JOURNAL OF INTERDISCIPLINARY METHODOLOGIES AND ISSUES IN SCIENCE

Identifying European Sustainable Mega-city Regions with Multi-dimensional Network Percolation

Juste RAIMBAULT[1,2,3,*]

[1]UPS CNRS 3611 ISC-PIF, France
[2]CASA, UCL, UK
[3]UMR CNRS 8504 Géographie-cités

*Corresponding author: juste.raimbault@polytechnique.edu

DOI: 10.18713/JIMIS-ddmmyy-v-a
Submitted: Month-in-letters Day Year - Published: Month-in-letters Day Year
Volume: **N** - Year: **YYYY**
Issue: **Title of the interdisciplinary issue**
Editors:

Abstract

The spatial distribution of human settlements into territories strongly conditions their sustainability. Recent forms of urbanization, in particular integrated mega-city regions, may imply different patterns of economic and transportation flows and thus exhibit various performances regarding different indicators of sustainability. This paper proposes to reconstruct endogenous urban regions from the bottom-up using network percolation, in a multi-dimensional way to take into account both urban form (spatial distribution of population) and urban function (properties of transportation networks). Variable parametrizations allows to consider patterns of optimization for two stylized contradictory sustainability indicators (economic performance and greenhouse gases emissions). This suggests a customizable spatial design of policies to develop sustainable territories.

Keywords

Road network; multilayer percolation; mega-city region

I INTRODUCTION

The structure of road networks both contains its past growth dynamics and has a significant impact on the sustainability of territories it irrigates. Diverse methods to characterize the structure of spatial networks, and more particularly road networks, have been developed in that context.

Lagesse *et al.* (2015)

A method to characterize topologies of these spatial networks is network percolation. Percolation in physics can be understood in a broad sense processes related to the progressive occupation or connection of nodes of a network, and is generally associated to a phase transition with the emergence of a giant cluster at a given connection probability (Stauffer and Aharony, 2014). Important applications include the quantification of network robustness (Callaway *et al.*, 2000).

Such approaches have been applied to urban systems. Makse *et al.* (1998) model urban growth with a local percolation model. It is also a method applied to the analysis of street networks, for example to extract endogenous urban regions (Arcaute *et al.*, 2016). In spatial statistics, this method can be used to characterize the spatial morphology of point patterns Huynh *et al.* (2018).

Piovani *et al.* (2017)

Existing heuristics however generally focus on a single morphological dimension of networks, and leave out the functional properties of urban systems (Burger and Meijers, 2012). However, urban systems are known to be multidimensional.

This paper addresses such a gap by introducing a multi-dimensional percolation heuristic, in order to combine different dimensions of the urban system, the same way that Cottineau *et al.* (2018) combines population density and commuting flows to produce multiple definitions of urban areas.

The rationale behind combining urban form with network topology measures relies on the capture of the link between urban form and function, assumed as distributed by transportation networks (Raimbault, 2018b).

Our contribution relies on several points: (i) this is to the best of our knowledge the first time a multi-dimensional percolation method is appplied to urban systems; (ii) we furthermore apply it on a significant spatial extent; and (iii) we link the cluster obtained with simple sustainability measures.

The rest of this paper is organized as follows: we first describe the multi-dimensional percolation heuristic, the data and variables to which it is applied, and the indicators used to characterized the sustainability of clusters produced. We then describe the results of applying this method to population and network variables for the whole European Union, focusing on the endogenous regions produced and their sustainability properties. We finally discuss possible developments and the implications of this methodology to the design of policies.

II METHODS

2.1 Multi-dimensional percolation

Given discrete spatial fields, site percolation is operated between two cells given a threshold parameter for each dimension and a distance threshold. This heuristic is similar to multilayer network percolation (Boccaletti *et al.*, 2014).

Son *et al.* (2012)

Hackett *et al.* (2016) bond percolation multiplex networks

The method works with an arbitrary number of layers.

The rationale

The parameters implied in this heuristic are the percolation radius r_0 and the percolation thresholds θ_i for each layer i.

