International Journal of
Geo-Information

Editorial

Spatial Heterogeneity, Scale, Data Character and Sustainable Transport in the Big Data Era

Bin Jiang

Faculty of Engineering and Sustainable Development, Division of GIScience, University of Gävle, SE-801 76 Gävle, Sweden; bin.jiang@hig.se

Received: 25 April 2018; Accepted: 28 April 2018; Published: 28 April 2018

In light of the emergence of big data, I have advocated and argued for a paradigm shift from Tobler's law to scaling law; from Euclidean geometry to fractal geometry; from Gaussian statistics to Paretian statistics; and more importantly, from Descartes' mechanistic thinking to Alexander's organic thinking. Fractal geometry falls under the third definition of fractal. Essentially, a set or pattern is fractal if the scaling of far more small things than large ones recurs multiple times [1]. Fractal geometry does not fall under the second definition of fractal, which requires a power law between the scales and details [2]. The new fractal geometry leans more towards living geometry that *"follows the rules, constraints, and contingent conditions that are, inevitably, encountered in the real world"* ([3], p. 395). This is important for understanding complexity and creating complex or living structures [4]. This editorial attempts to clarify why this paradigm shift is essential and to elaborate on several concepts, including spatial heterogeneity (scaling law), scale (or the fourth meaning of scale), data character (in contrast to data quality) and sustainable transport in the big data era.

The current geographic information systems (GIS), which were first conceived and developed in the 1970s, are still largely based on the legacy of conventional cartography [5]. Although computer technology has dramatically advanced since then, the legacy or the fundamental ways of thinking remains unchanged. For example, the GIS representations of raster and vector and even the so-called object-oriented representation are still constrained to the geometric primitives, such as pixels, points lines, and polygons [6]. These geometry-oriented representations help us to see things that are more or less similar, which is characterized by Tobler's law [7]. This is commonly known as the first law of geography. For example, the price of your house may be similar to those of your neighbors, but there are far more low house prices than high ones. This notion of far more lows than highs or far more smalls than larges in general is what underlies the scaling law for characterizing spatial heterogeneity. The concept of spatial heterogeneity, as conceived in current geography literature, is mistaken because it does not recognize the fact of far more smalls than larges. This notion of far more smalls than larges adds a fourth meaning of scale: a series of scales ranging from the smallest to the largest that forms the scaling hierarchy [8]. The scaling hierarchy can be further rephrased as: numerous smallest, very few largest and some in between the smallest and the largest. In order to see far more smalls than large ones, we must consider a topological perspective on the meaningful geographic features, such as streets and cities, instead of the geometric primitives. In other words, we must shift our eyes from the geometric details to the overall character of the data (which will be further elaborated on below). Tobler's law depicts a fact as a local scale. However, the geographic space is governed by not only Tobler's law but also by the scaling law. These two laws are complementary to each other (Table 1). Calling for a shift from Tobler's law to scaling law does not involve abandoning Tobler's law, but instead calls for a shift from thinking that is dominated by Tobler's law to thinking that is dominated by the scaling law as the scaling law is universal and global.

Table 1. Comparison of scaling law versus Tobler's law.

Scaling Law	Tobler's Law
far more small things than large ones	more or less similar things
across all scales	available on one scale
without an average scale (Pareto distribution)	with an average scale (Gauss distribution)
long tailed	short tailed
interdependence or spatial heterogeneity	spatial dependence or homogeneity
disproportion (80/20)	proportion (50/50)
complexity	simplicity
non-equilibrium	equilibrium

Note: These two laws complement each other and recur at different levels of scale in the geographic space or the Earth's surface.

