# CroScalar: A Muti-Scale Modeling Framework for Spatio-Temporal Data

**Yi Qiang[a]*, Barbara Buttenfield[b]**

[a] School of Geosciences, University of South Florida, Tampa, Florida
[b] Department of Geography, University of Colorado, Boulder, Colorado
* qiangy@usf.edu

**Keywords:** spatial scale**,** multi-scale analysis, MAUP, point pattern, spatial autocorrelation, regression

## 1 Introduction

Scale is one of the fundamental topics that differentiate geography from other disciplines. However, the handling of scale is a longstanding challenge in spatial analysis, as analyses conducted at different spatial scales may create different results. The importance of scale in spatial analysis has been epitomized in the well-known Modifiable Areal Unit Problem (MAUP) (Openshaw, 1983). Ideally, spatial data should be analyzed at the scale where spatial processes become evident, are best understood and/or where spatial relationships are maximized. However, most spatial analyses and modelling are conducted at a pragmatic scale or the scale where data is collected, which may miss critical spatial processes and relationships concealed at other scales. With the advent of the Big Data era, the increasing availability of geospatial data collected at different resolutions poses new challenges for multi-scale data fusion and analysis.

In GIS, space is conventionally represented in "flat layers" and most spatial analysis tools operate at a single scale. Multi-scale analysis tends to treat scale as a variable and analyze the variation of spatial metrics computed at different scales (Behrens et al., 2019; Lam et al., 2018). Some spatial analysis methods use an adaptive kernel bandwidth or search radius to detect spatial patterns and relations at multiple scales (Ester et al., 1996; Fotheringham et al., 2017; Kulldorff, 1997; Van Kerm, 2003). However, existing multi-scale analysis methods cannot fully reveal hierarchical structures and nested relations among spatial processes at diverging scales. The solution to this challenge requires a modelling framework that can seamlessly integrate space and scale to represent spatial patterns or relations varying in both location and scale.

To address the challenges of multi-scale analysis, Qiang et al. (2014) developed a Triangle Model (TM) and extended it into a Pyramid Model (PM) which integrates spatial location and spatial scale in a true 3D space (Qiang et al., 2018; Qiang & Van de Weghe, 2019)). The PM has been applied to multi-scale analysis of wetland fragmentation (Qiang & Van de Weghe, 2019), point pattern analysis (Qiang et al., 2022), and land cover classification (Carlson, 2021). These studies demonstrate that the PM seamlessly integrates space and scale in a 3D environment that supports simultaneous monitoring and assessment of spatial patterns across a range of scales. To date, the developed analytical tools primarily focus on exploratory analysis. In the 2022 AutoCarto conference, we will present experiments of implementing inferential methods in the PM to detect and quantify multi-scale spatial patterns and relations.

## 2 Method

### *The Pyramid Model*

The Pyramid Model (PM) extends the Triangle Model (TM) to a 3D framework for multi-scale spatial data. In the PM, each point represents a spatial unit in geographic space (Figure 1). The horizontal position (x, y) of the point represents the spatial location of the unit (e.g. centroid). The vertical position (z) indicates the scale of the unit. As an example, a raster can be represented as a 3D point lattice in a PM, where a point located at $(x, y, z)$ represents a $z \times z$ focal window centered at $(x, y)$. Each point is attributed with a spatial metric (e.g., focal statistics, point density or Moran's $I$) computed in the focal window defined by $(x, y, z)$, and the spatial metric can be denoted as $f(x, y, z)$. In such a way, the spatial metric $f(x, y, z)$ computed in all different locations and at different scales are represented in a linked structure in the PM. In addition to regularized raster space, the PM can be built in other configurations to represent spatial data in various tessellations (hexagons, Voronoi polygons and irregular polygons). In these cases, the (x, y, z) coordinates in the 3D space can be defined differently to meet specific analytical purposes. In this research, we will implement 1) inferential point pattern analysis, 2) spatial autocorrelation and 2) regression analysis in the PM. All three examples are configured in raster space.

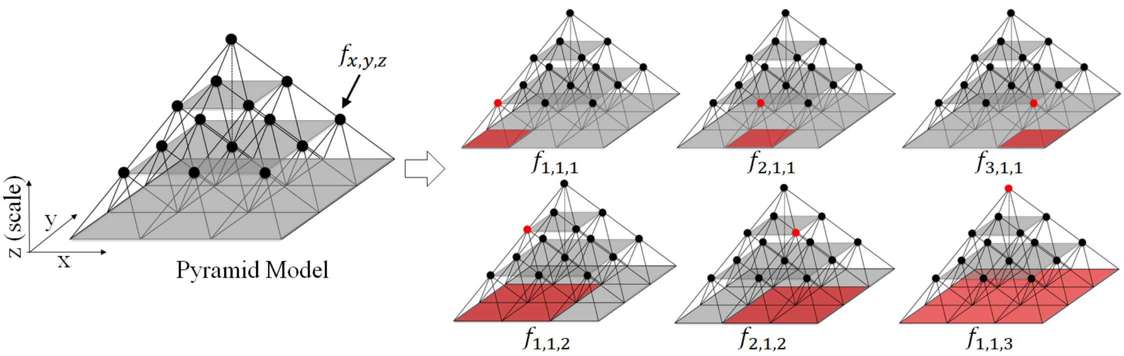


Figure 1: The Pyramid Model (PM) (Qiang et al., 2022).