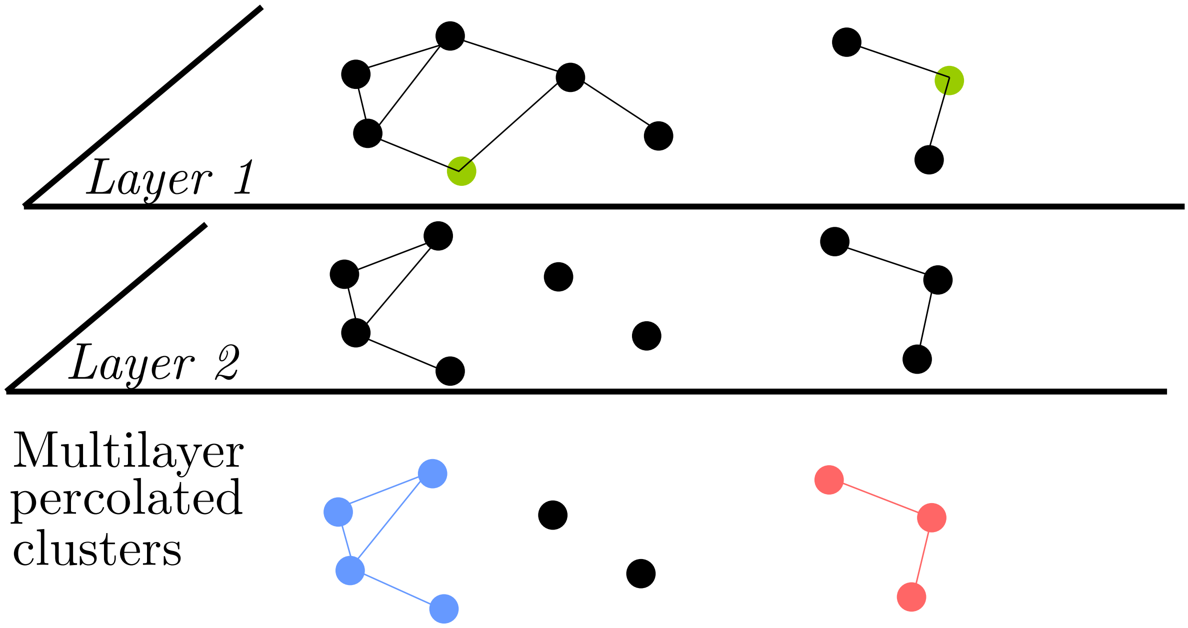

Figure 1: Schematic representation of the multi-dimensional network percolation heuristic.

2.2 Empirical data

We apply the heuristic to urban morphology and road network topology measures in Europe. More precisely, a grid with resolution 50km of population density morphology indicators and road network topology indicators, has been computed on spatial moving windows for all European Union by Raimbault (2019).

The percolation on such an abstract network is a necessary condition in our case to link the different dimensions considered, namely population distribution and local road network properties.

We percolate the population density layer with a network characteristic layer, that we test among number of edges, number of vertices, cyclomatic number and euclidian efficiency, which capture functional properties especially for the two last.

2.3 Sustainability indicators

We use this endogenous definition of regional urban systems produced by the percolation algorithm to evaluate their sustainability, in terms of conflicting objectives of economic integration and greenhouse gases emissions. The definition of sustainability, or sustainable development, is by essence multi-dimensional (Viguié and Hallegatte, 2012). Its characterization as quantitative indicators is even more subject to numerous degrees of freedom. We work here with two stylized indicators for two conflicting dimensions, as a proof-of-concept.

The EDGAR database (Janssens-Maenhout *et al.*, 2017) (version 4.3.2) is used for local estimates of greenhouse gases emissions.

Applying a gravity model to each region, we estimate abstract transportation flows within each and extrapolate emissions by coupling with the Edgar emission database (Janssens-Maenhout *et al.*, 2017) and economic activities with a scaling law of population. More precisely, greenhouse gases emissions derived from economic and transportation flows are estimated with the following expression

$$\phi_{ij} = \left(\frac{v_i v_j}{(\sum_k v_k)^2}\right)^\gamma \cdot \exp\left(\frac{-d_{ij}}{d_0}\right) \tag{1}$$

where v_k are either effective local GHG emissions or population. Indeed, the economic activity follows relatively well scaling laws of populations Bettencourt *et al.* (2007), the exponent being dependant on the activity and the definition of areas on which it is estimated (Cottineau *et al.*, 2017).