Benoit Mandelbrot, the father of fractal geometry, remarked that unlike the things we see in nature, Euclidean shapes are cold and dry. Christopher Alexander, the father of living geometry, referred to the structures with a higher degree of wholeness as living structures. Wholeness is defined mathematically as a recursive structure, which physically exists in space and matter and psychologically reflects in our minds and cognition [4,9]. A comparison between a cold and dry structure and a living one vividly describes the difference between Euclidean shapes and fractal or living structures. A shift from Euclidean geometry to fractal or living geometry does not involve abandoning Euclidean geometry as it is instead requires the abandonment of Euclidean geometric thinking. Euclidean geometry is essential for fractal geometry, since one must first measure it in order to see whether there are far more smalls than larges. However, Euclidean geometric thinking differs fundamentally from that of fractal geometry. For example, Euclidean geometric thinking tends to see things individually (rather than holistically) and non-recursively (rather than recursively). The work by Jiang and Brandt [8] is a good reference for a more detailed comparison about these two geometric ways of thinking. The shift from Euclidean to fractal or living geometry implies that the fractal or living geometry needs to be the dominant way of thinking. To present a practical example, a cartographic curve is commonly seen as a set of more or less similar line segments under the Euclidean geometric thinking. However, it should more correctly be viewed, under the fractal or living geometric thinking, as a set of far more small bends than large ones. More importantly, one must recognize that these small bends are recursively embedded in the large ones.

According to Tobler's law, the price of your house is similar to those of your neighbors. In other words, averaging your neighbors' housing prices would obtain your housing price. In this case, the average is suitable for predicting your housing price as all neighboring housing prices can be characterized by a well-defined mean. In order for the prediction to make sense, there is a condition to meet. Namely, the neighboring housing prices are more or less or with a well-defined mean. This is indeed true for the housing prices on a local scale. This condition is violated on a global scale, since there are far more low prices than high ones. In this case, using the average does not make sense as it lacks an average, scaling or being scale-free. Things with more or less similar sizes can be well modeled by Gaussian statistics, while things with far more smalls than larges should be well characterized by Paretian statistics. In this regard, we have developed a new classification scheme for the data with far more smalls than larges. This classification scheme is called the head/tail breaks [10], which recursively breaks the data into the head (for data values greater than an average) and the tail (for data values less than an average) until the condition of a small head and a long tail is violated. A head/tail breaks-induced index, which is called the ht-index [1], can be used to characterize the notion of far more smalls than larges or the underlying scaling hierarchy. As mentioned above, the new definition of fractal is based on the notion of far more smalls than larges.

Under the Euclidean geometric thinking that focuses on geometric details, the data quality or uncertainty is one of the priority issues in GIS. This type of thinking is evident in many scientific papers and talks about, for example, the data quality of OpenStreetMap (OSM). I believe that the

GIS field has over-emphasized the data quality issue, so I would like to add a different view, arguing that the data quality is less important than the data character. By data character, I mean the overall character of geospatial data, the wholeness or living structure as briefly mentioned above. To illustrate, this link (https://twitter.com/binjiangxp/status/985322539625967618) shows a cartoon and a photo of Kim Jong-Un. The photo on the right has the highest geometric details, while the cartoon on the left has the lowest geometric details. However, the cartoon on the left captures the highest character or personality. This link (https://twitter.com/binjiangxp/status/985322961342263296) further illustrates the living structure of the street network of a small neighborhood. The street network on the right has the highest data quality, while the graph on the left captures the highest data character: far more less-connected streets than well-connected ones. I want to argue that if the street network on the right suffers from some errors, this would not have a significant effect on the data character on the left. In this sense, the quality is not super-important compared to data character. If data quality or geometric details were compared to trees, the data character would be the forest. To present a specific example, Jiang and Liu [11] adopted a topological perspective and examined the scaling of the geographic space based on OSM data of three European countries: France, Germany and the UK. There is little doubt that there were numerous errors in the OSM data, but they have little effect on the finding, which was the scaling or living structure of geographic space in which there are far more small things than large ones.

