Process indicator-of-scale metrics – PRISMs

Taylor M. Oshan, Center for GISc./Geographical Sciences, UMD - College Park
Levi J. Wolf, Geographical Sciences, University of Bristol

Scale is a central concept in geography and the spatial sciences. Though it serves as an essential thread in the intricate tapestry of the physical and social sciences, the number and nature of the available definitions and deployments is as varied as the diverse disciplines that co-opt the concept. Nevertheless, a few consistent definitions have emerged. These are first highlighted in order to draw attention to an under studied type of spatial scale -- that of process scale. The focus is then shifted towards differentiating process scale from other types of scale in order to highlight trends and suggest a research agenda to develop metrics of process scale.

There is by now a large volume of literature describing the concept(s) of spatial scale and detailing its development across various disciplines. On the one hand, domain specific scale terminologies have arisen in conjunction with substantive research in areas such as landscape ecology (e.g., Turner, 1989), remote-sensing (e.g., Wu & Li, 2009), and segregation (e.g., Johnston et al., 2019). This has perhaps led to the idea that scale is an ambiguous concept and is dependent upon application context (Goodchild, 2001; Turner, 1989; Dabiri & Blaschke, 2019). However, on the other hand, there now exists several general treatments of the topic that attempt to bridge subjects and forge a common understanding (Harvey, 1968; Lam & Quattrochi, 1992; Marceau, 1999; Atkinson & Tate, 2000; Dabiri & Blaschke, 2019). An outcome of these efforts is a few definitions of spatial scale that are frequently encountered across the spatial sciences.

Dabiri and Blaschke (2019) provide a high level synthesis of some previous conceptualizations of spatial scale that result in three consistent definitions (amongst others) that are often described in the literature. First, geographic scale refers to the extent of an area of interest. Second, measurement or observation scale refers to the resolution of spatial units across an area of interest. Finally, the third definition refers to operational or process scale and is the dimension over which particular processes occur. It is this latter definition of scale that is of primary interest here that will be further differentiated and defined.

A more general definition of scale is provided by Marceau who states that, "scale refers to the spatial dimensions at which entities, patterns, and processes can be observed and characterized" (1999). Following this definition, another way to think about the three above mentioned conceptualizations of spatial scale is in terms of the way in which they modulate spatial entities, patterns, and processes. Commonly, in the spatial sciences the focus is on collecting and analyzing georeferenced data in order to measure patterns and ultimately learn about spatial processes. In this context, geographic scale and measurement scale are the dimensions that modulate spatial patterns. Geographic scale can be thought of as the macro attribute governing spatial patterns whereas measurement scale can be thought of as the micro attribute governing spatial patterns (Goodchild, 2001). The former controls the amount of area over which a pattern can vary where the latter controls the number and nature of potential spatial units over which a pattern can vary. In contrast, operational or process scale is the dimension that modulates the relationships that generate the data and patterns we observe.

Spatial patterns and processes are inherently linked since a single process can generate a variety of patterns and the same pattern can be generated by different processes. In addition, geographic scale and measurement scale intrinsically limit our ability to quantify spatial processes. That is, “scales of variation observable in spatial data are inextricably linked to the scales of measurement through which they were obtained” (Atkinson & Tate, 2000), which similarly limits our ability to infer spatially variable processes. However, a major difference between pattern and process is that patterns are typically concerned with a single variable (i.e., univariate) while process implies relationships and associations between two or more variables (i.e., multivariate). At the same time, it has long been recognized that different processes occur at different scales and that a single scale may not be sufficient to characterize complex multivariate processes (Harvey, 1968; Marceau, 1999). Despite this, a collective shift in focus towards multiscale patterns and processes is a relatively recent trend. It is also possible to differentiate between multiscale patterns, which can be captured by employing multiple geographic and/or measurement scales, and multiscale processes, which are typically captured via local multivariate statistical models.

Whereas scale is typically treated in the broader literature as a problem or issue to be dealt, local modeling treats scale as a feature that can be exploited. For example, there are many investigations of how to best fuse or change the scale of datasets, define an optimal scale associated with a pattern, or understand the effect of different measurement scales on analytical results (i.e., MAUP), which can each imply information about process scale. In contrast, local multivariate statistical models take a fixed geographical scale and observational scale in order to make explicit inferences about process scale and heterogeneity. In the local modeling paradigm, process is hypothesized to potentially vary across space in the same way that data observations may vary across space to form patterns. As a result, spatially varying processes may be characterized by a scale parameter that governs how constituent components interact, as well as patterns that describe the variation within the process. Two examples include geographically weighted regression (GWR) (Fotheringham et al., 2002) and Bayesian spatially varying coefficient models (Gelfand et al., 2003). Multiscale variants of local multivariate statistical models have recently been developed that produce relationship-specific estimates of scale parameters, allowing processes to spatially vary independent of each other (Fotheringham et al., 2017; Murakami et al.,, 2017; Wolf et al., 2018; Oshan et al., 2019b). As a result, multiscale local multivariate statistical models act as a prism for decomposing spatial variation amongst spatial data across a set of spatial processes.