The sum of all flows within the geographical span of the cluster (that we approximate as the convex Hull envelope of its points), allows us to approximate the cumulated potential emissions and economic activity.

Using these potential flows follows the logic of Arbabi *et al.* (2019) which shows a need for improved intra-city-region mobility in England and Wales. Considering the regions as entities in which such transportation development policies can more easily been developed, we look at the sustainability of different possible regions if these potential flows were realized. Varying the parameters γ and d_0 allows to control for the economic activity considered (high γ values correspond to high added-value activities) and the span of interactions through d_0.

III RESULTS

3.1 Implementation

Banos and Genre-Grandpierre (2012) network efficiency

In practice, the analysis is implemented using R and the igraph package. Source code, data and results are available on the open git repository of the project at `https://github.com/JusteRaimbault/UrbanMorphology`.

The network is constructed by superposing the population density layer with the network layer, starting from the 5km resolution spatial fields for morphological and network indicators. This network is filtered with the threshold parameters for each layer and with the radius parameter. Connected components yield the clusters that we interpret as endogenous regions.

The experience plan is a full grid, for parameters r_0, θ_P, θ_N and the network variable considered. We also make γ and d_0 vary.

We recall that the euclidian performance of the network is in our case $< d_e/d_n >$ where the average is taken on all origin-destination pairs in the network, d_e is the euclidian distance and d_n the network distance. Thus, it indeed increases with network performance, in consistence with the use done here through thresholding.

We systematically explore the clusters obtained for 4800 parameter configurations.

3.2 Extracting endogenous mega-city regions

We obtain different endogenous morphologies.

Maps reveal that some configurations resemble the actual distribution of European mega-city regions, which are functionally integrated polycentric urban areas (Hall and Pain, 2006). These are here defined endogenously from the bottom-up and have a priori no reason to coincide with these functional regions.

We show some examples in Fig. 2.

3.3 Percolation transition and fractal dimension

In its application to road networks by Arcaute *et al.* (2016), the structure of the national urban system for UK is captured by studying the percolation transition, i.e. the variation of the size of the largest cluster as a function of the percolation radius. As this signature is tightly linked to historical, cultural and geographical conditions, the application to different urban systems should yield different results. We study here this property, for different threshold parameter values. The relative size of the largest cluster is plotted in Fig. **??** as a function of the percolation radius.

This aspect furthermore gives methodological information on multilayer percolation. Indeed, comparing the result with $\theta_N = 0$ (single layer percolation) with $\theta_N = 0.8$

3.4 Pareto fronts for sustainability

We show therein that different population, network and distance thresholds yield different performances in terms of sustainability.

3.5 Linking urban morphology and sustainability

When considering the aggregated indicators for a parametrization of endogenous city regions, one can relate them to morphological indicators for population density spatial distribution, computed by Raimbault (2018a), that we average on areas. This establishes a link between urban morphology and sustainibility. A principal component analysis on considered points yield 96% of variance with two components, and 73% explained by the first component alone. The first component relates to a level of monocentricity ($PC1 = -0.3 \cdot I + 0.54 \cdot \bar{d} + 0.51 \cdot \varepsilon + 0.59 \cdot \alpha$).

As shown in Fig. 5, there seems to exist an optimal intermediate level of monocentricity regarding the normalized indicator for emissions only, except for long-range and low-hierarchy interactions.

However, when considering both emissions and economic indicators, urban form then acts as a compromise variable, moving points on Pareto fronts, as shown in Fig. 6. In some case, highly monocentric areas can be a good compromise, whereas the intermediate optimal for emissions may yield highly inefficient areas. This unveils morphological trad-offs, confirming that there is no optimal urban form, but different compromises regarding the conflicting indicators.

IV DISCUSSION

4.1 Developments

Further work can consist in the use of calibration heuristics to find in a more robust way optimal parameter values. The OpenMOLE model exploration platform provides a transparent access to genetic algorithms for multi-objective optimization (Reuillon *et al.*, 2013). The use of such calibration algorithms would allow to unveil the effective form of Pareto fronts, that we may have missed here through the grid sampling.

