving many people vulnerable to fluctuations in oil prices (Leung et al. 2018). The transport's dependence on fossil fuels also raises concerns about the worsening of greenhouse gas (GHG) emissions and the global climate crisis.

Most strategies to measure the resilience of urban mobility systems utilizes conceptual frameworks (Fernandes et al. 2017; Fernandes et al. 2019), indicator-based approaches (Martins et al. 2019; Azolin et al. 2020), and fuzzy logic models (Santos et al. 2020). However, holistic analyses considering key elements of transport systems and their interactions over space and time are not common. In addition, we observed a scarcity of studies considering the impacts of increased fuel prices, peak oil, and disrupted fossil fuel in several components of urban mobility systems (subsystems) and addressing the concurrent impacts of these threats on the system.

Given the above, developing an analysis framework that is capable of characterizing and quantifying the variables and dynamics involved in resilient urban mobility systems is foremost to prevent the systems from collapsing under shock events. Therefore, this study intends to answer two questions. What are the dynamics involved in urban mobility systems when exposed to fuel-related threats? How resilient are urban mobility systems in the face of fuel-related threats? For that, the objective of this research comprises the characterization of the resilience of urban mobility systems when exposed to fuel-related threats using qualitative system dynamics modeling.

2. Method

A framework for assessing the resilience of urban mobility when exposed to fuel-related threats using a system dynamics modeling (Sterman 2002) is proposed. A characterization of interrelations and interconnections was based on documentation of cause and effect relationships within and between subsystems. Next, a qualitative modeling process was developed, based on causal loop diagrams (CLDs), structure, and feedback processes under different conditions, namely: in the essential state of functionality, affected by a disruptive event, and still functioning due to its resilience.

3. Results

The present study comprises an extension of Lara and Rodrigues da Silva's (2021) study, which described and analyzed the dynamics involved in urban mobility systems when affected by environmental threats (such as climate change, flooding, and natural hazards).

3.1 Characterization of interrelations and interconnections

A comprehensive literature review indicated that urban mobility systems can be divided into eight subsystems: Institutional, Social, Economic, Material and energy flows, Infrastructure, Natural, Demand, and Transport mode.

3.2 Qualitative modeling: Essential state of functionality

The CLD regarding the urban mobility system in its essential state of functionality (Figure 1) corresponds to an archetype of urban mobility, in which *Social*, *Transport mode*, and *Institutional* subsystems have a greater influence over the diagram. Moreover, economic interests play an important role in directing transport policies, and measures to stabilize the dynamics involved in fuel consumption must be taken, as it is related to an aggravation of environmental issues and exposes the system to fuel restrictions.

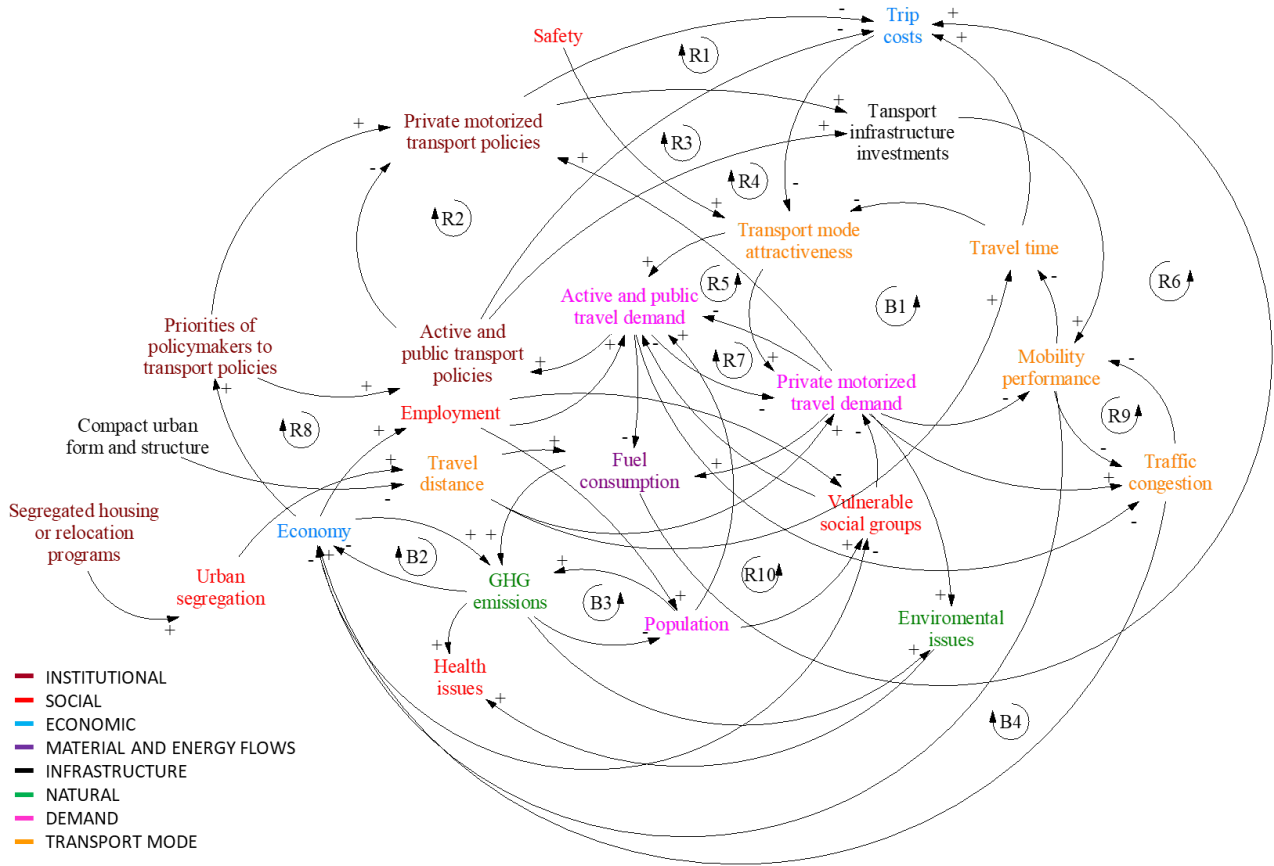


Figure 1: Causal loop diagram for urban mobility system in its essential state of functionality (Lara and Rodrigues da Silva 2021)

3.3 Qualitative modeling: Affected by fuel-related threats

The interrelations and connections among subsystems indicated that increased fuel price, peak oil, and oil vulnerability may concomitantly trigger a series of actions that cause the collapse of urban mobility systems. The CLD representing the system behavior when exposed to fuel-related threats (Figure 2) indicates that an increase in vulnerable social groups, in consumption, and price of fuels, and a decrease in fuel supply intensify the system's oil vulnerability. As the fuel supply decreases, the price increases, which in turn increases the system's vulnerability to the use of fossil fuels. Price and supply of fuels, when they are not part of the loops, are exogenous variables that cause the worsening of fuel consumption, economic problems and, consequently, weaken the support for technological strengths and fuel consumption reduction policies. Therefore, urban mobility systems exposed to fuel-related threats are driven by oil vulnerability, especially regarding fuel consumption. Furthermore, public awareness and willingness to change towards sustainable travel modes, technological strengths, and home office play an important role in stabilizing fuel consumption, trip costs, private motorized travel demand, and impacts on the economy.

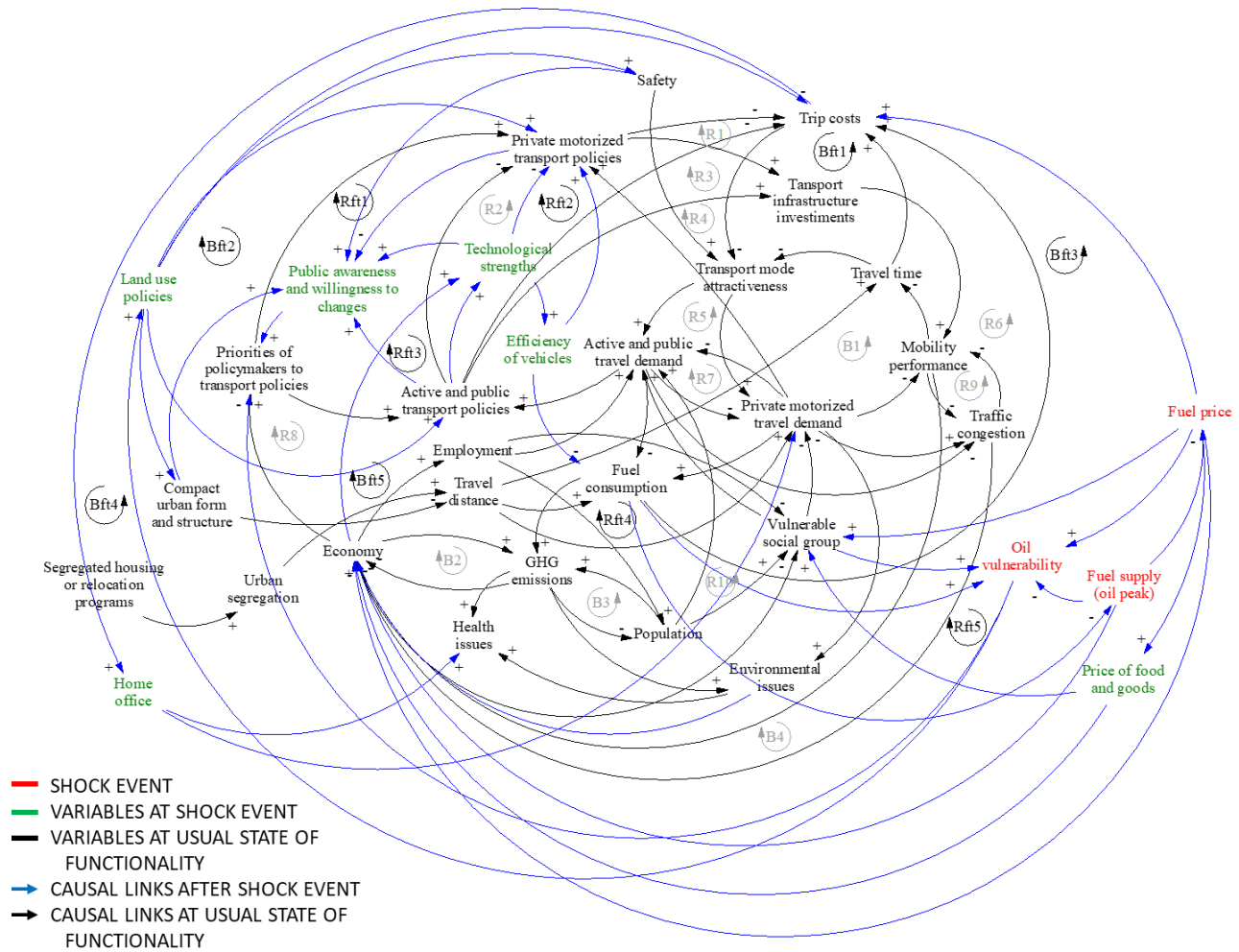


Figure 2: Causal loop diagram for urban mobility systems under fuel-related threats

3.4 Qualitative modeling: Resilient

In general, a resilient urban mobility system (as shown in Figure 3) should have attributes that allow the system to maintain its current conditions (persistence), to absorb the first damage and reduce impacts (adaptability), or to adapt to change and to quickly transform subsystems (transformability) after a shock event. The CLD for a resilient system has a preventive pattern, in which technological strengths, land use policies, and home office provide for the mitigation of fuel-related threats. In addition, public opinion plays an important role in balancing the system, as well as encouraging the use of sustainable transport.

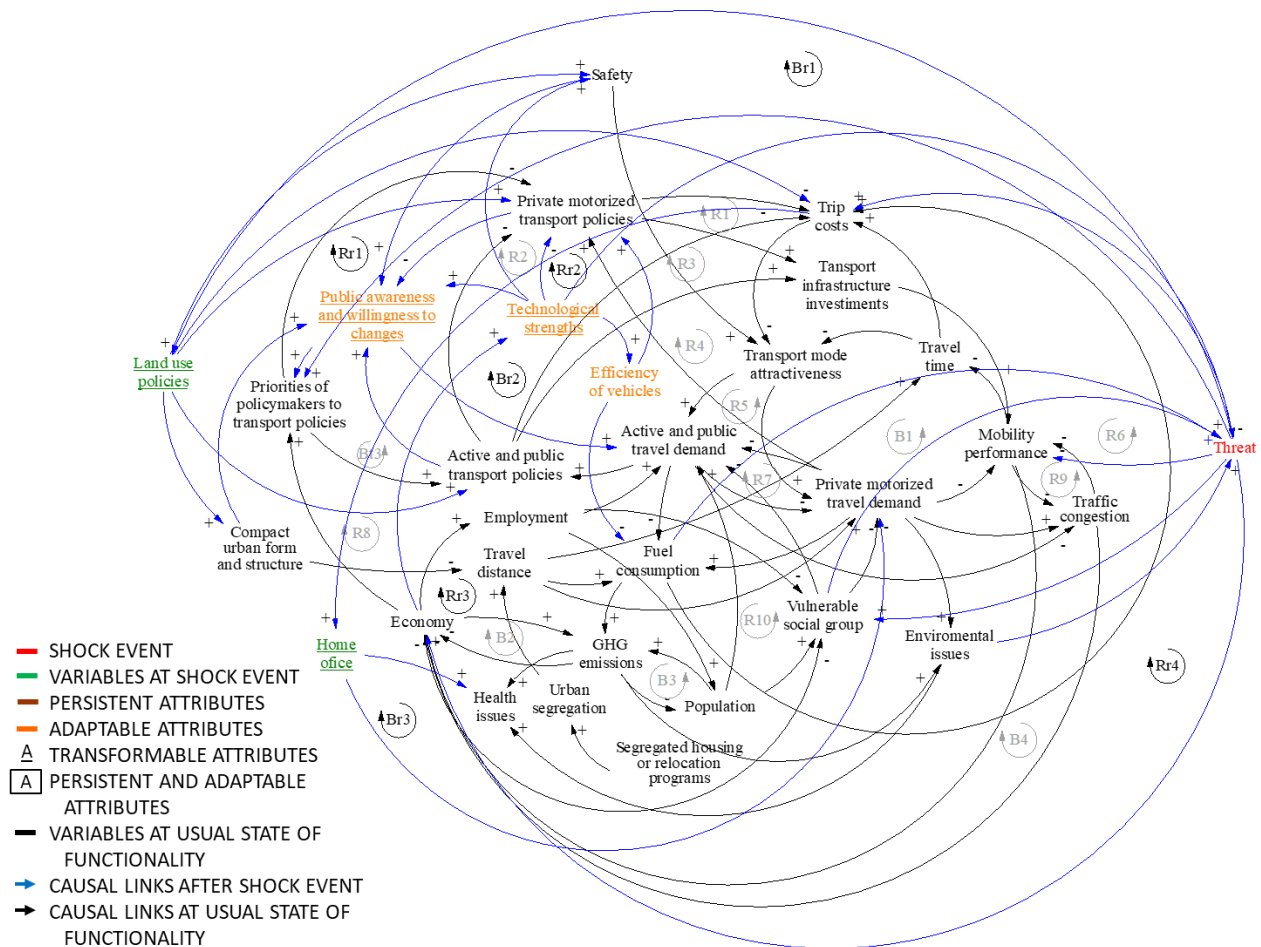


Figure 3: Causal loop diagram corresponding to a resilient urban mobility system

4. Conclusions

The study characterized the dynamics involved in a resilient urban mobility system when exposed to fuel-related threats using qualitative system dynamics modeling. This has the potential to help to develop holistic quantitative models, in which the consequences of policies and actions carried out in one part of the system can be identified and predicted in other parts of the system. Besides answering the questions formulated, this study explored in-depth holistic qualitative models. However, the models do not consider the dynamics over time.

The outcomes of this study may be used to support the process of planning and building resilient cities. For future investigations, we suggest extending this study for quantitative models based on stock and flow diagrams and validating them through case studies, for example.

References

Azolin, L.G., Rodrigues da Silva, A.N. and Pinto, N. (2020). Incorporating public transport in a methodology for assessing resilience in urban mobility. *Transportation Research Part D: Transport and Environment*, 85, p.102386. <https://doi.org/10.1016/j.trd.2020.102386>

- Cantillo, V., Macea, L.F. and Jaller, M. (2019). Assessing vulnerability of transportation networks for disaster response operations. *Networks and Spatial Economics*, 19(1), pp.243-273. <https://doi.org/10.1007/s11067-017-9382-x>
- Fernandes, V.A. et al. (2017). Resiliência da mobilidade urbana: uma proposta conceitual e de sistematização. *Revista Transportes*, 25(4), pp.147-160.
- Fernandes, V.A. et al. (2019). Urban resilience in the face of fossil fuel dependency: The case of Rio de Janeiro's urban mobility. *URBE-Revista Brasileira de Gestão Urbana*, 11. <https://doi.org/10.1590/2175-3369.011.e20180160>
- Kim, K., Francis, O. and Yamashita, E. (2018). Learning to build resilience into transportation systems. *Transportation Research Record*, 2672(1), pp.30-42. <https://doi.org/10.1177/0361198118786622>
- Lara, D.V.R. and Rodrigues da Silva, A.N. (2021). Resilience of urban mobility systems: Combining urban subsystems and threats with a system dynamics approach. In O. Gervasi et al., eds. *Computational Science and Its Applications. ICCSA 2021. Lecture Notes in Computer Science*. Springer, Cham, pp.93-108. https://doi.org/10.1007/978-3-030-87010-2_7
- Leung, A., Burke, M. and Cui, J. (2018). The tale of two (very different) cities - Mapping the urban transport oil vulnerability of Brisbane and Hong Kong. *Transportation Research Part D: Transport and Environment*, 65, pp.796-816. <https://doi.org/10.1016/j.trd.2017.10.011>
- Martins, M.C. da M., Rodrigues da Silva, A.N. and Pinto, N. (2019). An indicator-based methodology for assessing resilience in urban mobility. *Transportation Research Part D: Transport and Environment*, 77, pp.352-363. <https://doi.org/10.1016/j.trd.2019.01.004>
- Meerow, S., Newell, J.P. and Stults, M. (2016). Defining urban resilience: A review. *Landscape and Urban Planning*, 147, pp.38-49. <https://doi.org/10.1016/j.landurbplan.2015.11.011>
- Ribeiro, P.J.G. and Pena Jardim Gonçalves, L.A. (2019). Urban resilience: A conceptual framework. *Sustainable Cities and Society*, 50, p.101625. <https://doi.org/10.1016/j.scs.2019.101625>
- Santos, T. et al. (2020). Resilience and vulnerability of public transportation fare systems: The case of the city of Rio de Janeiro, Brazil. *Sustainability*, 12(2). <https://doi.org/10.3390/su12020647>
- Sterman, J.D. (2002). *System dynamics: systems thinking and modeling for a complex world*. (ESD-WP-2003-01.13; Working Paper Series).

Parallel Session PSF2

5 November 11:20 am -1:20 pm (GMT)

Health and Well-being

Exposure to air pollution and breast cancer risk: taking daily mobility into account in epidemiological studies

Marie RAMEL-DELOBEL^{1,2,3}; Thomas COUDON^{1,2,3}; Lény GRASSOT^{1,2}; Delphine PRAUD^{1,2}; Béatrice FERVERS^{1,2}; Pietro SALIZZONI³

¹Département Prévention, Cancer, Environnement, Centre Léon Bérard (CLB)

Cheney C, 28 rue Laennec, 69008 Lyon, France

+33 6 70 30 40 30, marie.ramel-delobel@lyon.unicancer.fr

²INSERM UMR 1296 Radiations, Défense, Santé, Environnement, CLB

Cheney A, 28 rue Laennec, 69008 Lyon, France

³CNRS UMR 5509 LMFA, Ecole Centrale de Lyon

36 avenue Guy de Collongue, 69132 Ecully, France

Keywords: urban geography, daily mobility, atmospheric pollution, breast cancer, NO₂, PM

1. Context

1.1 Air pollution exposure in epidemiological studies about cancer

Breast cancer is the most common cancer in women worldwide. In France, the incidence of this type of cancer has more than doubled in the last 30 years to 48,800 new cases in 2012 (Binder-Foucard et al., 2014). This increase is associated with widespread screening, menopausal hormone treatments or lifestyle changes. However, the role of environmental factors is also suspected. Several epidemiological studies suggest a link between long-term exposure to ambient air pollution and breast cancer risk (Rodgers et al., 2018; White et al., 2018).

Exposure to air pollutants occurs throughout the day, at home, at work, at other activities and during daily mobility (Reis et al., 2018; Steinle et al., 2013). However, due to a general lack of information on subject's activity and mobility patterns, epidemiological studies on air pollution usually estimate exposure based on their residential address. This can lead to inaccurate estimates (Basagaña et al., 2013).