*Case Study 1: Multi-Scale Point Pattern Analysis*

Our previous research (Qiang et al., 2022) shows that the PM can analyze continuous variation of point density across spatial scales and the nested structure of multi-scale patterns (Figure 1). The presented research implements statistical tests of quadrat analysis and nearest neighbor distance (NND) in the PM. The experiment includes the following steps: (1) a synthesis point set combines simulated point sets in different clustering patterns. The simulated point sets represent geographic processes operating at different scales. (2) A quadrat count, mean NND and their p-values in comparison with complete spatial randomness (CSR) are computed at each point $(x, y, z)$ in a PM. (3) Results are presented using 3D visualization tools to examine the metrics in the PM. The goal of the experiment is to develop quantitative methods to detect multi-scale point patterns, trends and outliers in the PM and associate detected patterns with the underlying processes (the simulated clustering patterns).

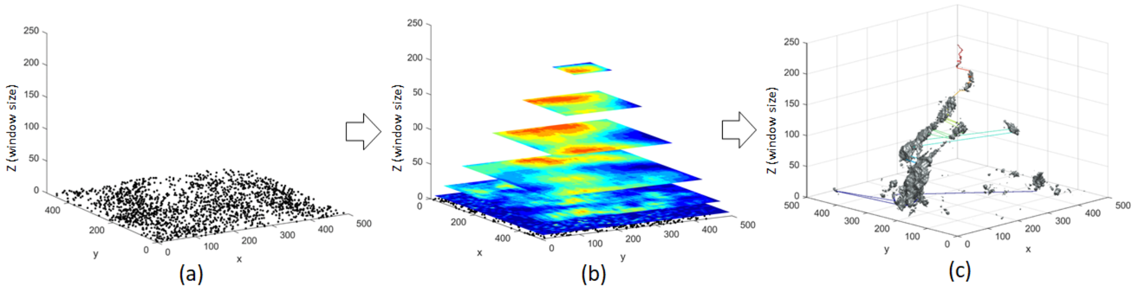


Figure 2: Representing multi-scale point density in a PM. (a) a synthetic point set in a 2D space. (b) point density in different sizes of focal windows. (c) isosurfaces of high-density focal windows in PM.

*Case Study 2: Spatial Autocorrelation*

Spatial autocorrelation (SA) may vary at different spatial scales. Current spatial analysis tools can only compute and display spatial autocorrelation at a fixed scale (i.e. the extent of neighborhood). The PM provides a framework to analyse the variation of spatial autocorrelation across spatial scales. Beginning with raster data (e.g., land cover data or gridded population density), SA indicators (Moran's $I$, Getis-Ord $G$, and their $p$-values) will be computed in different focal windows, which can be projected to 3D points in a PM. The 3D point lattice is then rasterized to volumes, which allows 3D tools to assess the continuous variation of SA indicators across scales. The analysis informs and highlights in which areas and at what scales local clustering/disperse patterns are significant. The unified view of space and spatial scale is expected to unveil the hierarchical structure of spatial autocorrelation at multiple scales and identify places and scales where spatial processes become prominent.

*Case Study 3: Regression Analysis*

The relationship among spatial variables may vary at different scales. We will implement regression analysis in the PM to analyze such relationships at multiple scales. As a notable approach to this problem, Multi-scale Geographically Weighted Regression (MGWR) selects optimal bandwidths to fit regression models for different variables. Unlike MGWR, the PM fits the regression model for spatial variables in different moving windows that are represented as 3D points in PM. Figure 3 illustrates a regression analysis between two spatial variables ($X$ and $Y$) in the focal window ($x$,$y$,$z$). The regression analysis repeats for all points in the PM and generates different regression coefficients, $R^2$ values and $p$-values for the points. Following rasterization, the point lattice is converted to voxels, which allows 3D tools to analyse the variation and hierarchical structure of the regression coefficients at different scales. Through this analysis, we can analyse the variable relationships in different locations and scales and identify areas and scales where the relation is significant.