Extrapolating transportation flows with a spatially explicit gravity and flow model can allow to compare these with actual transportation flows in the emission database, and yield a possible calibration. These extrapolated parameters could then be used within the economic and emissions potentials.

An other potential development would imply crossing our endogenous definitions of urban regions with socio-economic databases, and compute indicators implied in other dimensions of sustainability, for example related to socio-economic inequalities, spatial distribution of accessibilities, activities with different scaling exponents.

4.2 Towards policy applications

This suggests policies in terms of regional integration to increase the sustainability of mega-city regions.

The way such results can be transferred to policy-making recommandations remains an open question, but Pareto-optimal configurations can be used for the planning of regional transportation networks for example, or to design policies for the distribution of subsidies.

V CONCLUSION

In conclusion, our multilayer percolation approache captures in a way the multi-dimensionality of urban systems and a link between form and function.

References

Arbabi H., Mayfield M., McCann P. (2019). On the development logic of city-regions: inter-versus intra-city mobility in england and wales. *Spatial Economic Analysis*, 1–20.

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.

Banos A., Genre-Grandpierre C. (2012). Towards new metrics for urban road networks: Some preliminary evidence from agent-based simulations. In *Agent-based models of geographical systems*, pp. 627–641. Springer.

Bettencourt L. M., Lobo J., Helbing D., Kühnert C., West G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences 104*(17), 7301–7306.

Boccaletti S., Bianconi G., Criado R., Del Genio C. I., Gómez-Gardenes J., Romance M., Sendina-Nadal I., Wang Z., Zanin M. (2014). The structure and dynamics of multilayer networks. *Physics Reports 544*(1), 1–122.

Burger M., Meijers E. (2012). Form follows function? linking morphological and functional polycentricity. *Urban studies 49*(5), 1127–1149.

Callaway D. S., Newman M. E., Strogatz S. H., Watts D. J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical review letters 85*(25), 5468.

Cottineau C., Finance O., Hatna E., Arcaute E., Batty M. (2018). Defining urban clusters to detect agglomeration economies. *Environment and Planning B: Urban Analytics and City Science*, 2399808318755146.

Cottineau C., Hatna E., Arcaute E., Batty M. (2017). Diverse cities or the systematic paradox of urban scaling laws. *Computers, environment and urban systems 63*, 80–94.

Hackett A., Cellai D., Gómez S., Arenas A., Gleeson J. P. (2016). Bond percolation on multiplex networks. *Physical Review X 6*(2), 021002.

Hall P. G., Pain K. (2006). *The polycentric metropolis: learning from mega-city regions in Europe*. Routledge.

Huynh H. N., Makarov E., Legara E. F., Monterola C., Chew L. Y. (2018). Characterisation and comparison of spatial patterns in urban systems: A case study of us cities. *Journal of computational science 24*, 34–43.

Janssens-Maenhout G., Crippa M., Guizzardi D., Muntean M., Schaaf E., Dentener F., Bergamaschi P., Pagliari V., Olivier J., Peters J., *et al.* (2017). Edgar v4. 3.2 global atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth Syst. Sci. Data Discuss.*

Lagesse C., Bordin P., Douady S. (2015). A spatial multi-scale object to analyze road networks. *Network Science 3*(1), 156–181.

Makse H. A., Andrade J. S., Batty M., Havlin S., Stanley H. E., *et al.* (1998). Modeling urban growth patterns with correlated percolation. *Physical Review E 58*(6), 7054.

Piovani D., Molinero C., Wilson A. (2017). Urban retail location: insights from percolation theory and spatial interaction modeling. *PloS one 12*(10), e0185787.

Raimbault J. (2018a). Calibration of a density-based model of urban morphogenesis. *PloS one 13*(9), e0203516.

Raimbault J. (2018b). *Caractérisation et modélisation de la co-évolution des réseaux de transport et des territoires*. Ph. D. thesis, Université Paris 7 Denis Diderot.

Raimbault J. (2019). An urban morphogenesis model capturing interactions between networks and territories. In *The Mathematics of Urban Morphology, D'acci L., ed., Springer Birkhauser Mathematics*.