1.2 The APoPCo project (*Atmospheric Pollution and Physical activity linked with Commute*)

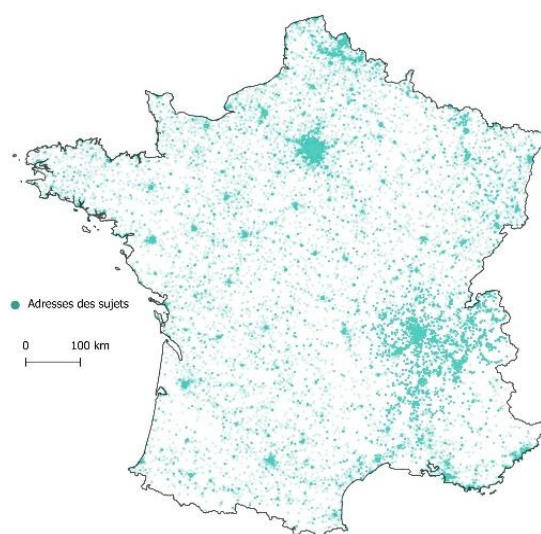
This work, carried out as part of a Master 2 internship and being pursued now as a PhD thesis, is part of the APoPCo project. This project is the result of collaboration between the Département Prévention Cancer Environnement of the Centre Léon Bérard and the Laboratoire de Mécaniques des Fluides et d'Acoustique of the Ecole Centrale de Lyon (CNRS UMR 5509). It is financed by ANSES (French Agency for the Environment, Health and Safety) and the Fondation de France.

The APoPCo project aims to estimate the association between breast cancer risk (10,000 subjects in the E3N French national cohort) and long-term exposure to air pollution (1990-2010, NO₂, PM₁₀, PM_{2.5}) by considering chronic exposure, at residential and occupational addresses and during commuting. The method for quantifying exposure during commuting is being developed to be applied at a national scale and to be reproducible in other epidemiological studies.

2. Materials and Methods

This project is based on data previously collected and produced by the team within the framework of the XENAIR project (2016-2021; funding by the Institut National du Cancer and the ARC Foundation):

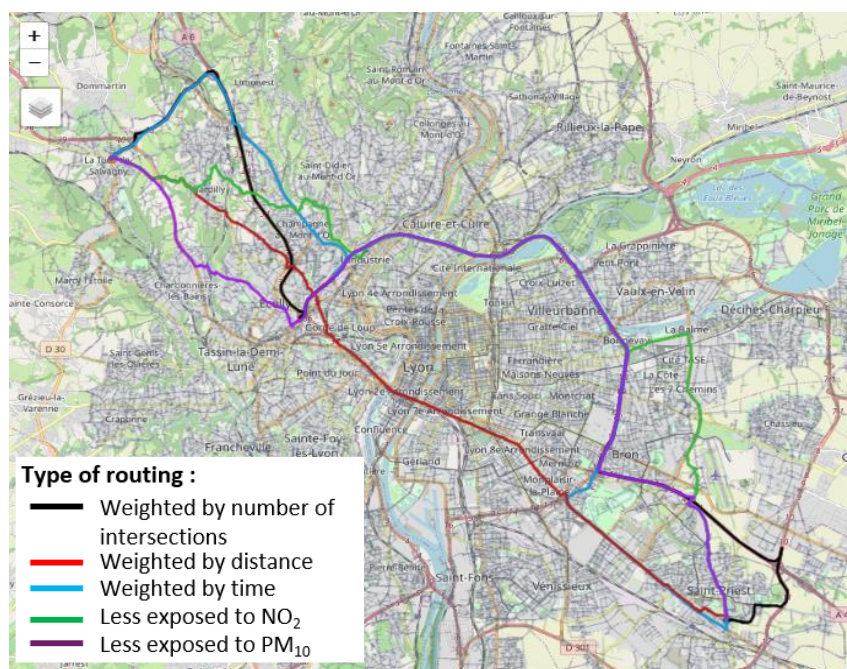
- A questionnaire on residential and professional mobility, and modes of transport used while commuting by the study subjects (10,444 women, i.e. 15,865 residential addresses between 1990-2010 (Map 1); professional addresses still geocoding) (Faure et al., 2017)
- NO₂ and PM (PM₁₀ ; PM_{2.5}) concentration data at local (SIRANE model, 10x10m, hourly) (Soulhac et al., 2017) and national (Land Used Regression (LUR) model, 50x50m, annual) scales over a 21-year period (1990-2010) (Amadou et al., 2020);



Map 1: Geocoding residential addresses from XENAIR project

Individual exposure to air pollution is calculated considering NO₂ and PM concentrations (estimated using the Land Use Regression Model (50 m resolution)), microenvironments of modes of transport, type of routing (Map 2), and traveling hours. Then, the inhaled dose is estimated based on the intensity of breathing volume. Finally, the variation in exposure induced by the choice of mode, as well as the proportion of the total daily exposure attributable to commuting, is evaluated. By identifying the

influence of each factor (mode, route, breathing volume, traveling hours etc.) on the classification of subjects, the optimal method can be defined to meet the needs of epidemiological studies.

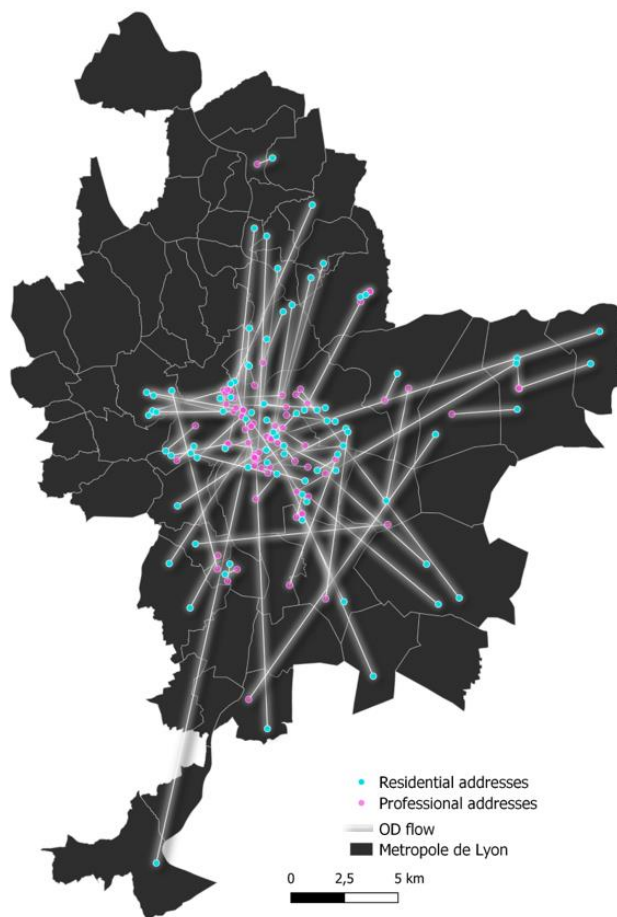


Map 2: Series of 5 routes calculated according to the Dijkstra algorithm

3. Results

3.1 Part of the daily inhaled dose attributable to commuting

First experiments carried out in the Metropole de Lyon (Map 3) – before application on the national scale – showed that the inhaled dose of pollutant while commuting contributes on average between 5 and 10% to the total daily dose. For each pollutant, commuting represents between 2 and 18% of the NO₂ dose, 1 and 16% of the PM₁₀ dose and 6 to 16% of the PM_{2.5} dose for a relatively small share of time (between 1 and 10% of the circadian time), compared to the shares of time and exposure estimated on average at home (57-66% of the circadian time, 52-62% of the daily dose) and at work (33% of the circadian time, 31 to 33% of the dose) (Fig 1).



Map 3 : Origin-Destination flow in the Metropole de Lyon

3.2 Hierarchy of mode exposure

Walking during commuting contributes the most to increase the inhaled dose (18% of the daily inhaled dose for NO₂, 16% for PM₁₀ and PM_{2.5}) (Fig 1), followed by cycling (NO₂: 5%, PM₁₀: 13%, PM_{2.5}: 6%) and finally driving a car (NO₂: 2%, PM₁₀:6%, PM_{2.5}:1%). This hierarchy is explained by the fact that active modes induce physical activity which increases the breathing volume (Bicycle : 23.5 L/min, Walk : 22.8 L/min, Car : 11.8 L/min (Zuurbier et al., 2009).

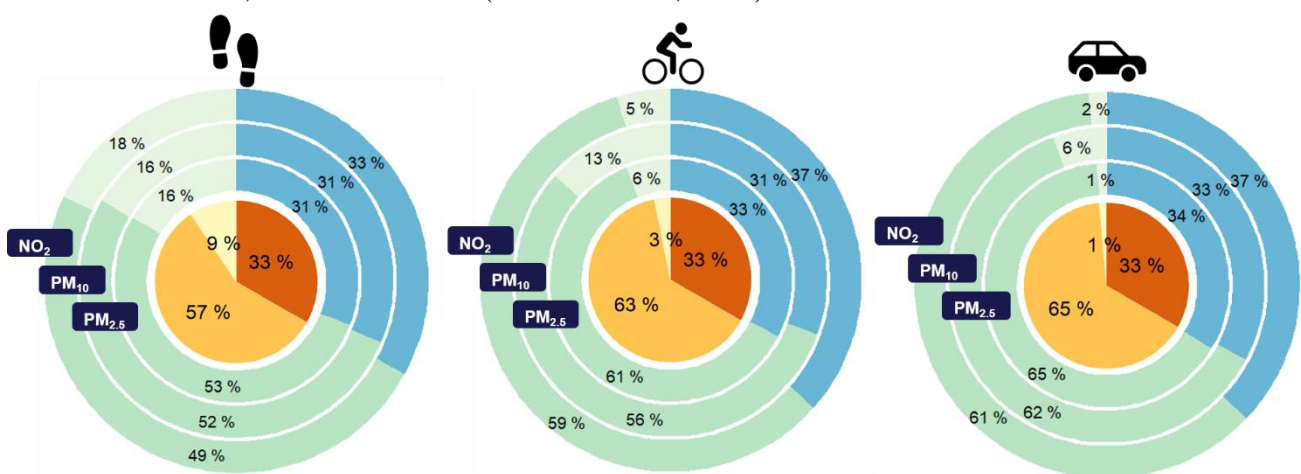


Fig 1: Part of daily exposure attributable to home, work and commutes and time distribution

4. Conclusions

Exposure to air pollutants during commutes varies according to the individual's characteristics (respiratory volume), of the individual's mobility (mode, route), as well as the characteristics of the urban environment (road network structure, speed limit, traffic). The method presented, which is still under development, provides a first estimation of exposure during travel and will be further developed. These developments will allow to classify modes according to the exposure of individuals, by integrating more parameters to define exposure in cars, as well as in public transport (metro, tramway, bus, etc.).

In addition, these intermediate results show that cycling or walking tend to increase exposure compared to car use. Nevertheless, previous studies have shown that practicing physical activity, even in a polluted environment, can help to reduce risk of overall mortality of up to 25% according to the study of Andersen et al. in 2015. The question of the benefits on breast cancer risk induced by active modes will therefore be explored, in addition to the negative impacts of pollution exposure.

Acknowledgments

The authors gratefully acknowledge the funding receive from the Ligue Contre le Cancer (Comité de Charente Maritime), the ANSES and the Fondation de France.

References

- Amadou, A., Praud, D., Coudon, T., Danjou, A.M.N., Faure, E., Leffondré, K., Le Romancer, M., Severi, G., Salizzoni, P., Mancini, F.R., Fervers, B., 2020. Chronic long-term exposure to cadmium air pollution and breast cancer risk in the French E3N cohort. *Int. J. Cancer* 146, 341–351. <https://doi.org/10.1002/ijc.32257>
- Andersen, Z.J., de Nazelle, A., Mendez, M.A., Garcia-Aymerich, J., Hertel, O., Tjønneland, A., Overvad, K., Raaschou-Nielsen, O., Nieuwenhuijsen, M.J., 2015. A study of the combined effects of physical activity and air pollution on mortality in elderly urban residents: the Danish Diet, Cancer, and Health Cohort. *Environ. Health Perspect.* 123, 557–563. <https://doi.org/10.1289/ehp.1408698>
- Basagaña, X., Aguilera, I., Rivera, M., Agis, D., Foraster, M., Marrugat, J., Elosua, R., Künzli, N., 2013. Measurement error in epidemiologic studies of air pollution based on land-use regression models. *Am. J. Epidemiol.* 178, 1342–1346. <https://doi.org/10.1093/aje/kwt127>
- Binder-Foucard, F., Bossard, N., Delafosse, P., Belot, A., Woronoff, A.-S., Remontet, L., French network of cancer registries (Francim), 2014. Cancer incidence and mortality in France over the 1980-2012 period: solid tumors. *Rev Epidemiol Sante Publique* 62, 95–108. <https://doi.org/10.1016/j.respe.2013.11.073>
- Faure, E., Danjou, A.M.N., Clavel-Chapelon, F., Boutron-Ruault, M.-C., Dossus, L., Fervers, B., 2017. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environ Health* 16, 15. <https://doi.org/10.1186/s12940-017-0217-5>
- Reis, S., Liška, T., Vieno, M., Carnell, E.J., Beck, R., Clemens, T., Dragosits, U., Tomlinson, S.J., Leaver, D., Heal, M.R., 2018. The influence of residential and workday population mobility

- on exposure to air pollution in the UK. *Environment International* 121, 803–813. <https://doi.org/10.1016/j.envint.2018.10.005>
- Rodgers, K.M., Udesky, J.O., Rudel, R.A., Brody, J.G., 2018. Environmental chemicals and breast cancer: An updated review of epidemiological literature informed by biological mechanisms. *Environmental Research* 160, 152–182. <https://doi.org/10.1016/j.envres.2017.08.045>
- Soulhac, L., Nguyen, C.V., Volta, P., Salizzoni, P., 2017. The model SIRANE for atmospheric urban pollutant dispersion. PART III: Validation against NO₂ yearly concentration measurements in a large urban agglomeration. *Atmospheric Environment* 167, 377–388. <https://doi.org/10.1016/j.atmosenv.2017.08.034>
- Steinle, S., Reis, S., Sabel, C.E., 2013. Quantifying human exposure to air pollution—Moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of The Total Environment* 443, 184–193. <https://doi.org/10.1016/j.scitotenv.2012.10.098>
- White, A.J., Bradshaw, P.T., Hamra, G.B., 2018. Air pollution and Breast Cancer: A Review. *Curr Epidemiol Rep* 5, 92–100. <https://doi.org/10.1007/s40471-018-0143-2>
- Zuurbier, M., Hoek, G., Hazel, P. van den, Brunekreef, B., 2009. Minute ventilation of cyclists, car and bus passengers: an experimental study. *Environ Health* 8, 48. <https://doi.org/10.1186/1476-069X-8-48>

Nonresponse in the analysis of the association between mental health and the urban environment in Brussels

Madeline GUYOT^{1,2}; Ingrid PELGRIMS³; Raf AERTS^{3,4,5}; Raf AERTS^{3,6}; Eva M. DE CLERCQ³; Isabelle THOMAS²; Sophie O. VANWAMBEKE¹

¹UCLouvain, Earth &Life Institute, Georges Lemaître Centre for Earth and Climate Research
Place Louis Pasteur 3 - L4.03.08, 1348 Louvain-la-Neuve (Belgium)
+32 (0)10 47 28 69, madeleine.guyot@uclouvain.be (correspondent author)

²UCLouvain, Louvain Institute of Data Analysis and Modeling in economics and statistics

³Sciensano (Belgian Institute of Health), Risk and health impact assessment

⁴Sciensano (Belgian Institute of Health), Epidemiology and Public health

⁵Ghent University, Applied Mathematics, Computer Science and Statistics

⁶KU Leuven, Division Ecology, Evolution and Biodiversity Conservation

Keywords: Mental Health, Non-response, Urban morphometrics, Brussels

1. Introduction

Understanding the relationship between urban environment and mental health is a challenging exercise (Pelgrims *et al.*, 2021). The issue of nonresponse and incomplete data, that often confronts population studies such as health surveys, further complicates the matter (Berete *et al.*, 2019; Boshuizen *et al.*, 2006; Burzykowski *et al.*, 1999; Volken, 2013). In the case of the Belgian Health Interview Survey (BHIS), three types of non-response can be highlighted: (i) initial nonresponse: refusal to participate in the survey (Van der Heyden *et al.*, 2014), not addressed in this paper, (ii) self-administered questionnaire (SAQ) nonresponse: for those over 15 years of age, the survey consists in two parts: a face-to-face (F2F) interview and a SAQ, the latter of which is not filled out, and (iii) item nonresponse: at different moments during the survey some questions were not answered by the participants. In this paper, the determinants of non-response are analysed, taking into account socio-economic as well as urban environmental factors; the impact of these non-responses on the analysis of the association between environment and mental health are discussed.

2. Background

For the BHIS, determinants of SAQ nonresponse have already been highlighted by Berete et al (2019): nonresponse is more frequent among youngsters, non-Belgians, lower educational levels and lower income, residents of Brussels and Wallonia, and people with poor perceived health. Non-response is also strongly associated with the interviewer. Other studies (Boshuizen *et al.*, 2006; Volken, 2013) on

partial non-response have additionally shown that non-response is greater for men and for unskilled workers. These studies also include urbanity indicators but the results do not converge.

3. Data and Method

Data used here are from the 2008 and 2013 BHIS, and are spatially limited to the Brussels-Capital Region, because the Region was oversampled compared to the rest of Belgium and because of data consistency issues for the urban environment indicators. We limit ourselves to participants over 15 years old (minimum age to fill the SAQ) and living at place of residence for at least one year ($n = 4\,355$). Logistic regressions with the nonresponse as dependent variable were computed: no missing data (0) vs. at least one nonresponse among six mental health indicators (1). As independent variables, socio-economic (reported household income, age, gender, family composition and highest educational level in the household) and urban environment (view of green, street canyon effect, noise, black carbon, street corridor effect, linear tree density, vegetation coverage and street visible vegetation coverage) indicators were selected (see Pelgrims *et al.*, 2021 for more information).

4. Results

In our sample of 4 355 individuals, the SAQ is not available for 36.5% of participants: for 19%, SAQ is required but not available and for 17.5%, SAQ is not required and not available (when the interview is done by a proxy such as a parent). For the available SAQ, non-response to mental health items ranged from 1% to 11%. If we analyse the subset of participants for whom the SAQ is available and who answered all mental health and socio-economic items, we have a set of 1 929 individuals (or 44% of the initial sample).

Partial nonresponse is higher among low income, older person, low educational level and people with children [Figure 1]. At the exception of family composition, these socio-economic determinants are the same as those associated with poor mental health (Silva *et al.*, 2016), so it is reasonable to assume that there will be more people with mental health problems in nonrespondents compared to the respondents.

Partial nonresponse is higher in low vegetation, more polluted (black carbon) and more urbanized areas (street canyon and street corridor effect) when adjusted for socio-economic variables [Figure 2].

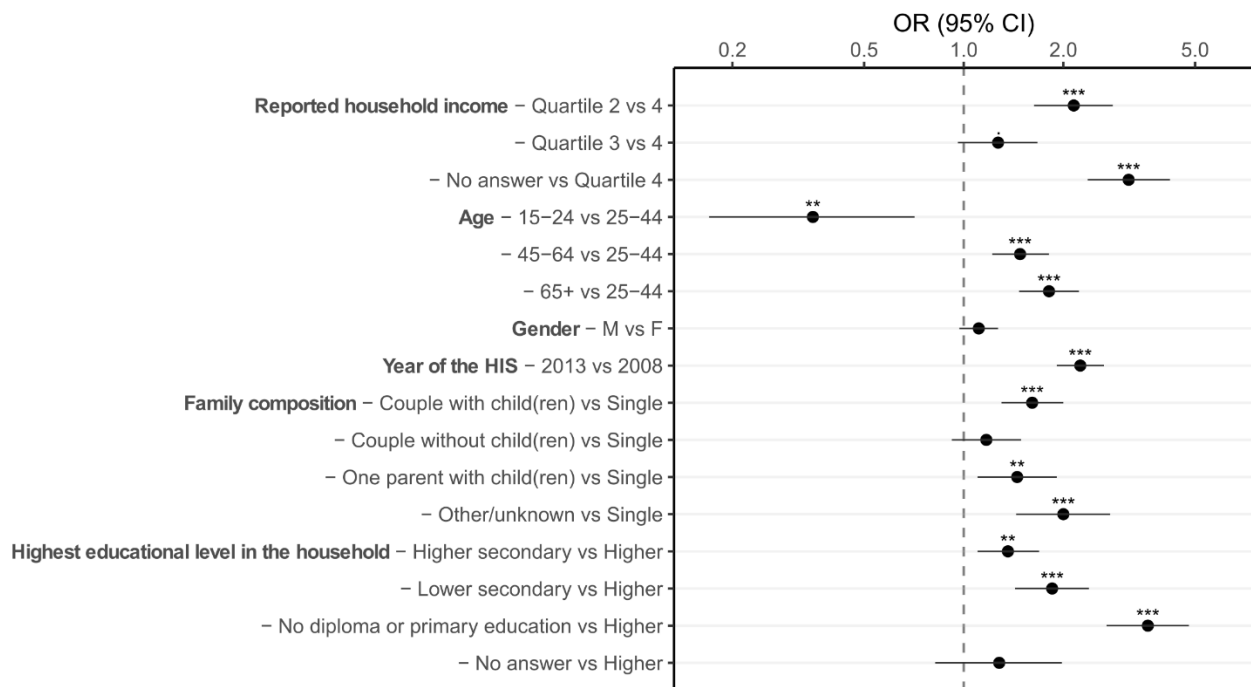


Figure 1. Odds ratios (and 95%CI) of nonresponse for socio-economic indicators. Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1.

