The effect of spatial and temporal scale on the inferences about human exposure to air pollution

Eun-Hye Enki Yoo
Department of Geography, University at Buffalo (SUNY),
Buffalo, NY, 14261-0055
eunhye@buffalo.edu

Position paper for Scale and Spatial Analytics
Feb. 10-11, 2020

The wide availability of location sensing technologies, such as Global Positioning System, has provided a unique opportunity to collect large-scale human activity data with fine spatiotemporal resolutions. Based on this unprecedented fine-grained time-activity (or mobility) data, some studies have highlighted the importance of human mobility on the air pollution exposure estimates. However, much work still remains in understanding the role of human spatial behavior in conditioning exposures to air pollution, because other studies also have shown that mobility-based exposure estimates do not necessarily differ from traditional residence-based exposure estimates.

I argue that these inconclusive findings are closely associated with the spatial and temporal scale at which exposure to air pollution was assessed. Individuals' exposure to air pollution is estimated by characterizing two processes simultaneously — individuals' time-activities and air pollution concentrations at which those activities occurred. Thus, fine-granularity of spatial behavior characteristics do not guarantee more accurate exposure estimates unless air quality is also modeled at fine spatial and temporal scales. Furthermore, the lack of understanding on the time window (study period) sufficient to capture the spatial variations of both individuals' time-activity patterns and air quality is likely to result in error and biased estimates of individuals' exposure to air pollution. For example, the spatial behaviors observed on a single day may not fully represent one's routine activity patterns (Stanley et al. 2018, Yoo 2019) nor the spatial variability of air pollutant concentrations.

The position of this paper emphasizes the importance of **choosing an optimal scale across multiple spatial and temporal processes**. Optimality is often determined by a specific research objective (e.g., population-based or individual-based exposure assessment for health studies), but it is also affected by the feasibility issues such as the availability of data or appropriate methods to transform or assimilate multiscale data with heterogeneous quality. Recent advancements in sensing technology enable investigators and communities to collect diverse types of air quality data. For example, the low-cost and real time sensing capability of portable sensors increase spatial and temporal resolution and data availability, and computer simulation models and remote sensing provide information on air quality with exhaustive spatial coverage at coarse spatial resolutions in the range of 1 to 100 km.

Specifically, for spatial data assimilation (or data fusion), several approaches have been proposed to integrate data obtained from disparate sources that vary in terms of their spatial resolutions. One of the key challenges around data integration lies in the different spatial resolutions of multi-source data, often referred to as the change-of-support problem. Previous works, including area-to-point interpolation (Kyriakidis 2004), Bayesian melding (Fuentes & Raftery 2005), and statistical calibration (Berrocal et al. 2010), brought attention to this issue within multiple disciplines, including geography, geostatistics, and biostatistics, and developed a potential solution. However, additional challenges remain to harness these diverse geospatial data that vary in space and time. Specifically, we can extend the existing data assimilation framework to incorporate a temporal domain, and/or to account for the heterogeneous data quality, while improving the computational efficiency. This gap between different fields of application points to several future research directions. These include the development of:

1. a general framework to easily accommodate spatio-temporal data. Given that most data in environmental fields are spatially and temporally indexed, a framework to accommodate spatio-temporal data should be sought that accommodate the mismatch of spatial and temporal resolutions of data.

2. a flexible statistical model to estimate the bias/error of each different data source. Different data sources are likely to contain different levels of uncertainty (e.g., point measurements from low-cost sensors are likely to contain higher error/uncertainty than point data from state or federal level regulatory monitoring station networks), and these differences in uncertainty need to be properly taken into account.

3. a computationally efficient statistical model to resolve spatio-temporal change of support problems. Bayesian melding has suffered from formidable computational costs, for example. Here we can adopt an efficient computational approach, such as the integrated nested Laplace approximation (INLA) for fixed rank kriging (Cressie & Johannesson 2008), which can provide temporally dynamic and spatially fine air quality predictions.

References

Berrocal, V. J., Gelfand, A. E. & Holland, D. M. (2010), 'A spatio-temporal downscaler for output from numerical models', *Journal of agricultural, biological, and environmental statistics* **15**(2), 176–197.

Cressie, N. & Johannesson, G. (2008), 'Fixed rank kriging for very large spatial data sets', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **70**(1), 209–226.

Fuentes, M. & Raftery, A. E. (2005), 'Model evaluation and spatial interpolation by bayesian combination of observations with outputs from numerical models', *Biometrics* **61**(1), 36–45.

Kyriakidis, P. C. (2004), 'A geostatistical framework for the area-to-point spatial interpolation', *Geographical Analysis* **36**(3), 41–50.

Stanley, K., Yoo, E.-H., Paul, T. & Bell, S. (2018), 'How many days are enough?: capturing routine human mobility', *International Journal of Geographical Information Science* pp. 1–20.

Yoo, E.-H. (2019), 'How short is long enough?: Modeling temporal aspects of human mobility behavior using mobile phone data', *Annals of the American Association of Geographers* (DOI: 10.1080/24694452.2019.1586516).