The legacy of GIS has been driven substantially by the mechanistic thinking of the past 300 years of science since the 17 century by Descartes and others [12]. Everything we have achieved in science and technology benefits greatly from mechanistic thinking, but it is now limited in terms of how to make a better built environment [4]. This mechanistic thinking is reflected in the GIS representations of raster and vector as well as in the box counting for calculating the fractal dimensions. It is also reflected in the top-down imposed geographic units, such as census tracts. These are clearly very useful for administration and management but are of little use for scientific purposes. Space is neither lifeless nor neutral but has a capacity to be more living or less living [4]. In other words, space is a living structure with a high degree of wholeness. A country is a living structure that consists of far more small cities than large ones. Figure 1 shows all the natural cities of Austria, which was derived from the street nodes of the country's OSM data. As seen from the figure, all cities have very natural boundaries and they are quite coherent, with a topographic surface that reveals the underlying living structure of far more small cities than large ones. The natural cities are objectively defined cities from a massive number of geographic locations, such as social media locations [13]: please refer to the Appendix of the paper for details on the derivation of natural cities.

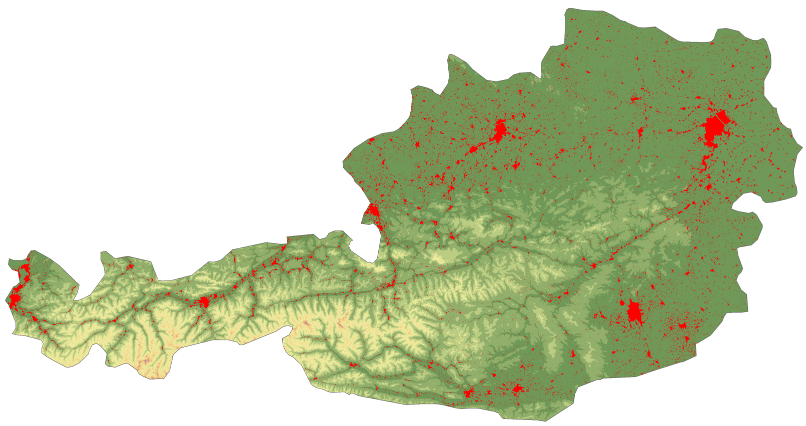

Figure 1. (Color online) Living structure of all natural cities derived from Austria's OSM data. Note: This topographic rendering is based on head/tail breaks [10,14]. Unlike the traditional renderings, this head/tail breaks-induced rendering clearly shows the living structure of far more low elevations than high ones.

The geographic space is a living structure at both the country and city scale. Figure 2 shows the natural streets of Vienna and Linz, which demonstrate striking living structures with far more less-connected streets than well-connected ones. The natural streets are able to capture the underlying scaling or living structure of the far more less-connected than the well-connected streets and therefore, are able to predict up to 80% of the traffic flow. In other words, the traffic flow is mainly shaped by the living structure and has little to do with human travel behavior. In this circumstance, human beings can be thought of as the atoms or molecules that interact with each other and with the natural streets to shape the traffic flow. Traffic is not a phenomenon, but an outcome of the living structure. This is in line with the famous statement by Winston Churchill: *"We shape our buildings and afterwards, our buildings shape us"*. With respect to sustainable transport, we can paraphrase Churchill: We shape our transport system and it will shape us, so make sure we shape it well so that we will be well-shaped too. To be more specific, we shape our transport system as a living structure and after this, it shapes our sustainable mobility.