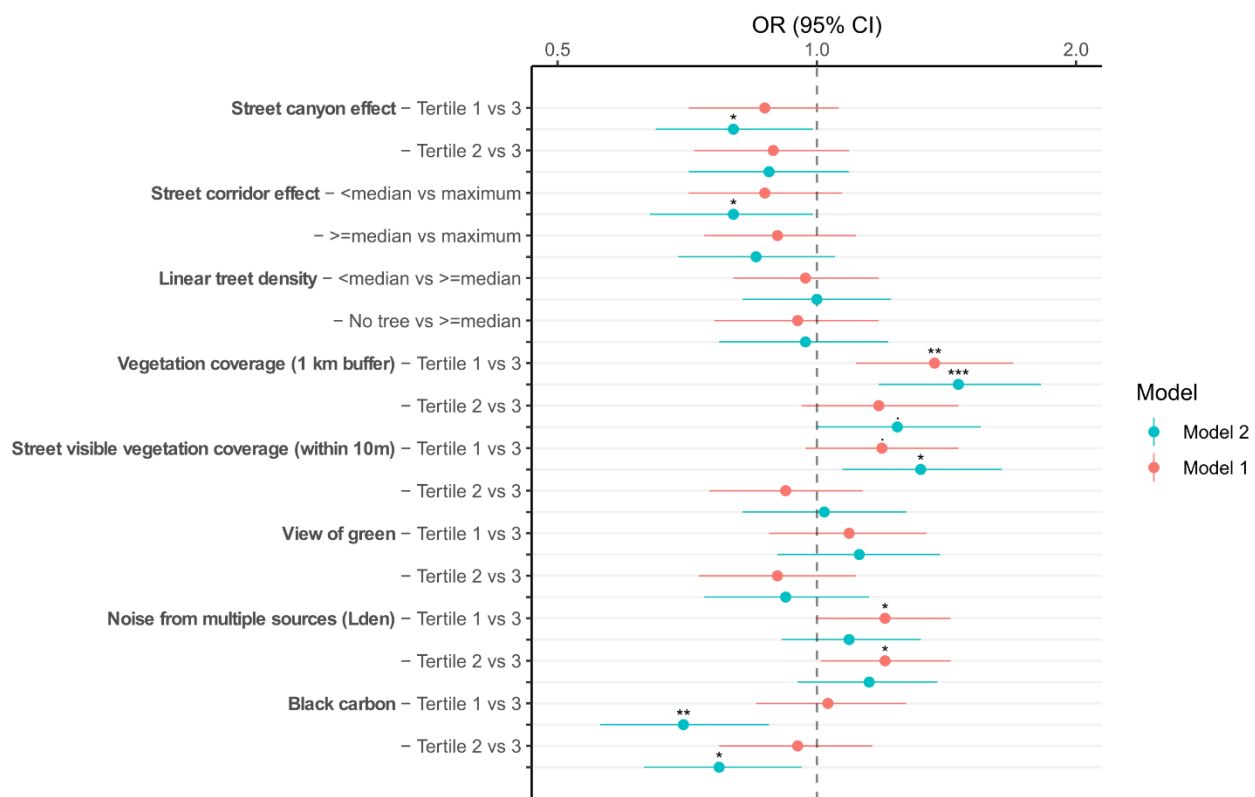


Figure 2. Odds ratios (and 95%CI) of nonresponse based on single-exposure models (Model 1). Model 2 is adjusted for sex, age, reported household income and year. Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1.

5. Discussion

Table 3. Example of a contingency table

	Mental disorder	No mental disorder
High vegetation coverage	A	B
Low vegetation coverage	C	D

With these results, we can extrapolate the effect of nonresponse on the association between mental health and the urban environment. Considering the association between vegetation and mental health, several studies on the topic have shown that vegetation is a protective factor for mental health (i.e. Gascon *et al.*, 2018; Lee and Maheswaran, 2011). However, in a previous study using the same data as this paper (Pelgrims *et al.*, 2021), no association could be found. With our results, it can be assumed that there would be proportionally a bigger pool of respondents in C [Table 3] than actually observed, i.e. $C_{true} > C_{observed}$. Indeed, the results show that the risk profiles for mental health have strong similarities with the risk profiles for nonresponse and significantly more nonresponses are found in places with lower vegetation cover. And considering the equation of an odd ratio:

$$OR = \frac{A * D}{B * C}$$

we may conclude that $OR_{true} < OR_{observed}$. Therefore, the protective effect of green spaces on mental disorder would be underestimated, i.e. the odd ratio observed is greater than the true odd ratio.

6. Conclusion

This paper deals with one of the major challenges encountered by research on the association between green space and health: the nonresponse in surveys. Because the spatial, as well as socio-economic, distribution of this bias is non-random, it is likely that it affects the research findings on the topic. We here show that the protective effect of green spaces on mental disorder may be underestimated.

Acknowledgement

This project was partially funded by the “Nature Impact on Mental Health Distribution (NAMED)” project (BR/175/A3/NAMED), a BRAIN-be (Belgian Research Action through Interdisciplinary Networks) project funded by BELSPO (Belgian Science Policy). This project aims to better understand the links between mental health and the urban residential environment (see Lauwers *et al.*, 2020, 2021; Pelgrims *et al.*, 2021).

References

- [1] Berete, F., Heyden, J. V. der, Demarest, S., Charafeddine, R., Gisle, L., Braekman, E., Tafforeau, J., and Molenberghs, G. (2019) Determinants of unit nonresponse in multi-mode data collection: A multilevel analysis. *PLOS ONE* 14(4): e0215652.
- [2] Boshuizen, H. C., Viet, A. L., Picavet, H. S. J., Botterweck, A., and van Loon, A. J. M. (2006) Non-response in a survey of cardiovascular risk factors in the Dutch population: Determinants and resulting biases. *Public Health* 120(4): 297–308.
- [3] Burzykowski, T., Molenberghs, G., Tafforeau, J., Van Oyen, H., Demarest, S., and Bellamammer, L. (1999) Missing data in the Health Interview Survey 1997 in Belgium. *Archives of Public Health* 57(3): 107–129.
- [4] Gascon, M., Sánchez-Benavides, G., Dadvand, P., Martínez, D., Gramunt, N., Gotsens, X., Cirach, M., Vert, C., Molinuevo, J. L., Crous-Bou, M., and Nieuwenhuijsen, M. (2018) Long-term exposure to residential green and blue spaces and anxiety and depression in adults: A cross-sectional study. *Environmental Research* 162: 231–239.
- [5] Lauwers, L., Leone, M., Guyot, M., Pelgrims, I., Remmen, R., Van den Broeck, K., Keune, H., and Bastiaens, H. (2021) Exploring how the urban neighborhood environment influences mental well-being using walking interviews. *Health & Place* 67: 102497.
- [6] Lauwers, L., Trabelsi, S., Pelgrims, I., Bastiaens, H., Clercq, E. D., Guilbert, A., Guyot, M., Leone, M., Nawrot, T., Nieuwenhuijse, A. V., Remmen, R., Saenen, N., Thomas, I., and Keune, H. (2020) Urban environment and mental health: the NAMED project, protocol for a mixed-method study. *BMJ Open* 10(2): e031963.
- [7] Lee, A. C. K., and Maheswaran, R. (2011) The health benefits of urban green spaces: a review of the evidence. *Journal of Public Health* 33(2): 212–222.
- [8] Pelgrims, I., Devleeschauwer, B., Guyot, M., Keune, H., Nawrot, T. S., Remmen, R., Saenen, N. D., Trabelsi, S., Thomas, I., Aerts, R., and De Clercq, E. M. (2021) Association between urban environment and mental health in Brussels, Belgium. *BMC Public Health* 21(1): 635.
- [9] Silva, M., Loureiro, A., and Cardoso, G. (2016) Social determinants of mental health: a review of the evidence. *The European Journal of Psychiatry* 30(4): 259–292.
- [10] Van der Heyden, J., Demarest, S., Van Herck, K., De Bacquer, D., Tafforeau, J., and Van Oyen, H. (2014) Association between variables used in the field substitution and post-stratification adjustment in the Belgian health interview survey and non-response. *International Journal of Public Health* 59(1): 197–206.
- [11] Volken, T. (2013) Second-stage non-response in the Swiss health survey: determinants and bias in outcomes. *BMC Public Health* 13(1): 167.

Preliminary spatial analysis of childhood asthma, air pollution, and socioeconomic status in Calgary, Canada

Stefania BERTAZZON¹; Caitlin CALDER-BELLAMY²; Rizwan SHAHID^{1,3}; Isabelle COULOIGNER¹; Yui TANAGI¹

¹Department of Geography, University of Calgary
2500 University Dr. NW, T2N 1N4, Canada

+1-403-220 7355, bertazzs@ucalgary.ca (corresponding author); icouloig@ucalgary.ca; yui.tanagi@ucalgary.ca

²Cumming School of Medicine, University of Calgary
2500 University Dr. NW, T2N 1N4, Canada

+1-403-220 7355, caitlin.calderbellam@ucalgary.ca

³Primary Health Services, Alberta Health Services,
10101 Southport Road SW, T2W 3N2, Canada

Keywords: Air Pollution, Asthma, Socioeconomic Status, Material Deprivation, Spatial Analysis

1. Introduction

Asthma is the most common chronic lung disease in Canadian children, with approximately 850,000 asthmatics under fourteen years old (Asthma Canada). Characterized by airway inflammation, wheezing, coughing, and difficulty breathing, asthma decreases quality of life, and is a primary cause of respiratory hospitalizations among Canadian adolescents (Asthma Canada; Bosonea et al., 2020).

Traffic-related air pollutants (TRAP) and environmental air pollutants are common causes of asthma exacerbation (AE) in adolescents. Children may be especially susceptible because their undeveloped lungs breathe in more polluted air per body weight (Bertazzon et al., 2020). Various traffic-related and environmental pollutants, namely nitrogen dioxide (NO₂), and particulate matter with both a diameter of 2.5 microns or less (PM_{2.5}) and a diameter of 10 microns or less (PM₁₀), are strongly associated with childhood AE. Other pollutants, namely black carbon (BC) and BTEX (a combination of 4 volatile organic compounds: benzene, toluene, ethylbenzene, xylene), have not been extensively studied, yet research indicates they can cause childhood AE. Often, families of lower socioeconomic status (SES)

live in areas of elevated air pollution (AP), leading to increased respiratory-related hospitalizations (Cakmak et al., 2006).

Asthma prevalence has been increasing in Albertan children aged 1 to 14 (Bosonea et al., 2020). In Calgary, Alberta, previous research has found PM_{2.5}, PM₁₀ and BTEX, to be consistently correlated with, and identified potential growing disparities across communities (Bertazzon et al., 2020).

The examination of the relationship between AP and AE in children is critical considering the COVID-19 pandemic; individuals infected with COVID-19 and diagnosed with asthma have an increased risk of mortality (Lee et al., 2020). We therefore performed a spatial analysis to further understand the associations between asthma, air pollution, and socioeconomic status in Calgary, Canada.

2. Data and Methods

2.1 Data

Asthma data, acquired by Alberta Health Services, includes asthma emergency room visits (AERV) for children aged 0-14 at the Sub-Local Geographic Area (SLGA) level, ie., communities with a population of at least 1000. Asthma ratios per SLGA were calculated by dividing total AERV per SLGA per year by the total number of children aged 0-14 per SLGA. AERV visits were used as proxy for AE. 2016 Census data was collected through the University of Calgary and was used throughout 2014-2018.

AP data was acquired from LUR estimates for summer and winter months, monitoring NO₂, PM_{2.5}, PM₁₀, BC and BTEX. AP levels and spatial distribution were consistent over a 5-year period (Bertazzon et al., 2015; Bertazzon et al., 2019; Bertazzon et al., 2021); therefore 2015/2016 estimates were used throughout this study (Figure 1).

The Pampalon index (Alberta Health, 2020) expresses socioeconomic inequalities in Canada. It is composed of a social deprivation index and a material deprivation index (MDI). The MDI is more relevant in Alberta's context, and combines three variables: education, employment, and income, which are sorted into quintiles (Alberta Health 2020). The MDI at the SLGA level represents SES within Calgary.

2.2 Methods

Descriptive statistics and maps were used to illustrate three variables: asthma ratios, AP, and the MDI. Correlations and spatial statistics (Moran’s I, Getis G, and hotspot analysis) were computed for all three variables at the city level, the quadrant level (i.e., NW, NE, SE, and SW), and the local geographic area (LGA) level. Spatial regression and further predictive models are currently underway to further understand spatial-temporal trends. ArcGIS Pro was used for all maps presented, and R Studio was used for statistical analysis.

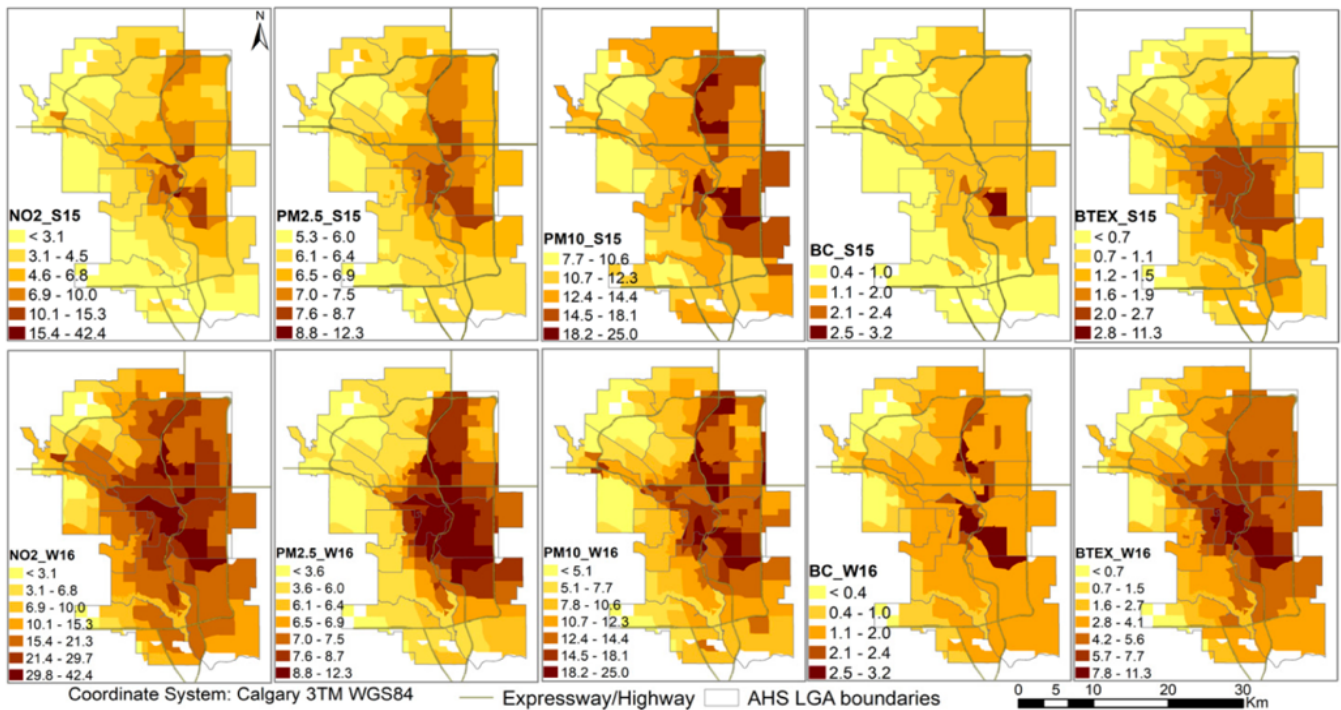


Figure 1: Maps depicting AP in the summer (S15) and winter months (S16)

3. Results and Discussion

Correlations were run between all 3 variables: AP, MDI, and asthma ratios.

3.1 SES and Asthma Ratios

There are positive, significant correlations at the city level, especially during the summer months, suggesting a potential spatial segregation of AE. However, over time, the correlations’ value decreases, possibly identifying decreased health inequalities.

At the quadrant level, most significant correlations were in the SW quadrant, yet correlations decreased over time, further indicating diminished health inequalities. The NE and SE quadrants have a higher MDI; the SW quadrant has a lower MDI. There may be fewer significant correlations in NE and SE quadrants because the MDI is consistent throughout. These results suggest indoor AP may have a greater impact on AE than outdoor AP. Areas of lower SES are commonly associated with greater indoor AP, resulting from dust, mold spores and poor ventilation.

3.2 AP and SES

PM10 was significantly correlated to the MDI during summer months at the city level. In the NE quadrant, all pollutants were negatively correlated to the MDI during summer and winter months. Despite the NE quadrant being a semi-industrial area, the MDI is negatively associated with AP, possibly because SLGAs have similar MDI values and consistent AP. However, the SE quadrant, another semi-industrial area, showed consistent positive correlations between the MDI and AP during summer and winter months, possibly because there is more variation in MDI throughout the quadrant, or outdoor AP is more irregular.

3.3 Asthma Ratios and AP

At the city level, there were slight significant negative correlations in 2015 during summer and winter months. Interestingly, there were significant negative correlations between all pollutants in the SW quadrant. PM2.5, a common TRAP, was most consistently negatively correlated, indicating outdoor AP may not be the primary cause of AE, but instead other factors, including indoor AP.

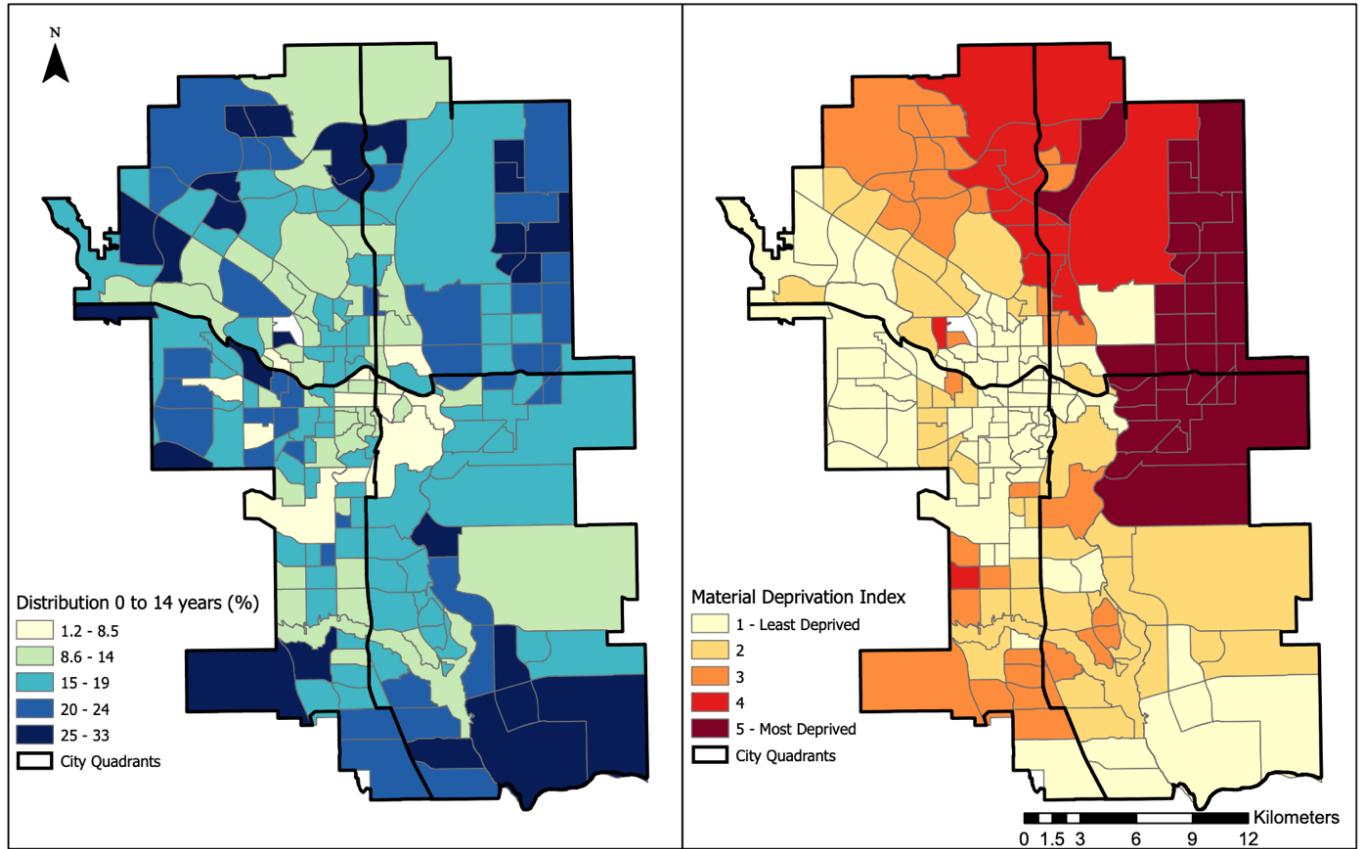


Figure 2: (a) Distribution of children aged 0-14 in Calgary per SLGA (b) Distribution of the MDI in Calgary per SLGA

3.4 Spatial Statistics

Spatial statistics were run for all three variables. Moran's I and Getis G analysis show significant clustering of MDI and AP. Mapping further shows evidence of MDI inequity in Calgary, where the NE quadrant experiences greater deprivation (Figure 2). AP shows a spatial-temporal trend, especially in winter months, where pollutants show a gradient-like increase with increasing concentrations in the NE (Figure 1). There was no significant asthma ratio clustering. However, mapping suggests a spatial-temporal trend where asthma ratios increase towards the NE quadrant in the summer months over time (Figure 3).

