Investigating the modifiable areal unit problem in Baltimore, Maryland

Proposal, GES Independent Study

Harrison DeFord

2025-03-17

### Introduction

Spatial data is an important component of decision-making, providing utility to disciplines including public health, urban planning, and disaster management. Areal units—polygons that provide a framework for pattern identification—are central to the use of spatial data across all domains. Well-known examples of spatial units used for analysis include Census geographies and neighborhoods: these spatial units allow for the upwards aggregation of raw geographic data. The aggregated data can be used to identify hotspots or clusters, and are often more relatable and useful in policy-writing.

Across all domains, the choice of spatial units in analysis can change the patterns that emerge, influencing both the conclusions drawn from data and the decisions based on them. This issue, known as the Modifiable Areal Unit Problem (MAUP), pervades spatial analysis and must be carefully managed. An analysis which does not consider the MAUP can misconstrue spatial patterns which exist in the real world, and can lead to conclusions which are, at best, misinformed.

As recently described by Wong (2009), the MAUP arises whenever real-world processes are analyzed within the framework of spatial units, whose arbitrary definitions can lead to vastly different, even contradictory, results. It manifests in two key ways: the scale problem, where patterns are exaggerated or diminished depending on the spatial scale of areal units, and the zoning (or boundary) problem, where arbitrary borders separate data in ways that may not reflect underlying spatial processes. With modern GIS tools, users—regardless of experience—can define spatial scales and boundaries with just a few clicks: arguably, the most powerful capability these software packages provide.

Given this, it is important to understand how the MAUP works its way into spatial data when GIS is used as the source for new areal units. While much research has focused on shifting conclusions based on existing administrative units, such as Census geographies, less attention has been devoted to the development of these novel spatial units, from definition to digitization to aggregation.

This paper will attempt to determine the explore different methods for defining spatial units and their impact on observed patterns, using 3-1-1 calls for service across Baltimore, Maryland, as a case study. 3-1-1 data lends itself well to expression of the MAUP for several reasons: first, because 3-1-1 calls are related to specific locations across urban space. Since 3-1-1 calls are related to a specific address, classic administrative boundaries, which are often drawn along geographic features such as roads and rivers, can easily split patterns: consider, for example, a rat infestation in which rodents nest on both sides of a street, the centerline of which demarcates the boundary between two spatial units. Additionally, 3-1-1 calls are important indicators of non-emergency needs for city services. Thus, identification of geographic areas for resource allocation is an important consideration for local government—however, these decisons can easily be influenced by the MAUP.

### Methods

Baltimore City makes historic 3-1-1 customer service requests available via their Open Baltimore data portal as an Esri Feature Service. This paper will focus on the “311 Customer Service Requests 2024” dataset (Baltimore City 2024), as it is the most recent complete calendar year at the time of publishing. U.S. Census Bureau administrative units, such as Census tracts (U.S. Census Bureau 2024), will be used as a framework from which to build novel areal units. In addition, Baltimore City maintains Neighborhood Statistical areas (Baltimore City Department of Planning 2024), which will be used as a starting point for the creation of novel units.

While there is much research available regarding the effect of the MAUP across various existing administrative units, less attention has been given to distinct geographic areas which are built as to not rely on predefined boundaries. R Statistical Software (R Core Team 2024) and ArcGIS Pro (ESRI 2022) will be used to generate novel aggregation units, which units will be used to summarize 3-1-1 calls for service. The latter two data sources mentioned above will be manipulated in order to minimize bias from the creation of their existing boundaries. For example, internal-point and bounding-box centroids will be created, to which data will be aggregated both inwards and outwards—this will minimize the influence of predefined spatial units. In addition, a tessellation will be computed across the region, which will provide raster-like spatial continuity.

A sensitivity analysis will be conducted by comparing descriptive statistics, as well as geographic and numeric distributions, across each type of novel spatial unit. An examination of similarities and differences between the statistical parameters of the various spatial units will provide a comprehensive assessment of how each aggregation method affects the spatial patterns of calls for service.

### Deliverable

The goal for this independent study is, as mentioned previously, a research paper suitable for publishing in an academic journal. The paper will contain written comparison of numeric distributions, as well as thematic maps depicting the geographic distribution of 3-1-1 calls when aggregated into various spatial units. This work is scheduled to be carried out over the summer, with an estimated completion in late August–early September, 2025. Maintaining this timeline will allow the authors to collaborate and edit the written portion of the document before submission to a peer-reviewed journal.

### References

Baltimore City. 2024. “311 Customer Service Requests 2024.” *Open Baltimore*. Baltimore City. <https://data.baltimorecity.gov/datasets/68a1136acff444bba6c93e845dfc00e1_0/explore>.

Baltimore City Department of Planning. 2024. “Neighborhood Statistical Area (NSA) Boundaries.” *Open Baltimore*. Baltimore City. <https://data.baltimorecity.gov/datasets/68a1136acff444bba6c93e845dfc00e1_0/explore>.

ESRI. 2022. *ArcGIS Pro*. Esri Inc. <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

U.S. Census Bureau. 2024. “tl\_2024\_24\_tract.” *TIGER/Line Shapefiles*. <https://www2.census.gov/geo/tiger/TIGER2024/TRACT/>.

Wong, David. 2009. “The Modifiable Areal Unit Problem (MAUP).” In, 105–23. SAGE Publications, Ltd. <https://doi.org/10.4135/9780857020130.n7>.