Several research directions can therefore be suggested in order to develop these relationship-specific scale parameter estimates as more formal process indicator-of-scale metrics (PRISMs). First, it is important to determine the connection between the spatial variability of processes (i.e., coefficient estimate surfaces) and scale parameter estimates (i.e., GWR bandwidth). A related second suggestion is to investigate scale uncertainty and how to interpret it (e.g., Oshan et al., 2019a; Li et al, 2019; Oshan et al., 2019c). These two tasks would help establish a better understanding of how to categorize spatial process as global, regional, or local. Third, more work is needed to ground metrics of scale in terms of the meaning of extreme values (i.e., minimum and maximum). Finally, it is necessary to develop diagnostics and identify potential limitations (e.g., Yu et al., 2019a; Yu et al., 2019b) when quantifying process scale using a statistical framework; the usual suspects, such as omitted variable bias, functional form misspecification, and multicollinearity, may affect the quantification of process scale and heterogeneity. Such endeavors would provide the foundation for PRISMs, enhancing the geography of relationships and contributing greatly towards a science of spatial scale.

References

Atkinson, P. M., & Tate, N. J. (2000). Spatial Scale Problems and Geostatistical Solutions: A Review. *The Professional Geographer*, *52*(4), 607–623.

Dabiri, Z., & Blaschke, T. (2019). Scale matters: A survey of the concepts of scale used in spatial disciplines. *European Journal of Remote Sensing*, *52*(1), 419–434.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multi-Scale Geographically Weighted Regression. *Annals of the American Association of Geographers*, *107*(6), 1247–1265.

Gelfand, A. E., Kim, H.-J., Sirmans, C. F., & Banerjee, S. (2003). Spatial Modeling With Spatially Varying Coefficient Processes. *Journal of the American Statistical Association*, *98*(462), 387–396.

Goodchild, M. F. (2001). Metrics of scale in remote sensing and GIS. *International Journal of Applied Earth Observation and Geoinformation*, *3*(2), 114–120.

Harvey, D. W. (1968). Pattern, Process, and the Scale Problem in Geographical Research. *Transactions of the Institute of British Geographers*, (45), 71.

Lam, N. S.-N., & Quattrochi, D. A. (1992). On the Issues of Scale, Resolution, and Fractal Analysis in the Mapping Sciences*. *The Professional Geographer*, *44*(1), 88–98.

Li, Z., Fotheringham, A. S., Oshan, T. M., & Wolf, L. J. (2019). Measuring Bandwidth Uncertainty in Multiscale Geographically Weighted Regression Using Akaike Weights. *OSF Preprints*. https://doi.org/10.31219/osf.io/h2qmy

Marceau, D. J. (1999). The scale issue in social and natural sciences. *Canadian Journal of Remote Sens*, *25*(4), 18.

Murakami, D., Yoshida, T., Seya, H., Griffith, D. A., & Yamagata, Y. (2017). A Moran coefficient-based mixed effects approach to investigate spatially varying relationships. *Spatial Statistics*, *19*, 68–89.

Oshan, Taylor, Wolf, Levi John, Fotheringham, A. Stewart, Kang, Wei, Li, Ziqi & Yu, Hanchen (2019a). A comment on geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 33:7

Oshan, T., Li, Z., Kang, W., Wolf, L., & Fotheringham, A. (2019b). mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating

Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information*, *8*(6), 269.

Oshan, T. M., Smith, J., & Fotheringham, A. S. (2019c). A multiscale spatial analysis of obesity determinants in Phoenix, Arizona. *OSF Preprints.* https://doi.org/10.31219/osf.io/unfwj

Turner, M. G. (1989). Landscape Ecology: The Effect of Pattern on Process. *Annual Review of Ecology and Systematics*, 29.

Wolf, L. J., Oshan, T. M., & Fotheringham, A. S. (2017). Single and Multiscale Models of Process Spatial Heterogeneity: Single and Multiscale Models. *Geographical Analysis*.

Wu, H., & Li, Z.-L. (2009). Scale Issues in Remote Sensing: A Review on Analysis, Processing and Modeling. *Sensors*, *9*(3), 1768–1793.

YYu, H. , Fotheringham, A. S., Li, Z. , Oshan, T. , Kang, W. and Wolf, L. J. (2019a). Inference in multiscale geographically weighted regression. *Geographical Analysis*.

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T. M., & Wolf, L. J. (2019b). On the Measurement of Bias in Geographically Weighted Regression Models. *OSF Preprints*. https://doi.org/10.31219/osf.io/etb42