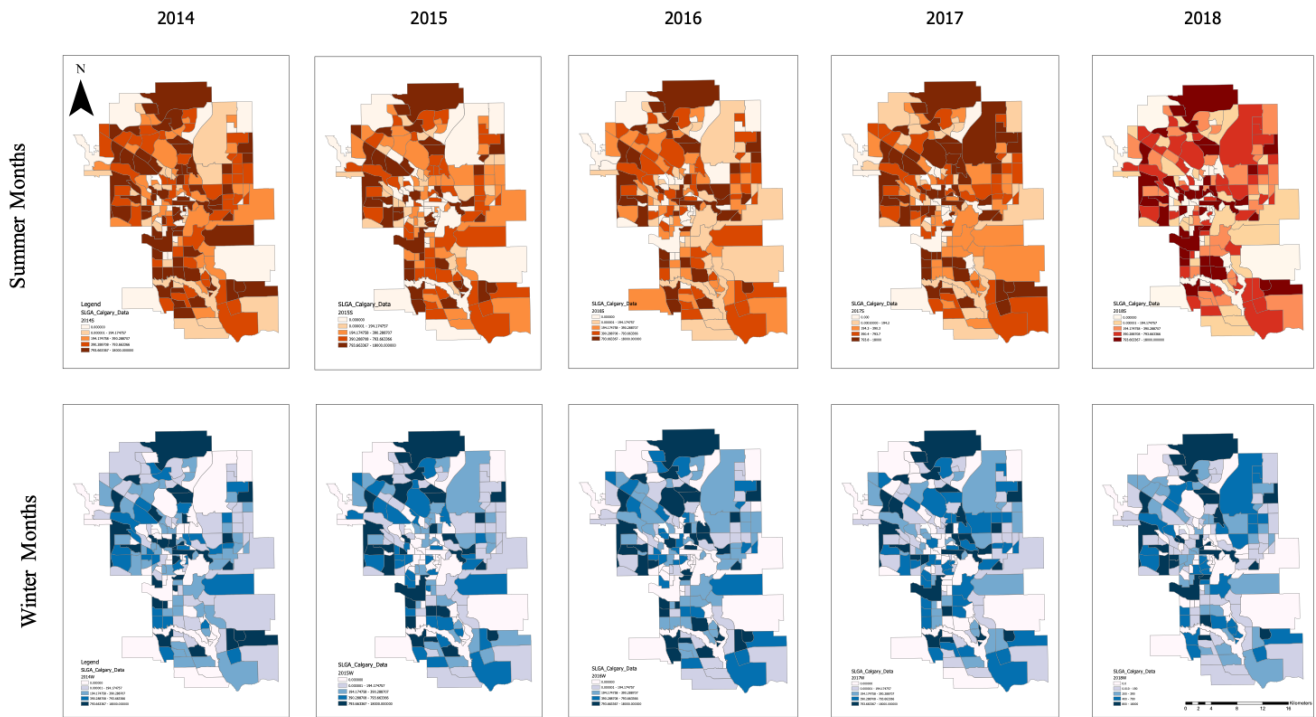


Figure 3: Spatial-temporal trend of asthma ratios between 2014 and 2018 during summer and winter months

4. Conclusion

The application of spatial analytical tools at the fine geographic level highlight critical and applicable health related findings. Correlations suggest some health inequity in Calgary, yet the primary cause of childhood AE may not be outdoor AP. Although health inequity may decrease over time, mapping suggests AE is becoming more prominent in areas of lower SES during summer months. Spatial statistics indicate clustering of outdoor AP and MDI in Calgary, yet AE was not, therefore inferring other factors may influence AE, like indoor AP, allergies, or cold temperatures.

There are some limitations to this study. AP estimates from 2015/16 were used rather than long-term dynamic AP data. Moreover, we only monitored 5 years, a rather short study period. Further lines of investigation include a longer study period to better understand health long-term inequity patterns, and further predictive models and spatial-temporal AP monitoring in areas of high AP exposure, including parks, schools, and areas of active transport.

Acknowledgments

We acknowledge the following funding sources: Alberta Innovates, the CIHR Institute for Population and Public Health, and the O'Brien Institute for Public Health.

References

- Alberta Health Services (2020). How to use the Pampalon Deprivation Index in Alberta. Calgary: Alberta Health Services. Available at: <https://www.arcgis.com/home/item.html?id=8df129ba3a5948d5a07f4e6c6215cd85> (Accessed: 10 June 2021).
- Asthma Canada. (no date). *Asthma Facts and Statistics*. Available at: <https://asthma.ca/wp-content/uploads/2020/07/Asthma-101.pdf> (Accessed: 25 May 2021).
- Bertazzon, S., Johnson, M., Eccles, K. and Kaplan, G. (2015). 'Accounting for spatial effects in land use regression for urban air pollution modelling', *Spat Spatiotemporal Epidemiol*, 14(15), pp.9-21.
- Bertazzon, S., Couloigner, I. and Underwood, F. (2019). 'Spatial land use regression of nitrogen dioxide over a 5-year interval in Calgary, Canada', *Int J Geogr Inf Sci*, 33(7), pp.1335-54.
- Bertazzon, S., et al. (2020). 'A Preliminary Spatial Analysis of the Association of Asthma and Traffic-Related Air Pollution in the Metropolitan Area of Calgary, Canada', *Atmosphere*, 11(10), pp.1066.
- Bertazzon, S., Couloigner, I. and Mirzaei, M. (2021). 'Spatial Regression Modelling of Particulate Pollution in Calgary, Canada', *Geojournal* [Preprint]. Available at: 10.1007/s10708-020-10345-7 (Accessed: 29 June 2021).
- Bosonea, A.M., et al. (2020). 'Developments in asthma incidence and prevalence in Alberta between 1995 and 2015', *Allergy Asthma Clin Immunol*, 16(87), pp.1-11.
- Cakmak, S., Dales, R.E. and Judek, S. (2006). 'Respiratory health effects of air pollution gases: modification by education and income', *Arch Environ Occup Health*, 61(1), pp.5-10.
- Lee, S.C., et al. (2020). 'Impact of comorbid asthma on severity of coronavirus disease (COVID-19)', *Sci Rep*, 10(1), pp.21805.

Which local historical information as a proxy for environmental exposure to air pollution in France throughout the 20^e century?

Lény GRASSOT^{1,2}; Yohan FAYET³; Delphine PRAUD^{1,2}; Béatrice FERVERS^{1,3}; Thomas COUDON^{1,2}

¹Department Prevention Cancer Environment, Centre Léon Bérard

28 rue Laennec, 69008, Lyon, France

+33 4 69 85 60 48, leny.grassot@lyon.unicancer.fr (correspondent author)

²INSERM U1296 Radiation: Defense, Health, Environment, Centre Léon Bérard

28 rue Laennec, 69008, Lyon, France

³Equipe EMS – Département de Sciences Humaines et Sociales, Centre Léon Bérard

28 rue Laennec, 69008, Lyon, France

Keywords: Environmental exposure, Air pollution, Historical assessment, Demographics, Epidemiology

1. Introduction

Analytical environmental epidemiology seeks to investigate risk factors for pathologies associated to the living environment of individuals. It is therefore important to integrate temporal and spatial dimensions of the methodological approach including individual and environmental information. While risk factors' exposure can only be prior to the pathology, this discipline requires a historical perspective but also a knowledge of the territory. Moreover, certain pathologies, and particularly cancers such as breast cancer, the most frequent cancer diagnosed in women in industrialized countries need a significant historical perspective (up to several decades) for the study of their etiologies (Wild, 2012). Reconstitute individual information is generally based on subjects' declaration— resulting in recall biases. Moreover environmental information is limited because of missing historical data. Environmental data currently available do not allow a direct and detailed analysis of the historical exposures of each subject in our studies. However, this methodology is essential to better considered risk factors for past, present and future cancers.

The RHQ study – *Residential History Questionnaire* – is an epidemiological project on breast cancer risk. It attempts to explore some possibilities of exposure assessment over a long period of time using to the scarce data available. Based on the E3N cohort (100,000 women, born between 1925 and 1950, insured by a national health insurance plan and followed-up since 1990 by auto-questionnaires), the study is a nested case-control study including data on 2,876 incident breast

cancer cases identified up to 2011 and 3,616 matching controls. The project aimed to study the association between life-long residential trajectory of women and breast cancer risk and therefore to take into account the life trajectories across places of residence and their precise chronology.

Previous studies conducted by the project team showed that living in an urban area at inclusion of the cohort is associated with a higher breast cancer risk as well as being born in an urban area (Binachon et al., 2014).

In the continuity of these studies, the rural/urban status have been analysed as a surrogate of particular lifestyles, environmental and spatial exposures but other indicators could be taken into account to understand the best historical opportunities.

The objective is therefore to use reliable, homogeneous and regular information to understand the evolutions and trends of the atmospheric environments since 1900; allowing a spatial and temporal analysis but also possible and coherent geographical and historical comparisons to find rational and comprehensible proxies to consider environmental exposures

2. Methods

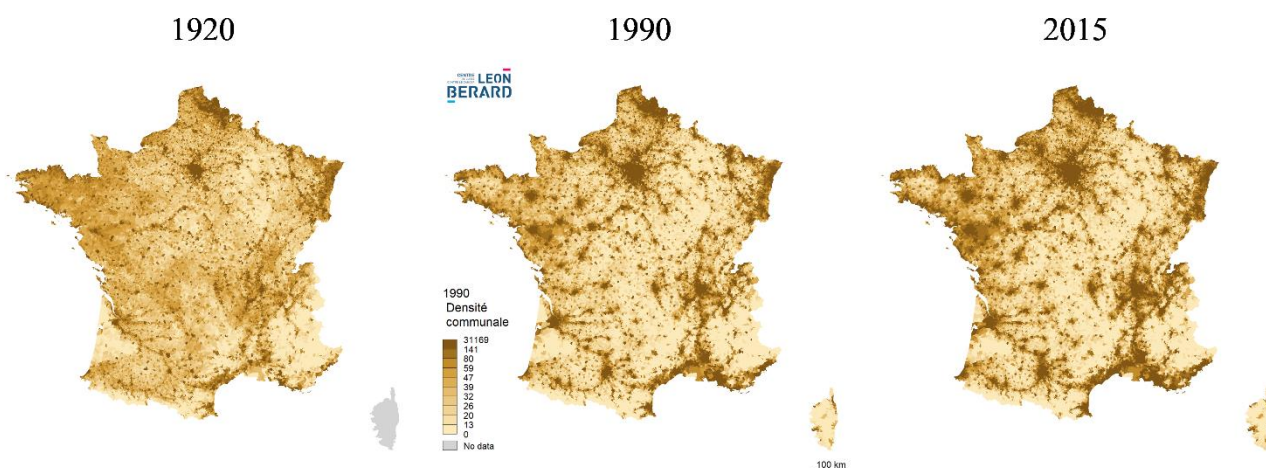
2.1 Demographics

INSEE (French National Institute of Statistics and Economic Studies) provided municipal demographic data since 1876 for the mainland France by considering as a stable territorial division in time the municipalities on January 1th, 2019. From the regular censuses [average gap of 5 years], a linear extrapolation allows us to reconstitute the annual demographics of each territory.

A supra-municipal indicator (named “local”) was generated to limit the impact of the administrative division: taking into account the demographic and surface information of neighbouring municipalities (maximum 80 kilometres between centroids) with a decrease factor related to the distance – in kilometres – squared.

From municipal and local population count, secondary data are produced:

- Boolean urban/rural status of the municipalities – based on a baseline commune/local population of 2,000 inhabitants
- Density – number of inhabitants per square kilometre
- Decadal and bi-decadal evolution of the municipality – percentage difference between the year analysed t_0 and $t_{-10\text{years}}$ OR $t_{-20\text{years}}$.
- Municipal rank of demography / density, commune / location



Map 1: Example of the evolution of municipal indicators – Density for the years 1920, 1990 and 2015

2.2 Environmental data

For an extended analysis, and a maximization of the spatial data, we relied on 3 distinct resolution grids based on the same model and a panel of 4 pollutants: Nitrogen dioxide (NO₂), Ozone (O₃), and Particulates Matters (PM_{2.5} and PM₁₀).

Air pollution data were produced, by the CHIMERE model (Couvidat et al., 2018; Terrenoire et al., 2015), chemistry-transport model developed by INERIS (French National Institute for the Industrial Environment and Risks). The 4 pollutants concentrations were simulated at various spatial resolutions over 28 years (1990-2019) on a regular grid:

- 0.0625° x 0.125° (about 8 x 8 km) for 1990-2010
- 0.03125° x 0.0625° (about 4 x 4 km) for 2000-2017
- 0.0078125° x 0.015625° (about 1 x 1 km) for 2009-2013

Municipal exposures were based on spatial overlays between the municipal area and pollutant grids.

2.3 Land use occupation

An assessment of the municipal land use was evaluated using data from the HILDA project (Fuchs et al., 2013, 2015). It is a ten-years grid (1 x 1 km) extending over 28 European countries (EU-27 plus Switzerland) and hierarchically divided into 6 thematic classes (Settlements, Cropland, Forest, Grassland, Other Land and Water).

Land use occupation are based on spatial overlays between the municipal area and land use grids.

2.4 Geographical data

Finally, geographical invariants were integrated to account for the spatial situation of the municipality within the metropolitan territory: Latitude, Elevation, and Distances to neighbouring

countries, to the ocean and seas. In addition to the above data, geographical invariants would allow the national situation to be taken into account to highlight possible natural, non-anthropogenic or foreign exposures.

2.5 Data mining

The experimentation was based on the use of classical and usual data mining methods to detect a proximity link between two data sets: e.g. correlation (linear relationship) and misclassification (weighted Kappa); and to select relevant multi-variable models proven on the usual indicators (adjusted R², BIC and CP/AIC).

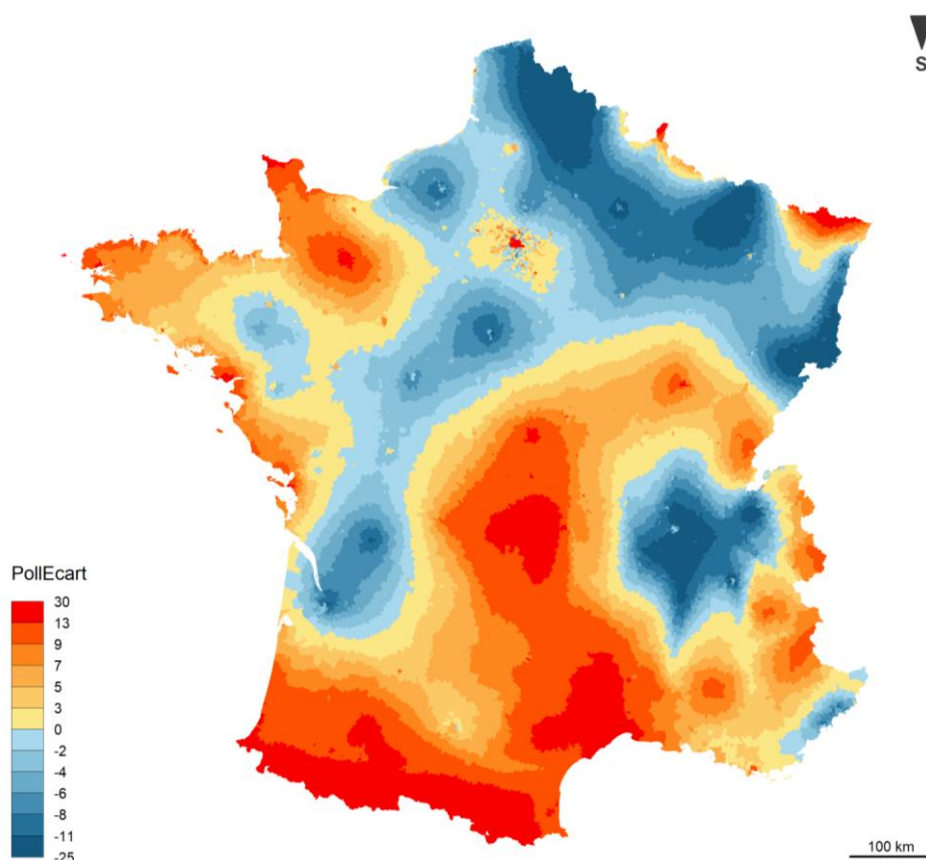
3. Results

Initial results showed variability in demographic and geographic information with each of the 4 pollutants but a strong historical stability of all the relationships. The first figure (Figure 1) illustrates the annual variations in correlations between the different pollutants and spatial scales and a selection of variables: trends between scales and across the 3 decades studied remain fairly constant.



Figure 1: Correlations (Pearson, on values) of selected variables and sources/pollutants

Systematic exploration of patterns from each demographic and geographic variable and pollutant accounts for spatially heterogeneous variabilities (Map 2). The various models and maps show heterogeneous variations according to the French territories: the changes remain progressive and illustrate the multifactorial tenants of the concentrations through space.



Map 2: Municipal density and PM_{2.5} (1x1km resolution): Variation (percentage of over- or under-estimation) between the regression model and the raw variable for the year 2010

Calculations and analyses are underway. More detailed results can be presented on the univariate relationships and the multi-variate models generated especially from the point of view of spatial analysis (e.g. global or local spatial autocorrelation).

2. Conclusion

The rural/urban status remains an opportunity, but more precise indicators, both demographic and geographic, seem to show more relevant results, with notably less inter-pollutant variability. For example, the density but also the demographic rank of the municipality or certain geographical invariants such as the latitude, are statistically more precise explanatory factors on the various pollutants.

Nevertheless, the creation of a single model for the different pollutants is complex, as the spatial and temporal characteristics and evolutions of each one remains heterogeneous. It will therefore be possible to consider a composite model that responds to historical and geographical constraints.

Acknowledgments

This work was supported by the National Institute for Cancer (INCa) (project SHS-ESP17-070).

References

- Binachon, B., et al. (2014). ‘Life in urban areas and breast cancer risk in the French E3N cohort’, *Eur. J. Epidemiol.*, 29:743–751; doi:10.1007/s10654-014-9942-z.
- Couvidat, F., et al. (2018), ‘Development of an inorganic and organic aerosol model (CHIMERE 2017 β v1.0): seasonal and spatial evaluation over Europe’, *Geosci. Model Dev.*, 11, :165–194; doi: 10.5194/gmd-11-165-2018.
- Fuchs, R., et al. (2015), ‘Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010’, *Global Change Biology* 21, 1. – ISSN 1354-1013 – pp. 299-313.
- Fuchs, R., Herold, M., Verburg, P.H., Clevers, J.G.P.W. (2013), ‘A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe’, *Biogeosciences* 10, ISSN 1726-4170 – pp. 1543-1559.
- Terrenoire, E., et al. (2015), ‘High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE’, *Geosci..Model Dev.*, 8:21–42; doi: 10.5194/gmd-8-21-2015.
- Wild, C.P., (2012), ‘The exposome: from concept to utility’, *Int. J. Epidemiol.*, 41, 24–32; Doi:10.1093/ije/dyr236.