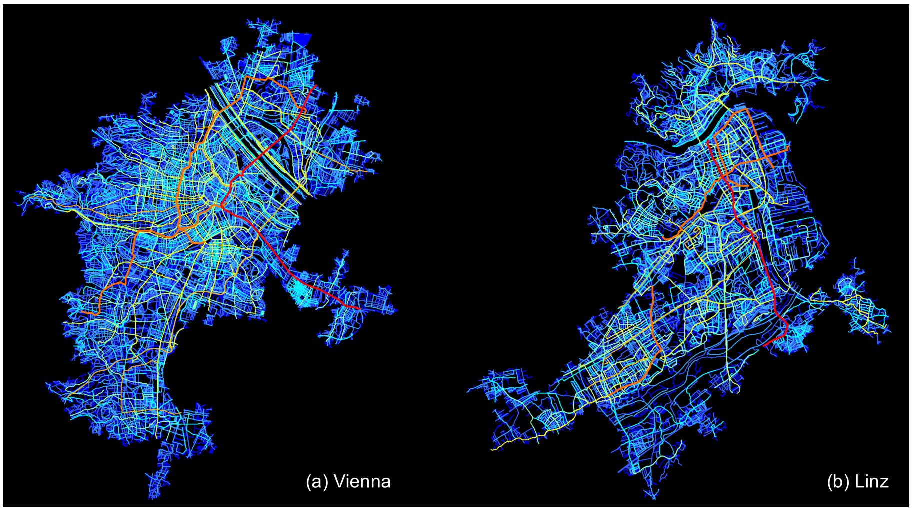

Figure 2. (**Color online**) Living structure of natural streets of (**a**) Vienna and (**b**) Linz. Note: The living structure shows far more less-connected streets (indicated by many cold colors) than well-connected ones (indicated by a very few warm colors).

This social physics perspective offers new insights into traffic flow. To this point, I would like to end this editorial with the following excerpt [15]:

"There's an old way of thinking that says the social world is complicated because people are complicated . . . We should think of people as if they were atoms or molecules following fairly simple rules and try to learn the patterns to which those rules lead . . . Seemingly complicated social happenings may often have quite simple origins . . . It's often not the parts but the pattern that is most important, and so it is with people."

Acknowledgments: This editorial was substantially inspired by my recent panel presentation *"On Spatiotemporal Thinking: Spatial heterogeneity, scale, and data character"*, presented at the panel session entitled "Spatiotemporal Study: Achievements, Gaps, and Future" with the AAG 2018 annual meeting, New Orleans, 10–15 April 2018, and my keynote *"A Geospatial Perspective on Sustainable Urban Mobility in the Era of BIG Data"*, to be presented at CSUM 2018: Conference on Sustainable Urban Mobility, 24–25 May, Skiathos Island, Greece.

References

1. Jiang, B.; Yin, J. Ht-index for quantifying the fractal or scaling structure of geographic features. *Ann. Assoc. Am. Geogr.* **2014**, *104*, 530–541. [CrossRef]
2. Mandelbrot, B.B. *The Fractal Geometry of Nature*; W.H. Freeman and Co.: New York, NY, USA, 1982.

3. Alexander, C.; Neis, H.; Alexander, M.M. *The Battle for the Life and Beauty of the Earth*; Oxford University Press: Oxford, UK, 2012.
4. Alexander, C. *The Nature of Order: An Essay on the Art of Building and the Nature of the Universe*; Center for Environmental Structure: Berkeley, CA, USA, 2002–2005.
5. Goodchild, M.F. Reimagining the history of GIS. *Ann. GIS* **2018**. [CrossRef]
6. Longley, P.A.; Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. *Geographic Information Science and Systems*; Wiley: Chichester, UK, 2015.
7. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [CrossRef]
8. Jiang, B.; Brandt, A. A fractal perspective on scale in geography. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 95. [CrossRef]
9. Jiang, B. A complex-network perspective on Alexander's wholeness. *Physica A* **2016**, *463*, 475–484. [CrossRef]
10. Jiang, B. Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *Prof. Geogr.* **2013**, *65*, 482–494. [CrossRef]
11. Jiang, B.; Liu, X. Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 215–229, Reprinted in *Big Data Comput* [CrossRef]
12. Descartes, R. *The Geometry of Rene Descartes*; Smith, D.E., Latham, M.L., Eds.; Dover Publications: New York, NY, USA, 1954.
13. Jiang, B.; Miao, Y. The evolution of natural cities from the perspective of location-based social media. *Prof. Geogr.* **2015**, *67*, 295–306, Reprinted in *ICT Social Networks and Travel Behaviour in Urban Environments*. [CrossRef]
14. Jiang, B. Head/tail breaks for visualization of city structure and dynamics. *Cities* **2015**, *43*, 69–77. [CrossRef]
15. Buchanan, M. *The Social Atoms: Why the Rich Get Richer, Cheaters Get Caught, and Your Neighbor Usually Looks Like You*; Bloomsbury: New York, NY, USA, 2007.