Areal interpolation as a proxy for polygon to polygon reaggregation: A case study for downscaling data from census tracts to dissemination areas

Michael Markieta

Department of Geography, Ryerson University, michael.markieta@ryerson.ca

Background and Relevance

Areal interpolation in geographic information systems (GIS) is a spatial statistics method that relies on theory generated in geostatistical mathematics. Ordinary kriging, a spatial optimal linear prediction method based on research since the 1930s and published in 1963 (Matheron), allows values at unknown locations to be estimated based on observed values at nearby locations (Cressie, 1990). By applying kriging theory to polygons, data presented in disparate geographic units can be interpolated and reaggregated to a comparable size for analysis (Gotway & Young, 2002).

Methods and Data

This paper explores an attempt to reaggregate census data, specifically average household income (AHI), from census tracts to dissemination areas for the City of Toronto. Data are obtained from the Canadian Census Analyser (2013) for both areal units. The reaggregated prediction of AHI is compared with the observed values to assess the reliability of areal interpolation for demographic data at disparate census geographies. Using the Areal Interpolation wizard in ArcGIS, a prediction surface of AHI is created by applying kriging theory to the census tract geography. The prediction surface is then reaggregated to the dissemination area geography. Observed average household income is compared to the predicted average household income from the areal interpolation.

Results

The results of the areal interpolation wizard produce an estimated spatial distribution of AHI across the City of Toronto at the census tract geography. Nuances in AHI across the City are not captured by the predicted model. Uncertainty metrics, such as error terms for the reaggregated estimates, are geovisualized to assess the strength of the prediction.

Conclusions

Due to the non-stationary characteristics of demographic data, areal interpolation is only able to translate the general patterns of AHI from census tracts to dissemination areas. Ancillary data can be incorporated in the areal interpolation for more accurate predictions, such as in methods like areal weighting or dasymetric modeling (Hawley & Moellering, 2005).

References

Canadian Census Analyser. (2013). Average Household Income, Dissemination Areas and Census Tracts, Toronto, 2006 Census. Retrieved October 16, 2013 from http://dc1.chass.utoronto.ca.ezproxy.lib.ryerson.ca/census/2006/index.html.

Cressie, N. (1990). The origins of kriging. Mathematical Geology, 22(3), 239-252. doi:10.1007/BF00889887

Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. Journal of the American Statistical Association, 97(458), 632-648. doi: 10.1198/016214502760047140

Hawley, K., & Moellering, H. (2005). A comparative analysis of areal interpolation methods. Cartography and Geographic Information Science, 32(4), 411-411. doi: 10.1559/152304005775194818

Matheron, G. (1963). Principles of geostatistics. Economic Geology, 58(8), 1246. doi:10.2113/gsecongeo.58.8.1246