Estimating health and social outcomes at small areas in Great Britain by generating and utilising synthetic population data

Guoqiang WU¹; Alison HEPPESTALL²; Nik LOMAX³

¹Leeds Institute for Data Analytics and School of Geography
University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK
+44 (0) 113 343 9680, g.wu@leeds.ac.uk (correspondent author)

² School of Social and Political Science and
MRC/CSO Social and Public Health Sciences Unit
University of Glasgow, Glasgow, G3 7HR, UK
alison.heppenstall@glasgow.ac.uk

³ Leeds Institute for Data Analytics and School of Geography
University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK
+44 (0) 113 343 3321, N.M.Lomax@leeds.ac.uk

Keywords: microsimulation, synthetic microdata, population estimation, health modelling, health inequalities

1. Background

In order to understand the health outcomes for distinct sub-groups of the population or across different geographies, it is advantageous to be able to build bespoke groupings from individual level data. Individuals possess distinct characteristics and behaviours and accumulate their own unique history of exposure or experiences. However, in most disciplines, not least public health, there is a lack of individual level data available outside of secure settings, especially covering large portions of the population. In this abstract, we present the rationale for, and methods used to construct a synthetic population dataset used by the SIPHER (*Systems Science in Public Health and Health Economics Research*) consortium, a collaboration of researchers from seven universities, three government partners and 12 practice partners in the UK. Data outlined in this abstract will be used as an input to models geared towards assessing outcomes within a system map of an inclusive economy and the impacts of policy interventions on a range of health and social outcomes. Here, we present the methods used to generate and validate the synthetic population for several geographical areas: the city region

of Greater Manchester (comprising 10 local authority districts), Sheffield local authority district, Glasgow council area and Cardiff local authority district.

2. Methods

2.1. Spatial microsimulation

This research applies spatial microsimulation to generate a synthetic population covering the whole of Great Britain (GB). Spatial microsimulation combines spatial constraints with more attribute rich individual level data to synthesise individuals within defined geographical zones (Lovelace & Dumont, 2016). This research utilises static spatial microsimulation to produce synthetic population of individuals at Lower Super Output Area (LSOA) scale (administrative areas of 1500 people) for England and Wales and Data Zones (500 - 1000 individuals) for Scotland, for the year 2018 (our base year). For microsimulation models, input data normally consists of a non-geographical but otherwise attribute-rich individual level dataset, for example a representative survey, and constraint tables containing aggregate counts across a number of attributes for a series of geographical zones (e.g. LSOAs), for example census data. To calculate the weights allocated to the individuals for each geographical zone, linking variables, which are shared between the individual and aggregated level datasets, are required for establishing a spatial microsimulation model (Lovelace & Dumont, 2016). In this study, two spatial microsimulation (sub-)models are developed and run separately for adults (aged 16 and over) and children (aged 15 and under). After each model is run, their population outputs are merged together to generate the full synthetic population and estimates of health outcomes at the LSOA level are created. The overall framework of generating synthetic population and health estimates in this study is shown in Figure 1.

2.2. Selecting and formatting linking variables in the microdata

The microdata (non-geographical individual level data) used for this study are derived from Wave 9 of a nationally representative longitudinal household survey – Understanding Society (the UK Household Longitudinal Study), conducted in 2018. Understanding Society collects a range of health and socio-economic information about adult individuals and a number of these align with the census data used as the spatial constraints in the spatial microsimulation. In our adult microsimulation model, seven linking variables are developed from the Understanding Society's adult dataset, including sex by age (sex/age), economic status, highest educational qualification, marital status, ethnicity, household composition, and housing tenure. The child microdata is formed by using the Understanding Society's child dataset which records information about the adult respondents' children (aged 15 and under). Compared to the adult microsimulation model which uses seven variables mapping to the

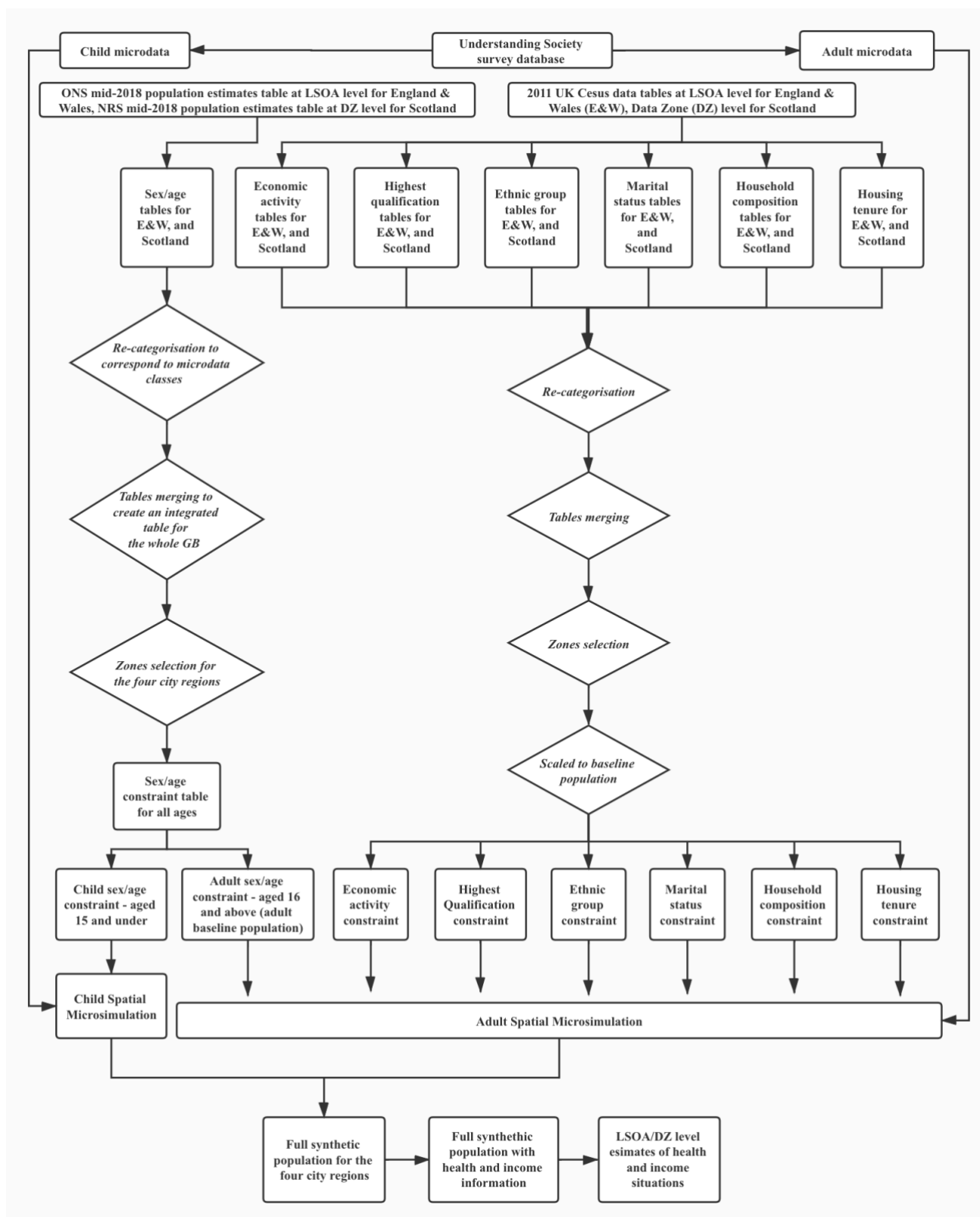


Figure 1: Methodological framework of generating synthetic population

Census data, the child microsimulation model is simpler because the linking variables shared between the Understanding Society survey data and geographically aggregated data are quite limited. Hence, only sex/age variable are included in child microsimulation model.

2.3. Formatting geographical constraint variables

Tables which report the total count of individuals for each of the seven linking variables for each geographical zone are used as constraints in the microsimulation process. The sex/age constraints come from the Office for National Statistics (ONS) mid-2018 population estimates for England and Wales, and the National Records of Scotland (NRS) mid-2018 population estimates for Scotland. These are formatted to match the sex/age categories in the microdata. This dataset is then split into an adult constraint dataset and a child constraint dataset. The child microsimulation model is then run using only sex/age constraint as discussed earlier. For the adult model, further constraints are derived from 2011 UK Census data. Because the total number of people in 2011 does not match the total in 2018, the Census tables are scaled to match the sex/age totals reported in the mid-2018 data for each geographical zone.

2.4. Software and algorithm

This study applies a Java-based application, the Flexible Modelling Framework (FMF) developed by the University of Leeds, to undertake microsimulation (Harland, 2013). The FMF incorporates a static spatial microsimulation algorithm based on simulated annealing which is a combinatorial optimisation algorithm proceeding by selecting an optimal configuration from a small sample population (e.g. survey data) constrained by observed aggregate population counts (population census) (Harland et al., 2012; Tanton, 2014). In addition, the FMF includes a model evaluation function that enables internal model validation by calculating goodness-of-fit statistics (e.g. R^2 , Total Absolute Error, Standard Absolute Error, Standardised Root Mean Square Error) at individual cell, category and overall attribute levels.

2.5. Utility of the synthetic data outputs

Once the microsimulation process is complete, the synthetic population data, containing individual personal identifiers and the codes of LSOA zones that each individual is allocated to, are generated. This means that the variables (e.g. income, physical and mental health performance, subjective wellbeing) that are not otherwise available at a high resolution geography can be made available to researchers by linking the synthetic population dataset with the original individual-level microdata (Understanding Society survey data) through the personal identifiers. The spatial distributions of these variables can be then mapped by joining the geographically aggregated dataset with the 2011 Census geography boundaries. Figure 2 gives an example of the spatial resolution of the data that is deposited across the four GB city regions.

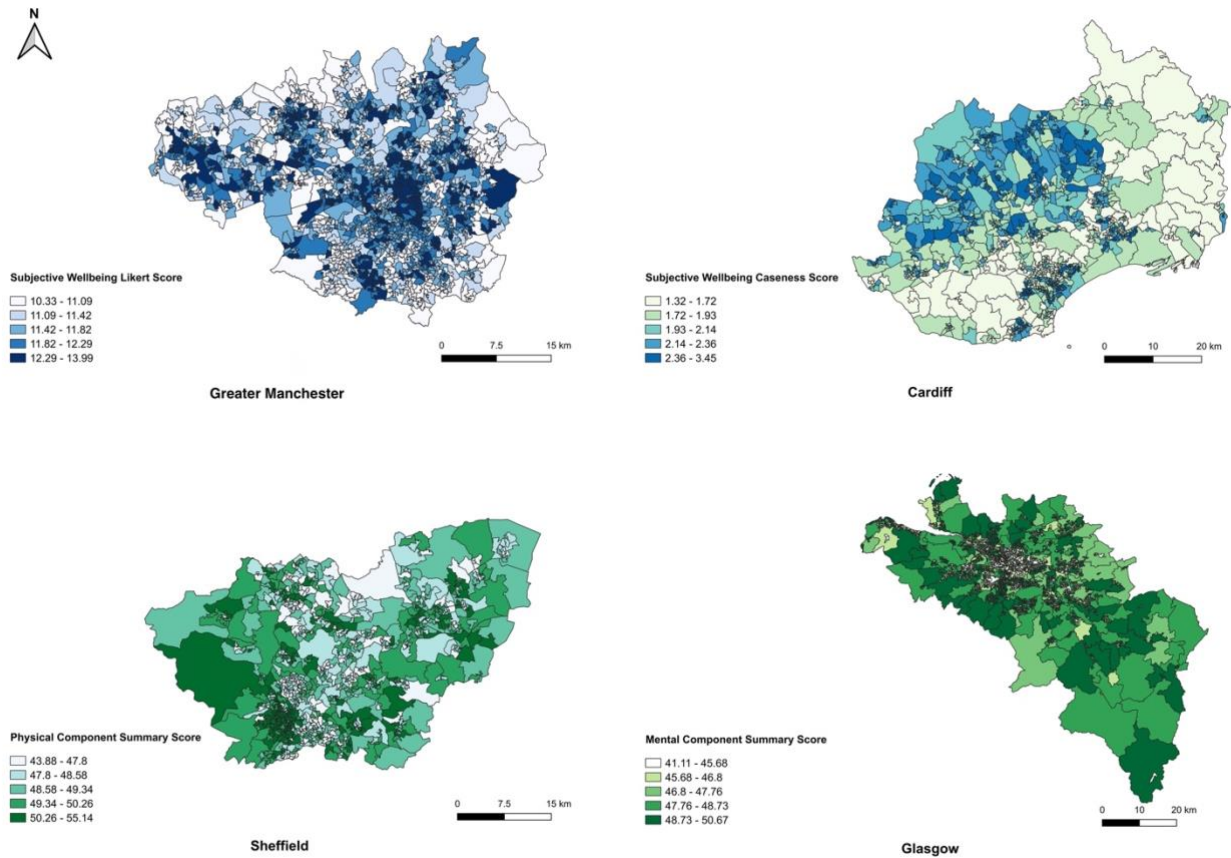


Figure 2: Examples of aggregated health conditions estimates at LSOA and equivalent level in the four selected city regions

3. Technical validations

Validation is a vital step that provides a level of confidence in the final synthetic data and normally performed internally and externally.

Internal validation refers to comparing values from the simulated dataset to the original datasets used in the simulation (Lovelace & Dumont, 2016). In practice, this process includes model calibration, whereby the model fit is assessed by comparing the observed and simulated values for constraint variables. In our case, the goodness-of-fit statistics at overall attribute level provided by the model evaluation function of the FMF show a good fit between the simulated and original data.

External validation is the process of comparing the simulated results to a different source of data that is external to the model. The first external datasets used to validate our simulated results are from the Index of Multiple Deprivation 2019 (IMD 2019), the Welsh Index of Multiple Deprivation 2019 (WIMD 2019), and the Scottish Index of Multiple Deprivation 2020 (SIMD 2020) which are the official measurements of relative deprivation for small areas (LSOAs or Data Zones) in England,

Wales, and Scotland. In the IMD/WIMD/SIMD datasets, all small areas are ranked according to their level of deprivation relative to that of other areas, and a higher rank suggests a more deprived situation. We use Spearman's test of rank correlation to examine the relationships between the IMD, WIMD, and SIMD ranks and our simulated income and health estimates at LSOA level for different city regions. These results suggest that the simulated income and health conditions at the small area level are in line with results from the independent IMD/WIMD/SIMD data.

The second dataset we use for external validation is the 2018 income estimates for small areas (IESA) for England and Wales published by the ONS. The 2018 IESA datasets are released at MSOA level and we aggregate our synthesised income for the three English and Welsh cities (Cardiff, Greater Manchester and Sheffield) to match, resulting in 715 MSOAs. We then compare our simulated household income with the confidence intervals of income estimates suggested by the IESA dataset. The comparison results show that the microsimulation derived income estimates fall within the ONS confidence intervals in 97% of MSOAs assessed (696 out of 715).

In summary, both the internal and external validations results above suggest that the simulated population has captured well the differences in individuals' health and income situations at the small area level.

4. Applications of data

There are many ways to use these synthetic data. In SIPHER project, we have utilised these data to calculate equivalent income, which is a preference-based multi-dimensional measure of wellbeing, for each individual and then aggregated for each LSOA in our example cities. We have also calculated SIPHER's inclusive growth indicators, which measure whether an economic system in a place is more or less inclusive compared to other places (e.g. level of employment, poverty, physical connectivity), for each LSOA based on the synthetic data.

Acknowledgments

This work was funded as part of the SIPHER (*Systems Science in Public Health and Health Economics Research*) project with support from the UK Prevention Research Partnership (grant number: MR/S037578/1).

References

Harland, K., Heppenstall, A., Smith, D. and Birkin, M. (2012). Creating realistic synthetic populations

at varying spatial scales: A comparative critique of population synthesis techniques, *Journal of Artificial Societies and Social Simulation*, 15 (1), pp. 1–15.

Harland, K. (2013). *Microsimulation model user guide: Flexible Modelling Framework*. National Centre for Research Methods Working Paper.

Heppenstall, A. and Smith, D. M. (2014). Spatial Microsimulation, in: Fischer, M. M. and Nijkamp, P. (eds.) *Handbook of Regional Science*. Berlin and Heidelberg: Springer, pp. 1235-1252.

Lovelace, R. and Dumont, M. (2016). *Spatial microsimulation with R*. New York: CRC Press.

Tanton, R. (2014). A Review of Spatial Microsimulation Methods, *International Journal of Microsimulation*, 7 (1), pp. 4-25.

Parallel Session PSF3

5 November 11:20 am -1:20 pm (GMT)

Diverse topics in quantitative geography

The Comparative Study of TOD in Metro Station Areas of Guangzhou and Shenzhen Using an Extended Node-Place Model

Yongxin Yang¹

¹Yongxin Yang

University College London, Gower St, Unite Kingdom

Keywords: transit-oriented development, node-place-design model, urban vibrancy, metro station areas

1. Introduction

Node-place (NP) model is an analytical approach to investigate the status quo of transit-oriented development (TOD) of station areas. Current research mainly expands the model by taking the walking environments around the station areas into consideration, while seldomly improves the model in regard of the compactness of spatial form. Besides, few attentions have been given to applying urban vibrancy as a regional review to improve the model. Finally, node-placed extended model generally uses public data as TOD indexes which make it possible to examine and compare the TOD development between different cities as a benchmark.

This work takes Guangzhou and Shenzhen as the study cases, to investigate how well the stations in Guangzhou and Shenzhen are aligned with transit-oriented development and what's their discrepancies. This work has three key findings: (1) Shenzhen is 20% better in terms of the overall performance of TOD. (2) there are significant discrepancies between the groups of stations within each city in terms of TOD (3) the stations in the same groups within each city showed differentiation and variations in terms of TOD and urban vibrancy.

2. Methodology Overview

Based on the interactions between transportation and land use explained by Bertolini (1996), five types of station areas are identified according to the balance of 'node' and 'place' (Fig. 1). First, 'Balance' stations can be found in the domain along the middle diagonal line, where node and place values are relatively equal. At the upper right corner of the diagonal line are station areas under 'Stress', they have tensions due to both very strong node-place value and may results in conflictions. On the contrary, at the bottom areas along the diagonal line, one finds stations labelled as 'Dependency', as

they are unable to sustain themselves, i.e., they may need government support or dependent upon larger transportation stations. Station areas above the domain of middle diagonal line where station areas have stronger node value are labelled as ‘Unsustained Nodes’. Conversely, ‘Unsustained Places’ below the domain of middle diagonal line represent station areas have stronger place value. Both two are the situations where the development dynamics of node-place can be expected, either positive (upgrading) or negative (downgrading). The situations could also suggest the more targeted planning policy for stations according to different upgrading (downgrading) pathways focusing on different aspects of TOD.

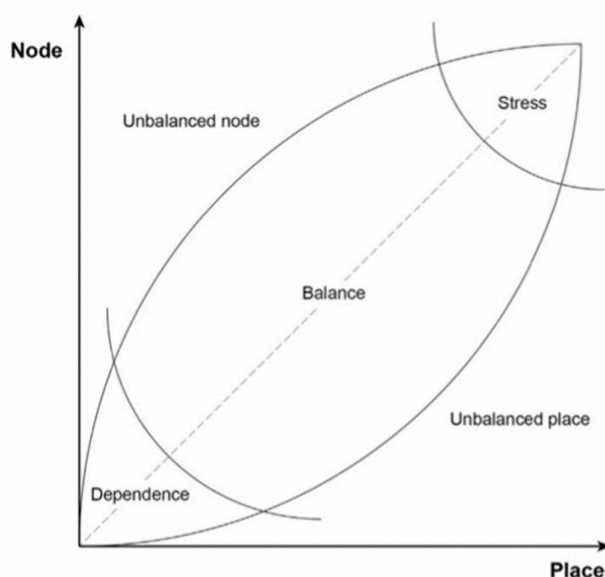


Fig. 1 Node-Place Diagram by Bertolini (1996)

In addition to the two original dimensions: node-place, many studies have added a third dimension, ‘design’, to current node-place model by interpreting the walkability of metro station areas (Vale, 2015; Lyu, Bertolini and Pfeffer, 2016). Following the same logic, this work expands the design dimension by adding indices focusing on the compactness of spatial form, e.g., average height and areas of buildings, functional distance from public facilities to station. Besides, despite all the advantages, the node-place model is still evaluating the metro station areas at local scale while lacking regional perspectives. To address that issue, Zhang, Marshall and Manley introduce the network criticality to node-place-design model (2019). Since few attentions are given to urban vibrancy in regard of TOD studies, this work introduces the concept of urban vibrancy as a regional review to the node-place-design mode. The indicators are calculated as table x

Indicator description	Calculation
Node index	average of indicators
Number of directions served by metro	x1=Number services offered by station
Time duration per train (Frequency)	x2=average time requery of the day
Number of bus stops	x3=bus stops around the station
Beweeness centrality of network (weighted by distance)	x4= local betweenness centrality weighted by line distances between stations
Closeness centrality of network	x5= local betweenness centrality unweighted
Place index	average of indicators
Number of residents	y1=polulation density assigned from raster grid and residential POIs
Number of workers	y2=number of working establishments POIs
Number of workers in leisure industry	y3=number of leisure POIs
Number of workers in catering and restaurant industry	y4=number of catering and restaurant POIs
Number of workers in administration	y5=number of administration POIs
Number of workers in education	y6=number of education POIs
Number of workers in retail industry	y7=number of retail POIs
Number of workers in hotel industry	y8=number of hotel POIs
Design index	average of indicators
Size of building areas	z1=size of areas of building shapes
Average height of building areas	z2=height of the buildings with station area
Length of predestrian network	z3=length of the road network with pedestrian walk paths
Number of intersections of road network	z4=polulation density assigned from raster grid and residential POIs
Functionality toward station	z5=standard distance ¹ calculated by Qgis standard distance plugin
Urban vibrancy	average of indicators
Number of trips originations from station area to whole city	calculated from mobile phone data network
Number of trips detinations to station from whole city	calculated from mobile phone data network
Night time light data	assigned from night light time data of the year

Table 1 The Indicators of TOD Indexes and Urban Vibrancy

3. Results

3.1 node-place-design model for Guangzhou and Shenzhen

Based on the diagram proposed by Bertolini, this work uses hierarchical clustering method to classify the station areas in Guangzhou and Shenzhen under a node-place-design framework. Scaling of indicators and PCA are also adopted to enhance the reliability of clustering analysis.