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.

Son S.-W., Bizhani G., Christensen C., Grassberger P., Paczuski M. (2012). Percolation theory on interdependent networks based on epidemic spreading. *EPL (Europhysics Letters) 97*(1), 16006.

Stauffer D., Aharony A. (2014). *Introduction to percolation theory*. Taylor & Francis.

Viguié V., Hallegatte S. (2012). Trade-offs and synergies in urban climate policies. *Nature Climate Change 2*(5), 334.

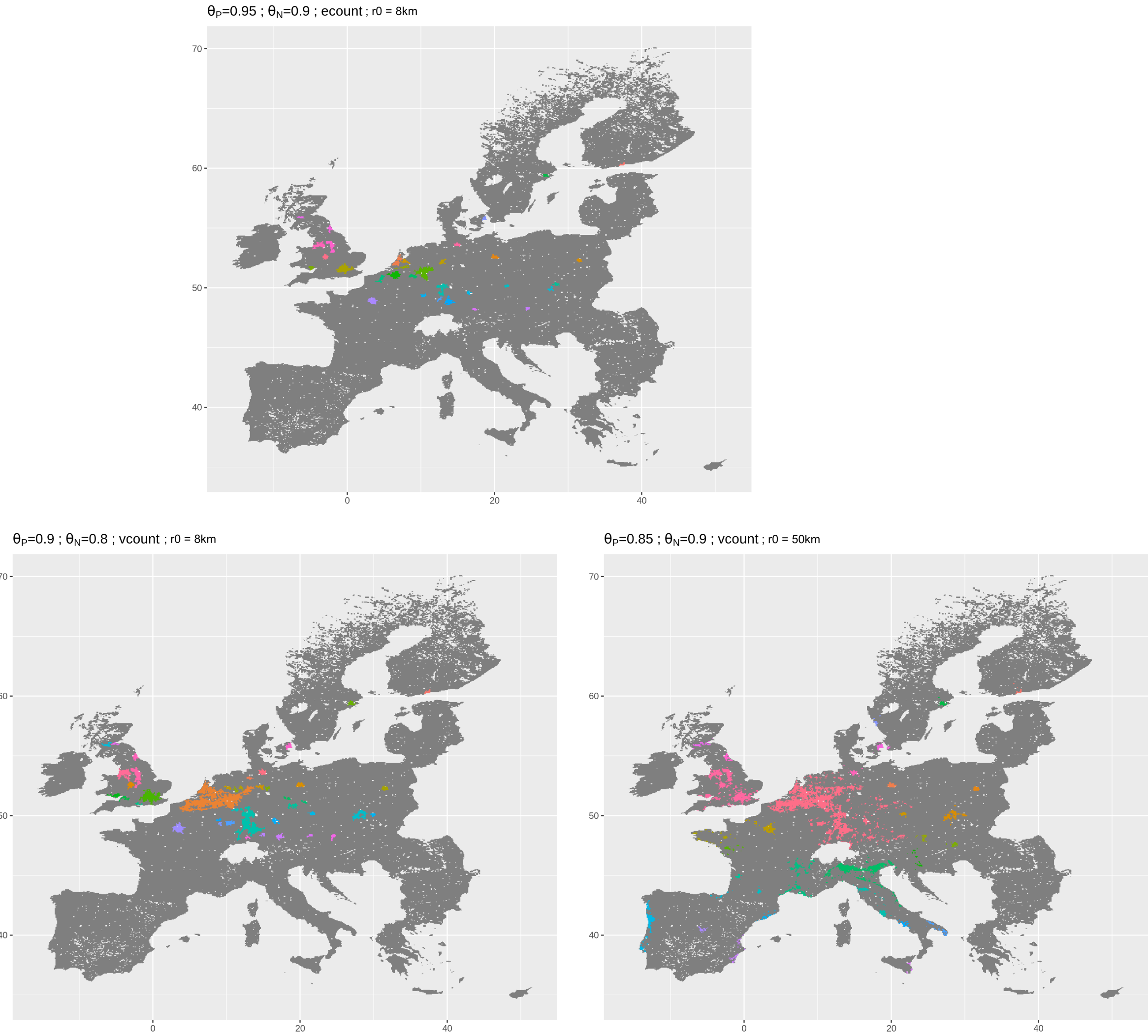

Figure 2: Examples of obtained clusters for different parameter values. In the third case for example, we obtain the urban regions of West midlands and London in the UK, Randstad merged with Rhein-Rhur and Rhein-Main in Germany, Paris in France, also with capital cities such as Copenhaguen, Stockholm and Helsinki. There is no cluster in South Europe in that case, due to the high population density threshold.

