**Uncertainty in geospatial data visualizations**

**Matthew L. Sisk**

**7-22-2019**

#### Summary of Research

There are nearly as many ways of representing uncertainty in data visualizations as there are sources that discuss these methods. These can range from simple calls to include error bars in all charts to highly complex methods for coding varying levels of uncertainty across multiple variables [(Brodlie, Allendes Osorio, & Lopes, 2012)](https://www.zotero.org/google-docs/?6z3oq2). Even within the somewhat smaller subset of geospatial data visualizations, there remains a continuum from very complex ways to include a factor representing uncertainty in raster models [(e.g. Chilès & Delfiner, 2009)](https://www.zotero.org/google-docs/?sUaqKk) to simple scaling of a circle relative to the spatial uncertainty of GPS points. Regardless, there are a few key types of uncertainty that are most often discussed in geospatial data visualization. These include error propagation (i.e. combining multiple data sources without combining the error ranges for them) [(Heuvelink, Burrough, & Stein, 1989)](https://www.zotero.org/google-docs/?qtRoNh), positional uncertainty (i.e. representing something as a fixed point when the location is not known at that resolution)[(Koo, Chun, & Griffith, 2018b)](https://www.zotero.org/google-docs/?1oOlSN) and uncertainty associated with the underlying variables or locations of boundaries (geo-semantic).These are, of course, in addition to the more commonly presented like the Modifiable Areal Unit Problem or the ecological fallacy which are more concerned with borders and how the are determined

While the techniques used to visualize spatial uncertainty are equally varied, a few that are frequently referenced are higher dimension symbologies[(Koo, Chun, & Griffith, 2018a)](https://www.zotero.org/google-docs/?uQrgUb), animation[(Keßler & Lotstein, 2018)](https://www.zotero.org/google-docs/?7PfYoJ) or fuzzy boundaries/edges[(Fisher, 1994)](https://www.zotero.org/google-docs/?iPh7I3). Of these, most require some advance knowledge of core GIS topics and all are most applicable to particular types of data. Training in how to interpret and how and when to use these techniques is thus an important part of GIS training.

#### Implications for Library Visualization Instruction

A primary outcome of this work is to incorporate a training in uncertainty into early GIS and data science workshops in a library setting, so this work will have direct implications for library visualization instruction. While most GIS textbooks make some reference to the effects of uncertainty and how to visualize it, many individuals trained in basic GIS learn it either on their own or through workshops, so putting some discussion of uncertainty into early training will increase exposure to this important topic.

#### Future Work

The goals for the remaining work on this project are threefold:

1. Develop an outline of best practices for working with geospatial uncertainty in data science and GIS contexts. This will encompass a list of basic methods, examples and a brief overview of how to implement them in major geospatial analysis platforms (qGIS, R, ArcGIS).
2. Sharing these best practices with the broader data/library community. This will take the form of a website (either in Notre Dame’s LibGuide platform or a GitHub site).
3. Development of a suite of examples for incorporating examples of uncertainty into GIS and data science training. These modules will be designed for GIS librarians, workshop instructors and other interested parties to drop directly into their existing training sessions. They will be released using a CC license and shared on the same platform chosen above

#### Works Cited

[Brodlie, K., Allendes Osorio, R., & Lopes, A. (2012). A Review of Uncertainty in Data Visualization. In J. Dill, R. Earnshaw, D. Kasik, J. Vince, & P. C. Wong (Eds.), *Expanding the Frontiers of Visual Analytics and Visualization* (pp. 81–109). https://doi.org/10.1007/978-1-4471-2804-5\_6](https://www.zotero.org/google-docs/?HG32ZR)

[Chilès, J.-P., & Delfiner, P. (2009). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons.](https://www.zotero.org/google-docs/?HG32ZR)

[Fisher, P. F. (1994). Probable and fuzzy models of the viewshed operation. In F. Worboys (Ed.), *Innovations in GIS: selected papers from the First National Conference on GIS Research UK*. London: Taylor and Francis.](https://www.zotero.org/google-docs/?HG32ZR)

[Heuvelink, G. B. M., Burrough, P. A., & Stein, A. (1989). Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Systems*, *3*(4), 303–322. https://doi.org/10.1080/02693798908941518](https://www.zotero.org/google-docs/?HG32ZR)

[Keßler, C., & Lotstein, E. (2018). Animation as a Visual Indicator of Positional Uncertainty in Geographic Information. In A. Mansourian, P. Pilesjö, L. Harrie, & R. van Lammeren (Eds.), *Geospatial Technologies for All* (pp. 365–382). https://doi.org/10.1007/978-3-319-78208-9\_19](https://www.zotero.org/google-docs/?HG32ZR)

[Koo, H., Chun, Y., & Griffith, D. A. (2018a). Geovisualizing attribute uncertainty of interval and ratio variables: A framework and an implementation for vector data. *Journal of Visual Languages & Computing*, *44*, 89–96. https://doi.org/10.1016/j.jvlc.2017.11.007](https://www.zotero.org/google-docs/?HG32ZR)

[Koo, H., Chun, Y., & Griffith, D. A. (2018b). Modeling Positional Uncertainty Acquired Through Street Geocoding. *International Journal of Applied Geospatial Research (IJAGR)*, *9*(4), 1–22. https://doi.org/10.4018/IJAGR.2018100101](https://www.zotero.org/google-docs/?HG32ZR)