To directly compare the overall performance, the score is calculated by the ratio of stations with different grade levels (high=3, moderate=2, low=1) in node, place and design dimension. E.g., a station with high grads in node, place and design is assigned with score 9. The results shows that Shenzhen is 20% better in TOD performance (the sum score). However, the basic overall performance didn't gain insights into where the discrepancies between the two cities.

Therefore, the clusters are examined based on the diagram of node-place, as shown in Fig. 2 (a) and Fig. 2 (b). Both cities have a clear “dependency” group (A3, B3), which is low on three dimensions compared to other groups, and a “stressed” group dominating on every dimension (A2, B1). The weaker groups in the two cities all have lower average values in place dimension rather than node and design dimensions, despite there are greater gaps between the better ones and the weaker ones for Guangzhou (A2 compared to A3). In addition, the stations of Shenzhen are more compact than Guangzhou in sub-dimensions scatter plot, indicating that the stations within the same group in Guangzhou have more variations than Shenzhen.

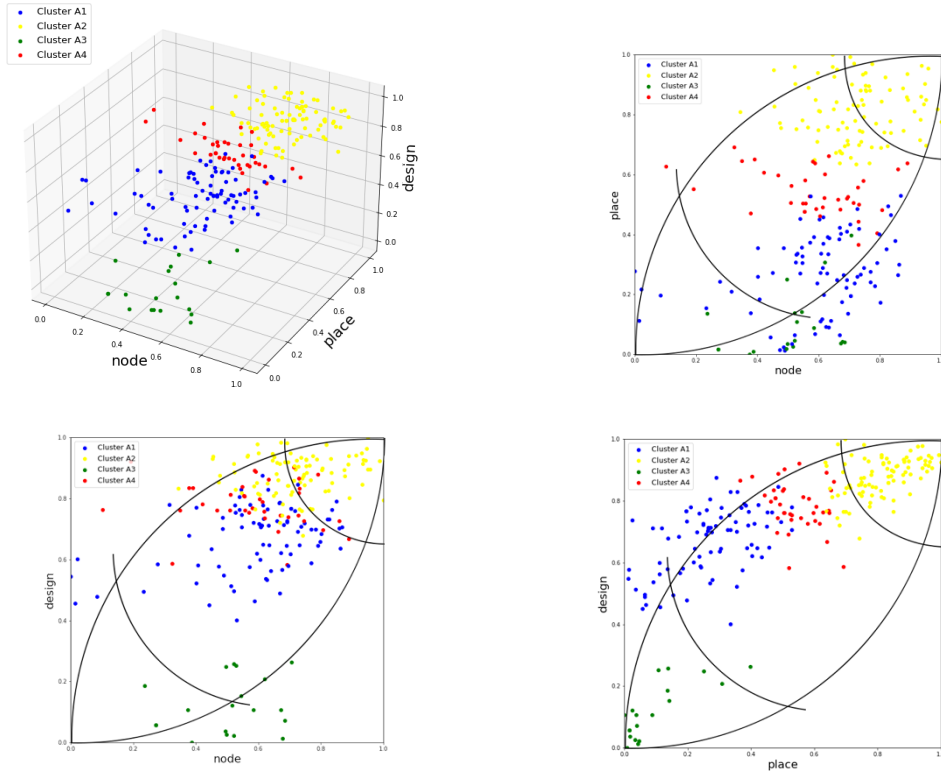


Fig. 2 (a) The Clusters of Stations in Guangzhou in Node-Place-Design Diagram

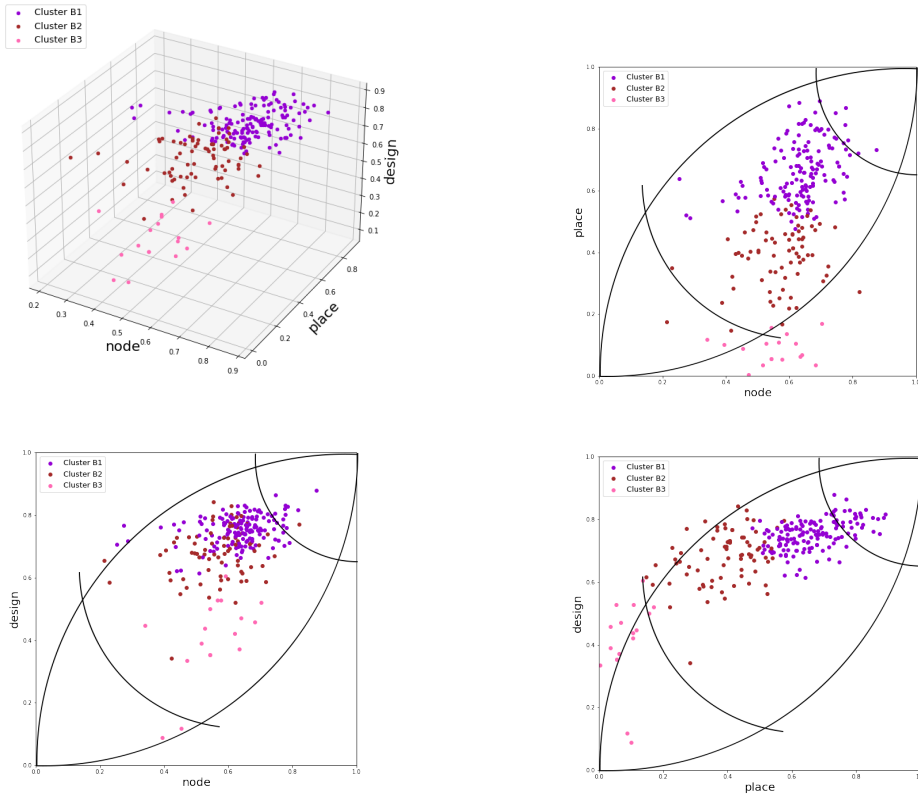


Fig. 2 (b) The Clusters of Stations in Shenzhen in Node-Place-Design Diagram

3.2 node-place-design and urban vibrancy

The urban vibrancy provides us a lens to gain deeper insights into the discrepancies of groups classified by node-place-design model the variations of station within the same group. As shown in Fig. 3 (a) and Fig. 4 (b), two cities exhibit differently in terms of TOD and urban vibrancy in cluster level and station level.

First, the urban vibrancy exhibits in different structure for Guangzhou and Shenzhen: Guangzhou exhibits a core-periphery structure, while Shenzhen exhibits a polycentric pattern. However, they all follow a same degressive logic from central areas to outskirts of the city.

In addition, combining the cluster results of TOD indexes and urban vibrancy, great correlation is seen in both cities. The station with higher urban vibrancy level is mostly grouped in the best performed TOD clusters of stations. However, Guangzhou showed greater variations within the same group, e.g., in cluster A1, certain stations in cluster A1 with high urban vibrancy level are located at the outskirts of the city. This may suggest these stations are critical in the outer regions of the city.

3. Discussions

The findings from this work provides policy implications from different scale and perspective. First, the strategic planning for different cities should be made according to the discrepancies between the groups of stations. As strategic planning is vital for a city's long-term prosperity and benefits, it should schedule the priority of development and investments. Groups like cluster A3 in Guangzhou apparently need to be more focus as it is depending on government's support to keep running, how to build a positive feedback of transportation and land-use in a form of compactness design should be given priority than keeping investing in the superior group like cluster A2. e. Furthermore, policymakers should consider the TOD evaluation of stations areas from both local and regional level, by applying regional concept as a strategic dimension in TOD planning. By doing so, it serves as an analytical tool to identify specific stations that are worth noticing. For example, stations have high urban vibrancy and node value indicating its importance in urban structure, while it is located at the periphery areas of the city. The disruption of these stations' function such as due to delay may cause more severe negative impacts on people who living in these areas. Finally, this work contributes to a more targeted, customized and precise procedure of planning in TOD of metro station areas.

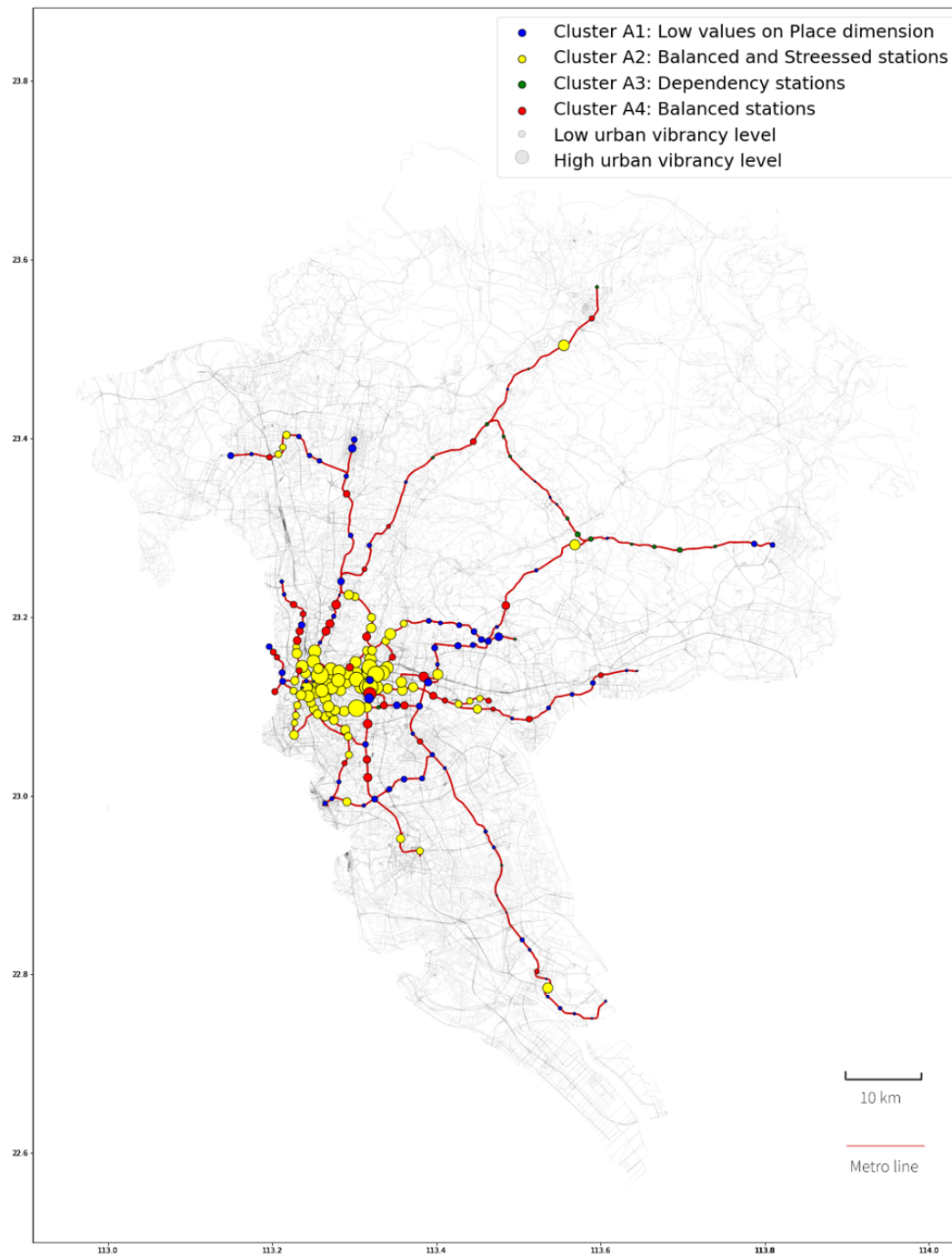


Fig. 3 (a) Spatial Typology of Guanzhou's Stations in TOD and Urban Vibrancy

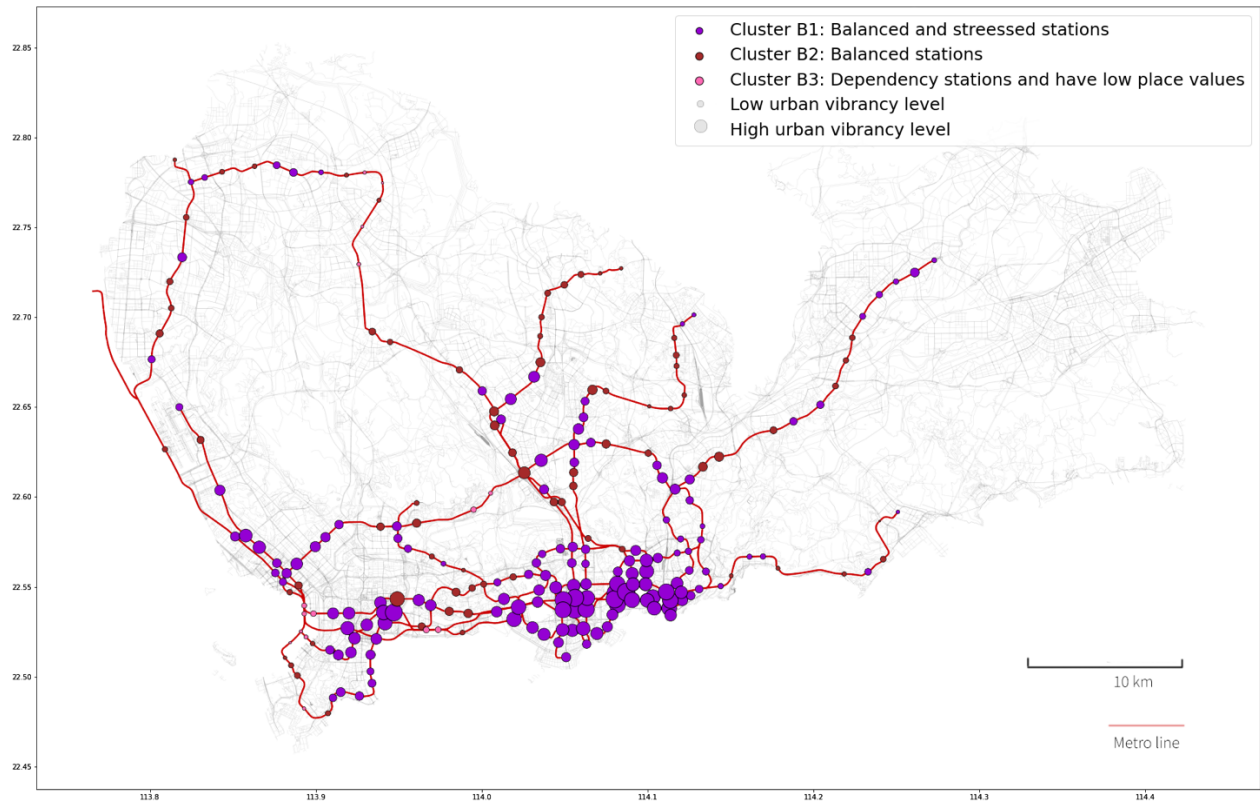


Fig. 4 (b) Spatial Typology of Shenzhen's Stations in TOD and Urban Vibrancy

Acknowledgments

The indicators are sources by open data and calculated by author

References

- Bertolini, L. (1996) 'Nodes and places: Complexities of railway station redevelopment', *European Planning Studies*, 4(3), pp. 331–345. doi: 10.1080/09654319608720349.
- Lyu, G., Bertolini, L. and Pfeffer, K. (2016) 'Developing a TOD typology for Beijing metro station areas', *Journal of Transport Geography*, 55, pp. 40–50. doi: 10.1016/j.jtrangeo.2016.07.002.
- Vale, D. S. (2015) 'Transit-oriented development, integration of land use and transport, and pedestrian accessibility: Combining node-place model with pedestrian shed ratio to evaluate and classify station areas in Lisbon', *Journal of Transport Geography*, 45, pp. 70–80. doi: 10.1016/j.jtrangeo.2015.04.009.
- Zhang, Y., Marshall, S. and Manley, E. (2019) 'Network criticality and the node-place-design model: Classifying metro station areas in Greater London', *Journal of Transport Geography*, 79(August), p. 102485. doi: 10.1016/j.jtrangeo.2019.102485.

Indian Traffic Congestion Model by Shock Wave

Tsutomu TSIBOI¹

¹Nagoya Electric Works Co. Ltd.

29-1 Mentoku Shinoda, Ama-shi, Aichi, 490-1294 Code, Japan

+81-52-443-1111, t_tsuboi@nagoya-denki.co.jp (correspondent author)

Keywords: traffic congestion 1, transportation 2, shock wave 3, traffic flow 4

This research is a series of Indian traffic flow analysis by Japan and Indian government funded project SATREPS (ID: JPMJSA160) since April 2017. The main topic in this research focuses on traffic congestion model by using shock wave theory [1] of fluid flow mechanism. The uniqueness of this research is introduction for two types of traffic congestion model based on more than one month traffic flow monitoring in one of major city “Ahmedabad” in Gujarat state of India. The shock wave analysis has been done several traffic flow analysis but most of research is comparison between traffic flow theory and actual observation [2] [3]. Most of shock wave study for traffic congestion is theoretical analysis plus small size of observation experience, not like our research more than one month observation. Based on recent Indian traffic observation in October 2020, author found two types of traffic congestion models. One is forward forming shock wave and there is no delay between traffic volume and occupancy peak point. The other backward recovery shock wave and there is some gap between traffic flow peak point and occupancy peak point. Occupancy is indicator for the level of traffic congestion in our previous study [4]. And there is unique traffic congestion time from in India especially in the evening time frame rather than in the morning with highest traffic volume [5]. From this shock wave mechanism study, the reason becomes clearer by its model of traffic congestion.

1. Traffic congestion

In general, traffic congestion is described by the one of traffic parameter “occupancy” from traffic theory. The occupancy shows how many vehicles are on the road and its value is percentage. From traffic flow experience, more than 25% occupancy is defined traffic congestion condition. At the same time, we use average vehicles speed on the road is used as reference of traffic congestion condition especially in highway. For human feeling for traffic congestion, we use traffic volume. And it is easy to understand that more traffic volume means traffic congestion as common sense. But sometimes we have different experience which we don’t feel traffic congestion even in heavy traffic flow. This is the reason traffic congestion is caused by not only heavy traffic volume but some other reasons for slow

traffic speed. The occupancy represents both parameter traffic volume and speed. Therefore, its makes more sense to show traffic congestion condition in general.

2. Shock Wave

The shock wave theory comes from fluid mechanism originally but it becomes common for traffic flow analysis as early mentioned. This shock wave provides us for traffic congestion condition when we are in a que of vehicle lane and waiting for movement. When we reach the end of lane of traffic congestion, it seems like its waiting end lane moves backward behind our vehicle. This is the shock wave phenomena in traffic congestion. In shock wave, there are two types of waves in which there are forward forming and backward recovery. The previous our feeling is backward recovery. In figure 1, it is illustrated shock wave type. When shock wave speed is defined as “ c ”, positive vale of “ c ” is forward forming shock wave and negative value of “ c ” is backward recovery.