Figure 3: Percolation transition.

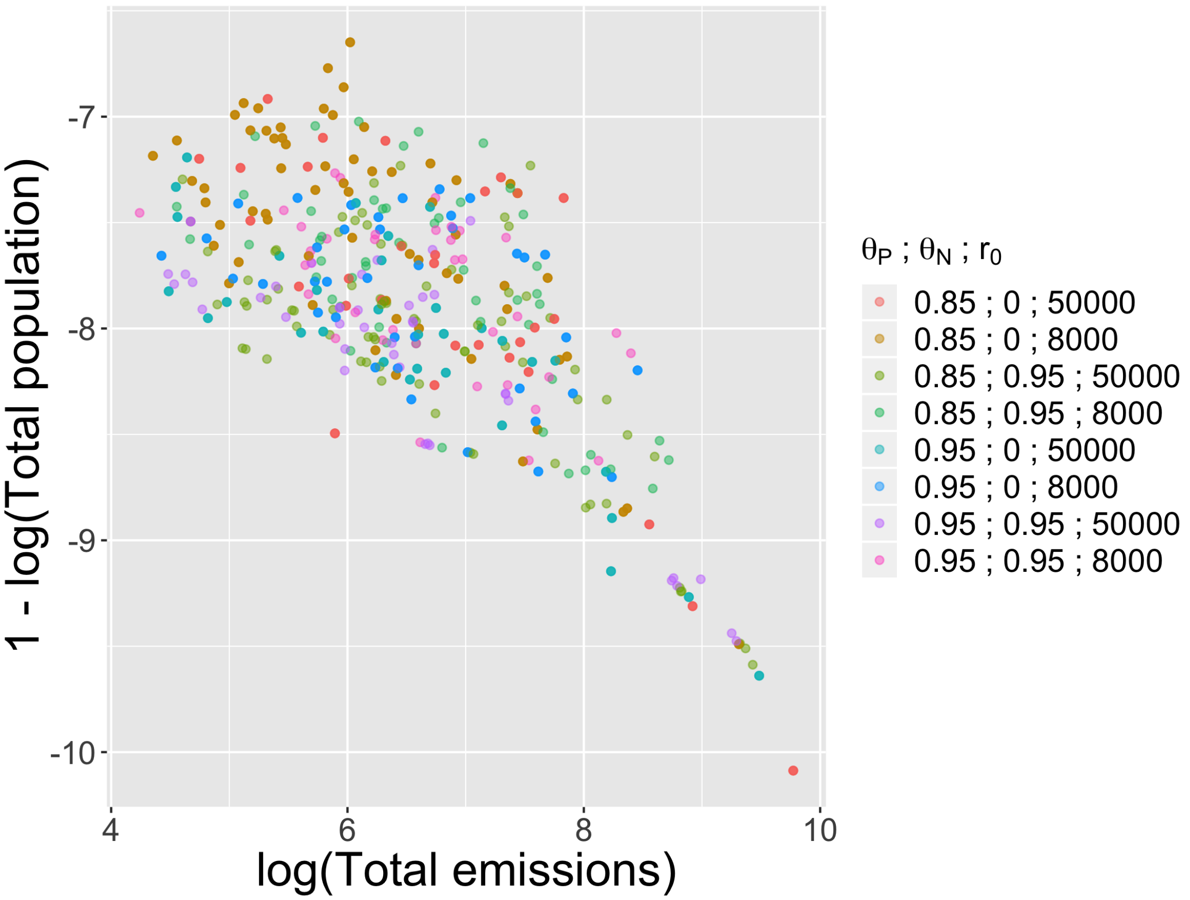

Figure 4: Point clouds of region-level indicators, namely population and emissions, for different parametrizations.