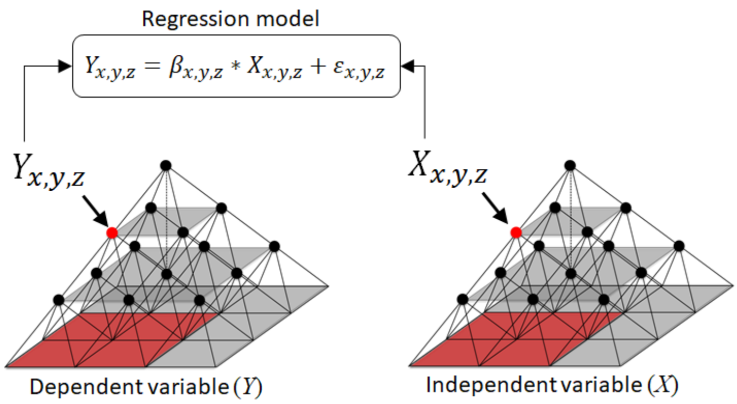


Figure 3: Representing regression analysis between two variables ($X$ and $Y$) in PM.

## 3 Discussion

In this research, we implement inferential statistical methods in the PM to analyze multi-scale spatial patterns and relations. Existing multi-scale analysis methods tend to treat space and spatial scale as separate variables, essentially controlling one variable and examining the change in the other variable. The PM improves upon the conventional representation by integrating space and spatial scale in a 3D environment. In other words, the PM combines “snapshots” of spatial metrics at discrete scales in a unified framework to represent the continuous variation and nested structures of spatial patterns/relations at multiple scales. The three case studies will demonstrate the utility of the PM in searching for areas and scales where inferential tools can explore specific spatial processes as they emerge or disappear. The research will lay a foundation for developing statistical or machine learning methods to detect and associate the metric variation in PM with realistic geographic processes. As the issue of scale underlies both spatial and temporal analysis, our future work includes integrating the PM and the TM into a higher-dimensional framework to support multi-scale spatio-temporal data.

## 4 Acknowledgements:

This research is supported by the Methodology, Measurement & Statistics (MMS) Program of the U.S. National Science Foundation (NSF Award No. 2102019). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

## References

Behrens, T., Viscarra Rossel, R. A., Kerry, R., MacMillan, R., Schmidt, K., Lee, J., Scholten, T., & Zhu, A.-X. (2019). The relevant range of scales for multi-scale contextual spatial modelling. *Scientific Reports*, *9*(1), 14800. https://doi.org/10.1038/s41598-019-51395-3

Carlson, K. (2021). Wetland Classification Accuracy and Scale: Visualizing Uncertainty Metrics across Multiple Resolutions [M.A., University of Colorado at Boulder]. In *ProQuest Dissertations and Theses* (2578122969). ProQuest Dissertations & Theses A&I; ProQuest Dissertations & Theses Global. http://ezproxy.lib.usf.edu/login?url=https://www.proquest.com/dissertations-theses/wetland-classification-accuracy-scale-visualizing/docview/2578122969/se-2?accountid=14745

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 226–231.

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods*, *26*(6), 1481–1496. https://doi.org/10.1080/03610929708831995

Lam, N. S.-N., Cheng, W., Zou, L., & Cai, H. (2018). Effects of landscape fragmentation on land loss. *Remote Sensing of Environment*, *209*, 253–262. https://doi.org/10.1016/j.rse.2017.12.034

Openshaw, S. (1983). *The modifiable areal unit problem*. Norwick: Geo Books.

Qiang, Y., Buttenfield, B. P., Lam, N., & Weghe, N. V. de. (2018). Novel Models for Multi-Scale Spatial and Temporal Analyses. In S. Winter, A. Griffin, & M. Sester (Eds.), *10th International Conference on Geographic Information Science (GIScience 2018)* (Vol. 114, p. 55:1-55:7). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. https://doi.org/10.4230/LIPIcs.GISCIENCE.2018.55

Qiang, Y., Buttenfield, B., & Xu, J. (2022). Analyzing multi-scale spatial point patterns in a pyramid modeling framework. *Cartography and Geographic Information Science*, 1–14. https://doi.org/10.1080/15230406.2022.2048419

Qiang, Y., Chavoshi, S. H., Logghe, S., De Maeyer, P., & Van De Weghe, N. (2014). Multi-scale analysis of linear data in a two-dimensional space. *Information Visualization*, *13*(3), 248–265. https://doi.org/10.1177/1473871613477853

Qiang, Y., & Van de Weghe, N. (2019). Re-Arranging Space, Time and Scales in GIS: Alternative Models for Multi-Scale Spatio-Temporal Modeling and Analyses. *ISPRS International Journal of Geo-Information*, *8*(2), 72. https://doi.org/10.3390/ijgi8020072

Van Kerm, P. (2003). Adaptive Kernel Density Estimation. *The Stata Journal*, *3*(2), 148–156. https://doi.org/10.1177/1536867X0300300204