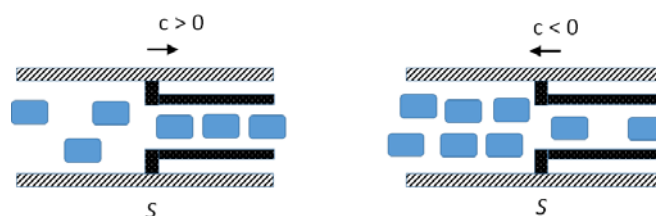
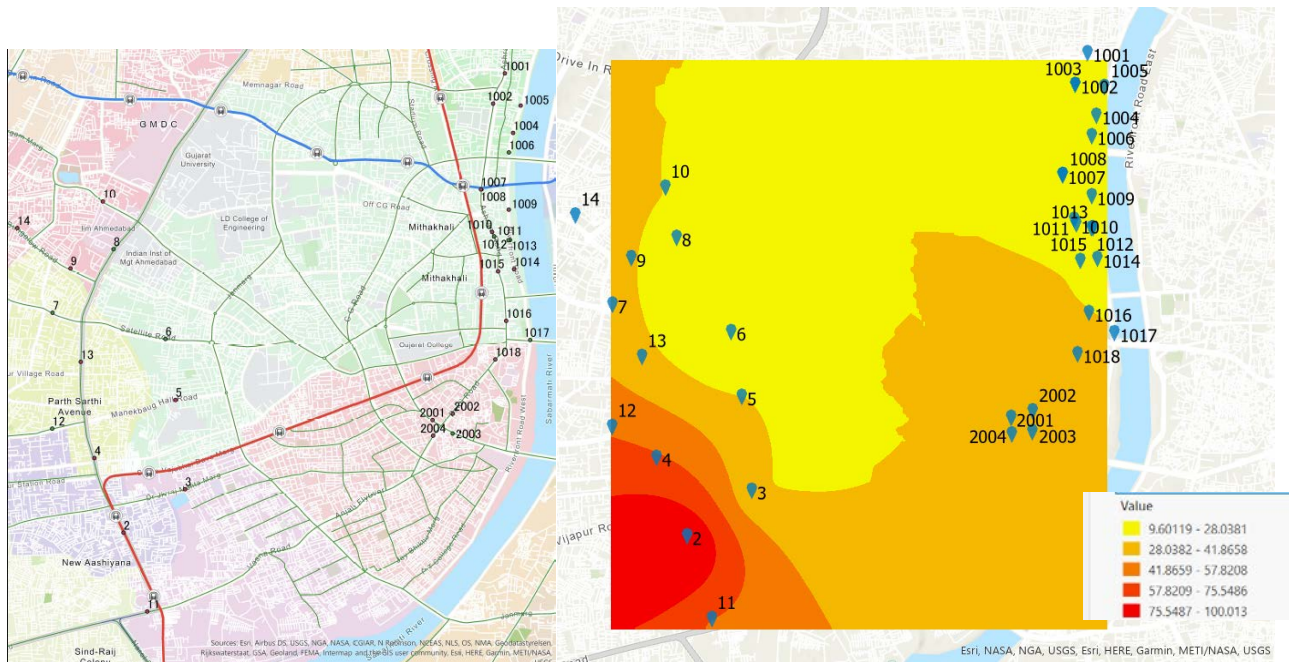


Figure 1: Forward forming shock wave and backward recovering shock wave

3. Traffic congestion observation field

Our traffic congestion field is one of major city in India—Ahmedabad city in Gujarat state of India—the most west located in Indian island. The current population is about 8 million and about 4 million vehicles are registered in 2017. The economic growth is between 5 to 7 % annually, which means rapidly growing city. Therefore, traffic congestion becomes issues for CO2 emission and economical loss by traffic congestion, environment damage and health problem by air pollution, and social loss such as fatality by traffic accidents. Map 1 left shows our traffic observation field in the west side of Ahmedabad where there are new building construction, more market places, and new road expansion etc. It is called “New Town” and the other east side is called “Old Town” where many residents lives from original. Map 1 right shows an example of traffic density at 19:00 in October 2020. The number of each map indicates the location of traffic monitoring camera which we use measure its traffic volume and other traffic flow parameters such as occupancy, traffic density, and average vehicle speed.



Map 1: Traffic condition observation field and an example of traffic density at 19:00 in October 2020.

4. Shock wave traffic congestion analysis

From Map 1 right, we see that there are typical area where there is heavy traffic and potential traffic congestion points, especially left bottom of the Map 1—camera#2 and #11. Table 1 shows the summary of traffic data in each location for only traffic congested time frame between 18:00 to 22:00, which is based on October 2020 one month observation. The highlighted data is more than 25% occupancy point and number of red under-bar is $c < 0$ condition, which means backward recovery shock wave condition. From Table 1, most of traffic congestion is forward forming condition and only camera#11 and camera#2 22:00 condition are in backward recovery condition. This result concludes Ahmedabad city traffic congestion is basically slow moving traffic congestion, not deadlocking condition.

Table 1. The summary of traffic monitoring data of each location

Time	18	19	20	21	22
Camera 2	33	59	47	37	26
3	12	18	15	12	9
4	19	29	24	20	15
5	16	18	14	11	9
6	3	4	4	5	5
7	15	20	15	12	10
8	13	14	9	10	6
9	20	25	21	18	17
10	11	13	9	8	6
11	24	35	25	19	15
13	15	21	16	12	9

1001	14	18	14	12	10
1002	12	15	11	9	8
1003	18	21	16	12	9
1004	6	8	6	6	5
1006	11	16	10	7	5
1007	9	10	7	5	5
1008	13	15	12	9	7
1009	8	11	8	7	5
1010	13	16	12	10	10
1011	11	13	10	8	8
1013	9	10	7	6	5
1014	7	9	6	5	5
1018	19	27	22	19	17
2003	14	22	19	16	14
2004	16	23	19	16	15

Based on all observation data, here is an important about traffic congestion time zone characteristics in Figure 2. There are two time zone based traffic volume and occupancy relationship at camera#2 and camera#11. There is no time lag between traffic volume peak and occupancy peak at camera#2 (Forward forming shock wave condition). But there is 1 hour time lag between traffic volume peak and occupancy peak camera#11 (Backward recovery shock wave condition).

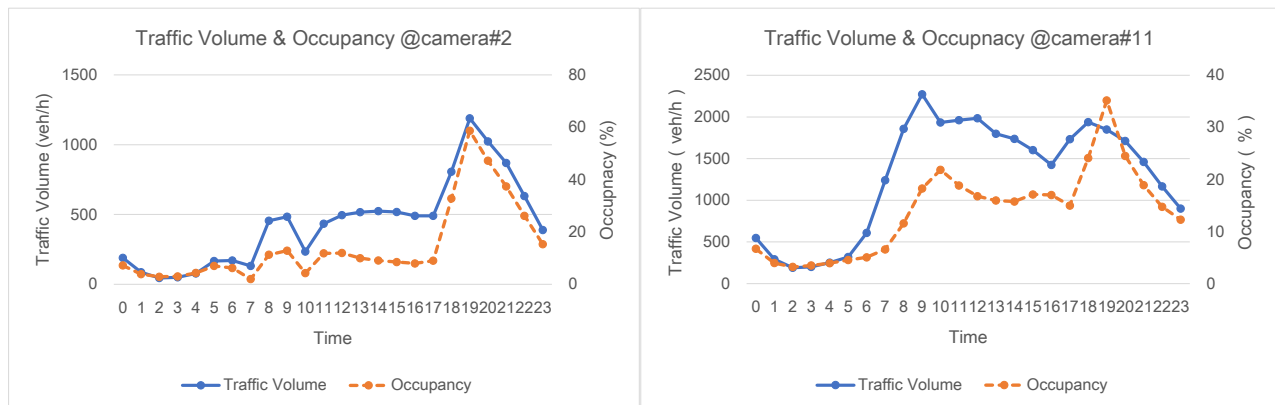


Figure 2. Time zone based traffic volume & occupancy relationship

5. Conclusion

In this study, we use shock wave mechanism for Indian traffic congestion analysis. And we have two types of traffic congestion model—forward forming shock wave condition and slow moving traffic congestion—backward recovery shock wave condition and traffic stack congestion. Most of traffic congestion in Ahmedabad city is forward forming shock wave and slow moving traffic congestion especially in the evening time frame. There is not so much traffic congestion in the morning even under more traffic volume in the morning. From this analysis, the traffic congestion in Ahmedabad has some unique mechanism of congestion and when this mechanism is found, there is a solution for improving traffic congestion condition in future. This research is still on-going project until September

2022. We will continue to analyse traffic congestion with more data collection and other spatial analysis such as Moran' I indication and investigation for social activities in the city in future.

Acknowledgments

Please write any acknowledgements or funding information here.

References

- [1] Pipes L.A., 1961, 'Hydrodynamic Approaches—Part I, An Introduction to Traffic Flow Theory, H.R., Special Report 79, pp.3-6.
- [2] Kune R. Michalopolos P., 'Continuum Flow Models Chapter 5' pp.3-5, accessed 13 September 2021, < <https://www.fhwa.dot.gov/publications/research/operations/tft/chap5.pdf>>
- [3] Sock wave analysis, Chapter 8, accessed 13 September 2021, <<https://ocw.tudelft.nl/wp-content/uploads/Chapter-8.-Shock-wave-analysis.pdf>>
- [4] Tsuboi.T, 2020, 'Traffic Congestion Analysis and Occupancy Parameter in India', International Conference on Data Analytics and Management.
- [5] Tsuboi.T, 2021, 'Visualization and Analysis of Traffic Flow and Congestion in India', MPDI Open Access Journal Infrastructures

Preliminary discoveries on the joint impacts of the modifiable areal unit problem (MAUP) and commission error on the classical linear regression model (CLRM)

Xiang YE^{1,2}; Renzhong GUO^{1,2}; Siru LI^{1,2}; Ding MA^{1,2}; Zhigang ZHAO^{*,1,2}

¹Research Institute for Smart Cities, Shenzhen University

No. 3688 Nanhai Boulevard, Shenzhen, Guangdong 5518060, P. R. China

²School of Architecture and Urban Planning, Shenzhen University

No. 3688 Nanhai Boulevard, Shenzhen, Guangdong 5518060, P. R. China

zhaozgrisc@szu.edu.cn (correspondent author*)

Keywords: modifiable areal unit problem, linear regression, commission error, irrelevant variable, moment

1. Motivation

The classical linear regression model (CLRM) is one of the most frequently used statistical models for social science. By adopting the model, the direction and magnitude of causality of a group of independent variables to a dependent variable are estimated. By default, the estimation is implemented via the ordinary least square (OLS) method, which is reputed for providing unbiased and efficient estimates (Greene 2012, pp. 100-101). Nevertheless, these properties may be no longer available if the regression model is misspecified.

Among all sorts of model specification errors, commission error (Wooldridge 2015, pp. 78-81) happens when irrelevant independent variables are included in a regression model. In this case, OLS estimators of regression coefficients are no longer efficient (Greene 2012, p. 98), and the coefficient of determination, R^2 , increases (Ohtani 1994). Meanwhile, commission error interacts with other issues like multicollinearity (Sifuentes-Amaya and Ramírez-Valverde 2010) and omission error (Basu 2020) and complicates the situation. However, whether commission error interacts with the modifiable areal unit problem (MAUP) (Gehlke and Biehl 1934; Openshaw and Taylor 1979; Wong 2009) has not been studied.

The MAUP is one of the most fundamental and long-lasting issues in Geography (Páez and Scott 2005). It refers to the sensitivity and/or inconsistency of analysis results due to different configurations of the same study area (Openshaw 1983). The impacts of the MAUP in empirical studies are well documented (Jelinski and Wu 1996; Zhang and Kukadia 2005; Dark and Bram 2007; Xu et al. 2018), but very few studies aim at the root causes of such impacts from the theoretical side.

In practice, both commission error and the MAUP can occur simultaneously whenever a linear regression model is applied to a spatially aggregated dataset. Will the impacts of the two issues interact? Does it matter? Should we care? To address all these concerns, it is of significance to discover the joint impacts of the two issues on linear regression via a theoretical and quantitative approach.

2. A regression model with commission error¹

A commission error occurs when irrelevant variables are included in the regression:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \ddot{\mathbf{X}}\ddot{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{1}$$

where \mathbf{y} is an $n \times 1$ vector of the dependent variable, \mathbf{X} is an $n \times k$ matrix for (necessary) independent variables, and $\boldsymbol{\beta}$ is a $k \times 1$ vector of coefficients for \mathbf{X} . Without loss of generality, $\ddot{\mathbf{X}}$ is an $n \times k_0$ matrix of irrelevant variables and $\ddot{\boldsymbol{\beta}}$ is a $k_0 \times 1$ vector of coefficients for $\ddot{\mathbf{X}}$. By definition, $\ddot{\boldsymbol{\beta}} = \mathbf{0}_{k_0 \times 1}$. The last term, $\boldsymbol{\varepsilon}$, records the disturbance.

Unfortunately, from the eyes of a researcher, equation (9) is not observable, as \mathbf{X} and $\ddot{\mathbf{X}}$ are not distinguishable but have to be treated together (otherwise, he/she would not commit this commission error):

$$\mathbf{y} = \dot{\mathbf{X}}\dot{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{2}$$

where $\dot{\mathbf{X}} = [\mathbf{X} \ \ddot{\mathbf{X}}]$ and $\dot{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{\beta} \\ \ddot{\boldsymbol{\beta}} \end{bmatrix}$. Correspondingly, for estimation purposes, the sample regression model is

$$\mathbf{y} = \dot{\mathbf{X}}\dot{\mathbf{b}} + \dot{\mathbf{e}} \tag{3}$$

where $\dot{\mathbf{b}} = \begin{bmatrix} \dot{\mathbf{b}}_1 \\ \dot{\mathbf{b}}_2 \end{bmatrix}$ estimates the two parts of $\dot{\boldsymbol{\beta}}$, and $\dot{\mathbf{e}}$ is the vector of residuals.

3. A regression model with the MAUP

When the MAUP is present, individual observations are no longer available, but merged to the aggregate level, like counties or census tracts. Typically the process can be represented by the merging matrix (Ye and Rogerson 2021), \mathbf{M} :

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n_1} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & m_{21} & m_{22} & \cdots & m_{2n_2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & m_{m1} & m_{m2} & \cdots & m_{mn_m} \end{bmatrix} \tag{4}$$

¹ The setting of the commission error is introduced in any econometric textbook, but it is included here to serve a better basis for the following discussions.

In equation (4), \mathbf{M} is an $m \times n$ matrix, with n represents the number of individual observations before aggregation, and m represents the number of regions after aggregation. Each element of \mathbf{M} , m_{ij} , represents the weight applied to the individual observation j when it is merged into region i , $1 \leq j \leq n_i$, $1 \leq i \leq m$, and n_i is the total number of individuals in region i .

Hence, the linear regression at the aggregated level can be modelled as (Ye and Rogerson 2021)

$$\mathbf{M}\mathbf{y} = \mathbf{M}\mathbf{X}\boldsymbol{\beta} + \mathbf{M}\boldsymbol{\varepsilon} \quad (5)$$

When the researcher does not realize the appearance of the MAUP, he/she will not notice \mathbf{M} in equation (5). So the regression model at the aggregate level looks like (Ye and Rogerson 2021):

$$\mathbf{y}^\circ = \mathbf{X}^\circ\boldsymbol{\beta} + \boldsymbol{\varepsilon}^\circ \quad (6)$$

where $\mathbf{y}^\circ = \mathbf{M}\mathbf{y}$, $\mathbf{X}^\circ = \mathbf{M}\mathbf{X}$, and $\boldsymbol{\varepsilon}^\circ = \mathbf{M}\boldsymbol{\varepsilon}$ are of dimension $m \times 1$, $m \times k$, and $m \times 1$, respectively. A circle is added to denote the notation is at the aggregate level.

Hence, the corresponding sample regression model is (Ye and Rogerson 2021)

$$\mathbf{y}^\circ = \mathbf{X}^\circ\mathbf{b}^\circ + \mathbf{e}^\circ \quad (7)$$

where \mathbf{b}° estimates $\boldsymbol{\beta}$, and \mathbf{e}° conveys residuals.

4. A regression model impacted by both issues

When the linear regression model is suffered from both commission error and the MAUP, the actual model can be described as

$$\mathbf{y}^\circ = \check{\mathbf{X}}^\circ\check{\boldsymbol{\beta}} + \check{\boldsymbol{\varepsilon}}^\circ \quad (8)$$

where $\check{\mathbf{X}}^\circ = [\mathbf{X}^\circ \quad \check{\mathbf{X}}^\circ]$, an $n \times (k + k_0)$ matrix, is the mixture of both necessary and irrelevant independent variables at the aggregate level, and $\check{\boldsymbol{\beta}}$ and $\check{\boldsymbol{\varepsilon}}^\circ$ correspondingly represent coefficients and disturbances. Therefore, the researcher will solve the following sample model:

$$\mathbf{y}^\circ = \check{\mathbf{X}}^\circ\mathbf{b}^\circ + \mathbf{e}^\circ \quad (9)$$

where $\mathbf{b}^\circ = \begin{bmatrix} \mathbf{b}_1^\circ \\ \mathbf{b}_2^\circ \end{bmatrix}$ estimates the two parts of $\check{\boldsymbol{\beta}}$, leaving \mathbf{e}° as residuals.

5. The expression and moments of the estimators

When both issues are not aware of, the OLS method will be adopted by default:

$$\mathbf{b}^\circ = (\check{\mathbf{X}}^{\circ T}\check{\mathbf{X}}^\circ)^{-1}\check{\mathbf{X}}^{\circ T}\mathbf{y}^\circ \quad (10)$$

According to the previous discussion, we know that \mathbf{b}° contains two parts: \mathbf{b}_1° estimates the true regression coefficients $\boldsymbol{\beta}$, and \mathbf{b}_2° estimates $\check{\boldsymbol{\beta}}$, which would be a vector of zeros. It can be proved that

$$\mathbf{b}^\circ = \begin{bmatrix} \mathbf{b}_1^\circ \\ \mathbf{b}_2^\circ \end{bmatrix} = \begin{bmatrix} (\mathbf{X}^{\circ T} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \mathbf{X}^\circ)^{-1} \mathbf{X}^{\circ T} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \mathbf{y}^\circ \\ (\ddot{\mathbf{X}}^{\circ T} \mathbf{R}(\mathbf{X}^\circ) \ddot{\mathbf{X}}^\circ)^{-1} \ddot{\mathbf{X}}^{\circ T} \mathbf{R}(\mathbf{X}^\circ) \mathbf{y}^\circ \end{bmatrix} \quad (11)$$

where $\mathbf{R}(\mathbf{X}) = \mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ and \mathbf{X} can be replaced as needed.

We care if \mathbf{b}° still behave. The good news is that \mathbf{b}° is still unbiased for $\boldsymbol{\beta}$, elementwise:

$$\mathbb{E}(\mathbf{b}_1^\circ | \dot{\mathbf{X}}, \mathbf{M}) = \boldsymbol{\beta} \quad (12)$$

$$\mathbb{E}(\mathbf{b}_2^\circ | \dot{\mathbf{X}}, \mathbf{M}) = \mathbf{0}_{k_0 \times 1} \quad (13)$$

But its variance is way more entangled:

$$\mathbb{V}(\mathbf{b}_1^\circ | \dot{\mathbf{X}}, \mathbf{M}) = \sigma^2 (\mathbf{X}^{\circ T} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \mathbf{X}^\circ)^{-1} \mathbf{X}^{\circ T} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \boldsymbol{\Lambda} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \mathbf{X}^\circ (\mathbf{X}^{\circ T} \mathbf{R}(\ddot{\mathbf{X}}^\circ) \mathbf{X}^\circ)^{-1} \quad (14)$$

$$\mathbb{V}(\mathbf{b}_2^\circ | \dot{\mathbf{X}}, \mathbf{M}) = \sigma^2 (\ddot{\mathbf{X}}^{\circ T} \mathbf{R}(\mathbf{X}^\circ) \ddot{\mathbf{X}}^\circ)^{-1} \ddot{\mathbf{X}}^{\circ T} \mathbf{R}(\mathbf{X}^\circ) \boldsymbol{\Lambda} \mathbf{R}(\mathbf{X}^\circ) \ddot{\mathbf{X}}^\circ (\ddot{\mathbf{X}}^{\circ T} \mathbf{R}(\mathbf{X}^\circ) \ddot{\mathbf{X}}^\circ)^{-1} \quad (15)$$

where $\boldsymbol{\Lambda} = \mathbf{M} \mathbf{M}^T$. By the Gauss-Markov theorem (Greene 2012, p. 100), $\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ is the best linear unbiased estimator (BLUE) for $\boldsymbol{\beta}$, so \mathbf{b}_1° will be less efficient for $\boldsymbol{\beta}$ with no surprise. However, it is an interesting question if \mathbf{b}_1° is also less efficient than \mathbf{b}_1 and \mathbf{b}° ; put it in plain words, will the inefficiency brought by commission error and the MAUP accumulate? Results from Monte Carlo simulation said yes, but we haven't secured a solid proof for that. If the answer is true, then the behavior of the MAUP is understood as expected. If the answer is false, however, it could be more exciting: the cumulative effect from the two issues are not additive; it opens up the possibility that in some cases, the MAUP can shrink the variance of estimators and improve efficiency.