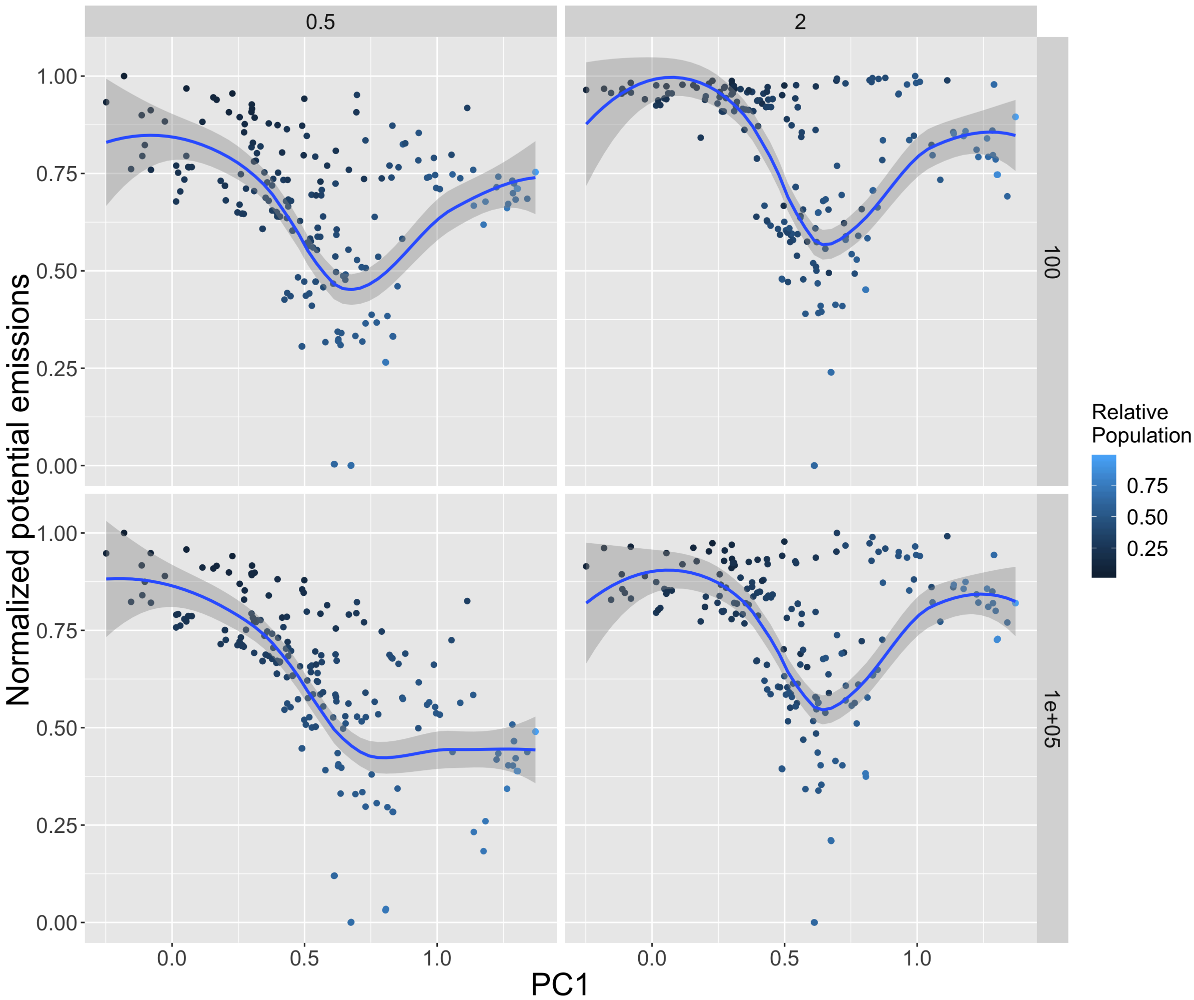

Figure 5: Aggregated values of normalized potential emissions, as a function of the first morphological principal component (PC1), for varying values of parameters d_G (rows) and γ_G (columns). Other intermediate values for these parameters yield similar behaviors. As PC1 is mainly linked to monocentricity, there seems to exist an optimal intermediate level of monocentricity for emissions alone.

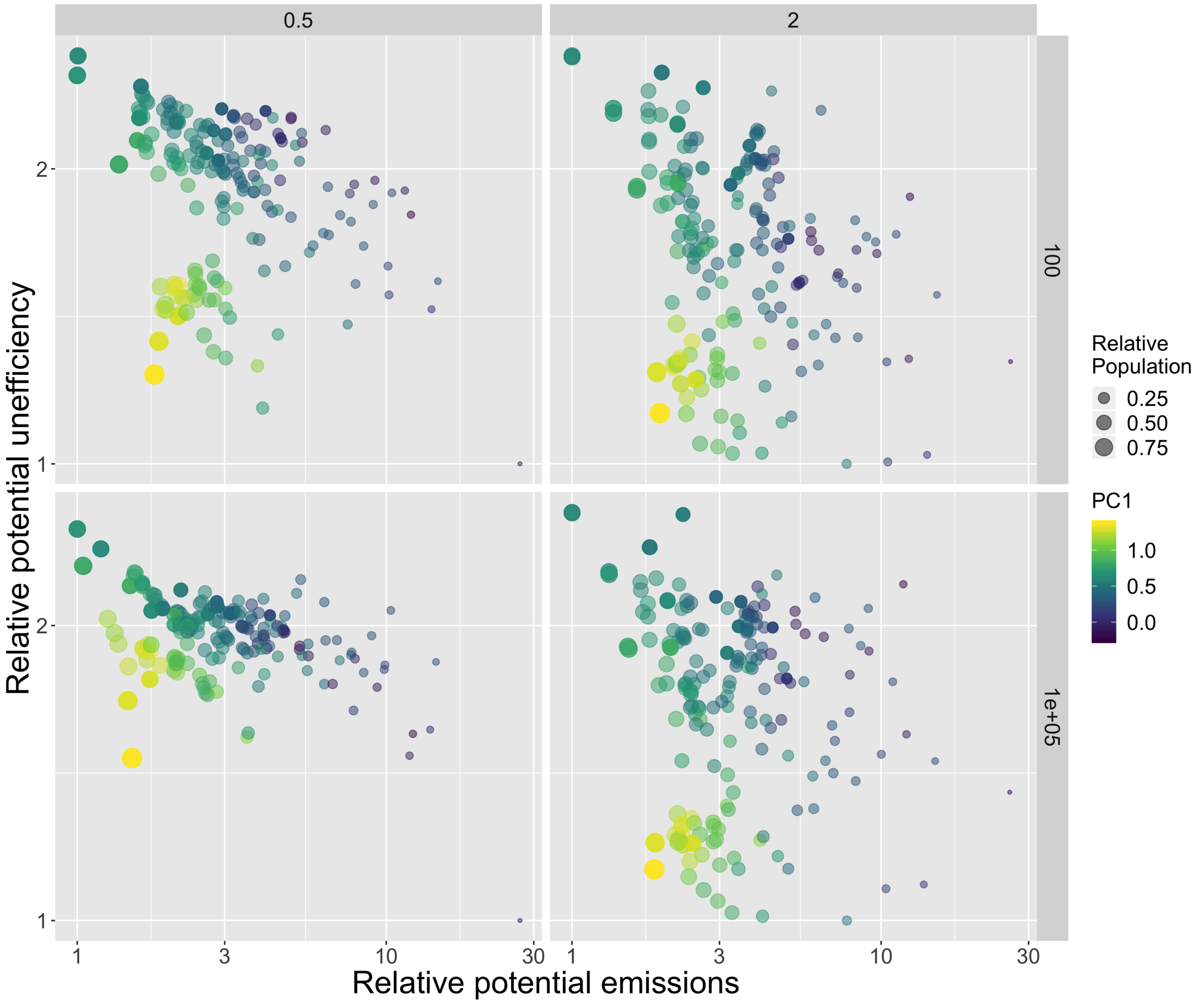

Figure 6: Relative potential emissions against relative potential economic unefficiency (both indicators should be minimized), for varying values of γ_G (columns) and d_G (rows). Color level gives the value of PC1, whereas point size gives the share of total population contained within considered areas.