References

- Basu, D. (2020). 'Bias of ols estimators due to exclusion of relevant variables and inclusion of irrelevant variables' *Oxford Bulletin of Economics and Statistics*, 82 (1), pp. 209-234. DOI: <https://doi.org/10.1111/obes.12322>.
- Dark, S. J. & Bram, D. (2007). 'The modifiable areal unit problem (maup) in physical geography' *Progress in Physical Geography*, 31 (5), pp. 471-479.
- Gehlke, C. E. & Biehl, K. (1934). 'Certain effects of grouping upon the size of the correlation coefficient in census tract material' *Journal of the American Statistical Association*, 29 (185), pp. 169-170. DOI: 10.2307/2277827.
- Greene, W. H. (2012). *Econometric analysis*. 7th edn. Essex, England: Pearson Education Limited, p. 1241.
- Jelinski, D. E. & Wu, J. (1996). 'The modifiable areal unit problem and implications for landscape ecology' *Landscape Ecology*, 11 (3), pp. 129-140.
- Ohtani, K. (1994). 'The density functions of r^2 and r^2 , and their risk performance under asymmetric loss in misspecified linear regression models' *Economic Modelling*, 11 (4), pp. 463-471. DOI: [https://doi.org/10.1016/0264-9993\(94\)90003-5](https://doi.org/10.1016/0264-9993(94)90003-5).
- Openshaw, S. (1983). *The modifiable areal unit problem. Concepts and techniques in modern geography (camtog)* Norwick: Geo Books.

- Openshaw, S. & Taylor, P. J. (1979). 'A million or so correlation coefficients: Three experiments on the modifiable areal unit problem', in Wrigley, N. & Bennett, R. J. (eds.) *Statistical applications in the spatial sciences*. London: Pion, pp. 127-144.
- Páez, A. & Scott, D. M. (2005). 'Spatial statistics for urban analysis: A review of techniques with examples' *GeoJournal*, 61 (1), pp. 53-67. DOI: 10.1007/s10708-005-0877-5.
- Sifuentes-Amaya, R. & Ramírez-Valverde, G. (2010). 'Efectos de especificar un modelo incorrecto para regresión logística, con dos variables independientes correlacionadas' *Agrociencia*, 44 pp. 197-207.
- Wong, D. W.-S. (2009). 'The modifiable areal unit problem (maup)', in Fotheringham, A. S. & Rogerson, P. A. (eds.) *The sage handbook of spatial analysis*. London: Sage Publications Ltd., pp. 105-123.
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. 6th edn. Boston, MA: Cengage Learning.
- Xu, P., Huang, H. & Dong, N. (2018). 'The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research' *Journal of Traffic and Transportation Engineering (English Edition)*, 5 (1), pp. 73-82. DOI: <https://doi.org/10.1016/j.jtte.2015.09.010>.
- Ye, X. & Rogerson, P. (2021). 'The impacts of the modifiable areal unit problem (maup) on omission error' *Geographical Analysis*, pp. 1-26. DOI: <https://doi.org/10.1111/gean.12269>.
- Zhang, M. & Kukadia, N. (2005). 'Metrics of urban form and the modifiable areal unit problem' *Transportation Research Record: Journal of the Transportation Research Board*, 1902 (1), pp. 71-79. DOI: 10.3141/1902-09.

A geophysical model for estimating CO₂ balance in tree crops in Mediterranean region

Angelos MIMIS¹; Kostas BITHAS²

¹Associate Professor

Institute of Urban Environment and Human Resources (UEHR)

Dept. Economics and Regional Development,

Panteion University, Athens

Email address: mimis@panteion.gr

²Professor

Institute of Urban Environment and Human Resources (UEHR)

Dept. Economics and Regional Development,

Panteion University, Athens

Email address: kbithas@panteion.gr

Keywords: carbon sequestration, model, tree crops, Mediterranean region

Abstract

In this research, a CO₂ sequestration model is described, which is dynamic and runs on a monthly timestep. Spatially, the model can run for any region in NUTS 1, 2 and 3 spatial level of detail. The model consists of the biomass (trunk and roots), the debris pool and the soil. The biomass grows with time and accordingly the roots. The pruning (if left on the field) and the crop losses are feeding the debris and then together with the root's exudates are feeding the soil. The soil processes are modeled by RothC which are affected by climatic data and soil characteristics. The model is customized for five species of perennial trees most common in Mediterranean area and are the olive, orange, apple, almond, peach trees.

1. Introduction

The Article 3.4 of the Kyoto protocol has generated a high interest in quantification of the carbon sequestration produced by different land uses. As Smith et al. (2000) conclude there is a considerable potential for carbon dioxide mitigation in European agriculture. In that respect, a dynamic model simulating carbon dioxide (CO₂) sequestration by tree-crops is developed and consists by three pools (Figure 1). The first one will include biomass (BM), the second one, debris pool (DP) and the third one the soil organic matter (SOM).

The 1st pool is connected to 2nd pool through pruning, and crops left to ground. It is also connected through the roots to the soil directly. The material in DP decomposes and feeds the Soil. The conceptual description of the model can be seen in Figure 1.

Having shown the conceptual model, we will continue in the next sections with the equations and the dataset supporting the model, and finally the illustration of basic results produced by the model.

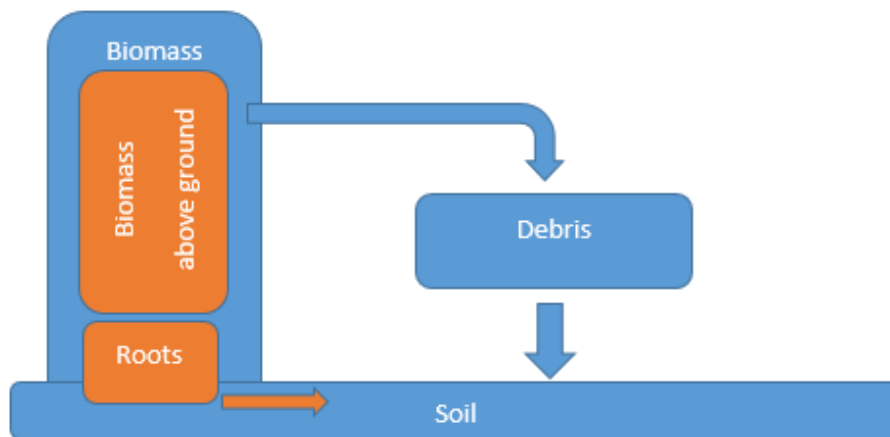


Figure1. Conceptual model.

2. Methodology and datasets

The three pools of the model will be briefly described along with the relevant datasets.

2.1 Biomass

The 1st pool consists of two sub-pools. The biomass above ground and the roots.

In our model, the trunk biomass grows at a variable rate given by the equation:

$$trunk\ biomass(t) = a * t^{0.3} \quad (1)$$

where t is the time and a is a constant specified for each tree in Table 1. ¹

	a
<i>Olive</i>	10.3
<i>Orange</i>	24.87
<i>Apple</i>	53.97
<i>Almond</i>	5.08
<i>Peach</i>	3.39

Table 1. Constant values for trunk growth rate.

¹ Constant a is determined by measurements in the project CLIMATREE and Scandellari et al. (2016).

Having in mind, the calculation for biomass above ground, the biomass formulas in roots are adopted by Farina et al. (2017; 2013).

2.2 Debris pool

The 2nd pool, the debris pool, has as an input the plant residues (Mg C ha⁻¹) including pruning and fruit products left on the field. The pruning values and crop losses can be seen in Table 2 and are divided in the relevant months that these actions take place.

	pruning (tn ha⁻¹ year⁻¹)	% crop losses
<i>Olive</i>	1.35	10%
<i>Orange</i>	1.27	20%
<i>Apple</i>	0.91	20%
<i>Almond</i>	0.85	10%
<i>Peach</i>	1.45	20%

Table 2. Pruning values (tn ha⁻¹ year⁻¹).

3.3 Soil

In our model, the dynamics of carbon in soil, are assessed spatially and temporally at regional scale based on RothC model, version 26.3 (Coleman and Jenkinson, 1996). The original model is extended in order to run spatially and is combined with (a) the administrative boundaries defined by NUTS 1, 2 and 3, for the countries under study, and (b) a spatial database including soil and climate characteristics.

In order to run the model, we need to define the input carbon, and datasets used to setup the parameters of the model (e.g. soil characteristics).

The monthly input carbon (t C ha⁻¹) that flows into that pool, as it has been described, is a combination of the carbon in debris pool, the carbon in roots exudates and the carbon weeds.

The data required to setup the parameters of the model are:

- Climatological data. These include monthly rainfall (mm), open pan evaporation (mm) and average air temperature (°C). In our case, these climatic data are used for two periods. The first is a monthly average for the period 2008-2012 and the second for the period 2048-2052 (simulated).

- Soil characteristics. The clay content on topsoil (%). For that purpose, the LUCAS 2009 TOPSOIL dataset is used (Orgiazzi et al., 2018).
- An estimate of the decomposability of the incoming plant material (DPM/RPM ratio). The value proposed by Coleman and Jenkinson, 1996 is used.
- Monthly input of monthly farmyard manure (FYM, in t C ha⁻¹), if any.

So firstly, the model is implemented in programming language R and is distributed as an open source software in github². The repository includes all the data produced and the secondary data collected and needed to make the carbon calculations.

Secondly, a dashboard is developed in Shiny (shiny.rstudio.com) and a web deployment is available for use by avoiding all the technicalities of the model³.

3. Results

In this section, typical results, that can be produced by the model, will be illustrated.

The results are displayed on an aggregate country level, although the runs were made on a larger scale. The first table of results (Table 3) shows the total carbon sequestration for the countries in our study, in 50 years' time. For these calculations, the actual statistical data for surface, yield, and plant density has been used.

	Greece	Spain	Italy
<i>olive</i>	27,67	220,71	36,49
<i>orange</i>	1,96	11,55	3,35
<i>apple</i>	1,35	2,95	12,71
<i>almond</i>	0,43	14,64	1,54
<i>peach</i>	1,34	1,68	2,82

Table 3. Total carbon (in Mt)

It should be noted here that approximately 80% of the carbon, in Table 3, is stored in soil.

The next table (Table 4) shows the effect of future temperature change on carbon sequestration.

² The code and the data are available in <https://github.com/amimis/climatree> with GPL-3.0 license.

³ The web application is in <https://amimis.shinyapps.io/climatree/>

	+ 0.0 C	+1.0 C	+2.0 C	+5.0 C
<i>Greece</i>	31,69	27,23	26,20	23,31
<i>Spain</i>	227,65	219,12	210,84	187,58
<i>Italy</i>	37,52	36,13	34,78	30,94

Table 4. Total carbon for olives (in Mt)

The last table (Table 5) displays the effect of (a) keeping the pruning material on the field and (b) keeping the field vegetated, have on the carbon sequestration.

	Greece	Spain	Italy
<i>keep pruning & keep vegetation</i>	27,67	220,71	36,49
<i>no pruning & keep vegetation</i>	25,05	196,50	32,80
<i>no pruning & no vegetation</i>	14,31	106,89	18,75
<i>keep pruning & no vegetation</i>	16,21	124,16	21,43

Table 5. Total carbon for olives (in Mt)

Similar results can be acquired for regions in NUTS1, 2 or 3 and for different tree species present in our study.

4. Conclusions

The model can run on an aggregate level, for the entire country, or on specific regions. There are several other options useful to make comparisons and assess management practices. These are compiled on the following list:

- Five different tree types are available.
- Location characteristics affects the carbon sequestration.
- Soil vegetation affects the carbon capacity to hold carbon in soil.
- Pruning left on the field.
- Age of the trees. So new planting options can be assessed.
- Future climatological scenarios.
- Plant density.

Lastly, the model has been developed in such a way that permits its use by other researchers by providing the source code, the full dataset and an illustrative web application.

Acknowledgments

This work is funded by the CLIMATREE Project «A novel approach for accounting & monitoring carbon sequestration of tree crops and their potential as carbon sink areas» (CLIMATREE LIFE14 CCM/GR/000635).

References

- Coleman K., Jenkinson D.S. (1996) RothC-26.3 - A Model for the turnover of carbon in soil. In: Powlson D.S., Smith P., Smith J.U. (eds) Evaluation of Soil Organic Matter Models. NATO ASI Series (Series I: Global Environmental Change), vol 38. Springer, Berlin, Heidelberg.
- Farina R., Marchetti A., Francaviglia R., Napoli R. and Di Bene C. (2017). Modeling regional soil C stocks and CO₂ emissions under Mediterranean cropping systems and soil types. *Agriculture, Ecosystems and Environment* 238, 128-141.
- Farina R., Coleman K., Whitmore AP. (2013). Modification of RothC model for simulations of soil organic C dynamics in dryland regions. *Geoderma* 200-201, 18-30.
- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O. 2018. LUCAS Soil, the largest expandable soil dataset for Europe: A review. *European Journal of Soil Science*, 69(1): 140-153.
- Scandellari F., Caruso G., Meggio F, Palese Assunta M, Zanotelli D., Celano G., Gucci R., Inglese P., Pitacco A. and Tagliavini M. (2016). A survey of carbon sequestration potential of orchards and vineyards in Italy, *European Journal of Horticultural Science* 81(2), 106-114.
- Smith, P., Powlson, D.S., Smith, J.U., Falloon, P., Coleman K., 2000. Meeting Europe's climate change commitments: quantitative estimates of the potential for carbon mitigation by agriculture. *Global Change Biology*, 6, 525-539.

Identifying trip purpose from a dockless bike-sharing system in Manchester

Antonio Ross¹; Neil Walton²; Nuno Pinto³

¹ Department of Social Statistics, The University of Manchester

Manchester M13 9PL, United Kingdom

luis.rossperez@manchester.ac.uk

² Alan Turing Building, The University of Manchester

Manchester M13 9PL, United Kingdom

neil.walton@manchester.ac.uk

³ Spatial Policy and Analysis Lab, Manchester Urban Institute, The University of Manchester

Manchester M13 9PL, United Kingdom

nuno.pinto@manchester.ac.uk

Keywords: Dockless bike-sharing system, Trip purpose, Bayes' rule, Activity inference, GPS

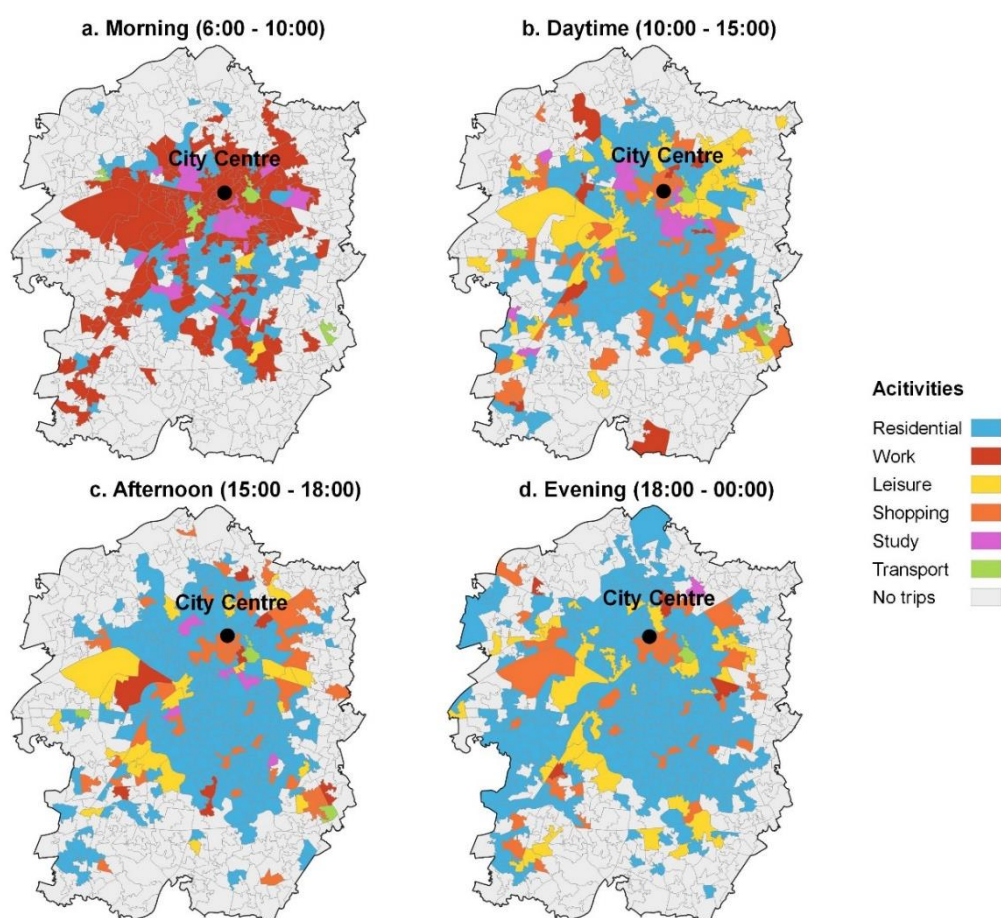
This study explores the application of a Bayes ruled-based algorithm to uncover the purpose of trips from a dockless bike-sharing system in Manchester. The purpose is to produce a picture of the activity patterns at an aggregated level to understand the role of such service. We relied on GPS data generated from every bicycle and a geospatial dataset of buildings and parks in the Manchester central area. The method proved to uncover trip purposes that, otherwise, would only be known through a survey and capable of distinguishing between activities in mixed land-use areas.

Two main datasets were used for the research. A dataset containing individual bike trips and a geospatial dataset containing all the buildings in the Manchester Central area. The first exclusively includes trips from working days in the period of June to September 2017, resulting in 57,584 records. The variables available include a trip's destination (longitude and latitude coordinates) and timestamps. The second contains 426,892 properties and variables include the general purpose of the building (building use), the size of a property and height in square metres. In contrast to research based on docked BSS, we expect users to cycle as close as possible to their destination drawing from the widely-researched concept that people travel to engage in activities rather than for pleasure (Anda et al., 2017; Axhausen and Gärling, 1992; Beecham et al., 2014; Kitamura, 1988). Therefore, our assumption is that a bike's final location is directly linked to the surrounding available activities that a user can engage in.

To our knowledge, the amount of research that has attempted to estimate the purpose of a trip in dockless BSS is limited. The method used in this paper is based on the study by Gong et al. (2016) who estimated the purpose of taxi trips based on the drop-off points obtained through GPS data. Their rule-based method is underpinned by Bayes rules and calculates the probability of visiting a surrounding building given: i) a walking distance radius, ii) a distance decay function to every Point-of-Interest (POI), and iii) the time dynamics of every location. Our study draws from this approach and uses GPS data from dockless bikes as well as a geospatial dataset containing all the buildings and green infrastructure in Central Manchester. As a result, this paper has two main contributions. Firstly, to obtain a reliable estimation of main user's activities at an aggregated level which are to be surmised based on probability. Secondly, to enable the discussion of NMS travel patterns and their role in modern cities.

In this paper, the method is employed to establish an association between bike trips and land use. Specifically, it calculates the weighted relationship between a property, either a building or a green area (O_i), and a bike's final location (B) using their geographical coordinates and the time of the day (t). However, only those locations within a walkable distance radius (δ) from the bicycle will be considered and they are defined as candidates (OC_i). As a result, the final allocation between a bike and a property can be understood as the most likely location that a user visited.

Our findings indicate that the user's relationship with the service is complex. Trips to residential areas are the most important ones, especially in the evening whilst trips to work are predominant in the morning (see Map 1). This produces an inward-outward movement of bicycles throughout the day. With over a quarter of the total trips, retail activities proved to be more popular than cycling to education. The study provides an insight into people's travel preferences when allowed to cycle freely through the city and searches to bridge the gap between new mobility systems and traditional transport strategies.



Map 1: Temporal progression of main activity allocations per LSOA from the proposed model.

References

- Anda, C., Erath, A., Fourie, P.J., 2017. Transport modelling in the age of big data. *Int. J. Urban Sci.* 21, 19–42. <https://doi.org/10.1080/12265934.2017.1281150>
- Axhausen, K., Gärling, T., 1992. Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. *Transp. Rev.* 12, 323–341.
- Beecham, R., Wood, J., Bowerman, A., 2014. Studying commuting behaviours using collaborative visual analytics. *Comput. Environ. Urban Syst.* 47, 5–15.
- Gong, L., Liu, X., Wu, L., Liu, Y., 2016. Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartogr. Geogr. Inf. Sci.* 43, 103–114. <https://doi.org/10.1080/15230406.2015.1014424>
- Kitamura, R., 1988. An evaluation of activity-based travel analysis. *Transportation (Amst.)* 15, 9–34.