Remote Sensing in Health
Defining or describing context and measuring exposure while maintaining confidentiality

Panel contribution to the Population-Environment Research Network Cyberseminar,
"People and Pixels Revisited"
(20-27 February 2018)
https://populationenvironmentresearch.org/cyberseminars/10516

By Kathryn Grace
Department of Geography, Environment and Society and the Minnesota Population Institute
University of Minnesota, Twin Cities

There are many important ways that remotely sensed data can be used to support health research, however, commitments to maintaining confidentiality in spatial data collection have brought forward new challenges in aligning data of different spatial and temporal scales (Van Wey et al. 2005, Grace 2017). **In this review I highlight that, because maintaining confidentiality is often of prime concern with geo-referenced health information, remotely sensed data supports contemporary health research in two important ways – 1) by providing contextual information, and 2) serving as a means for identifying vulnerable communities.**

In his chapter in People and Pixels, "Health Applications of Remote Sensing and Climate Modeling", Paul R. Epstein describes five distinct applications of remote sensing relating to health research (Epstein 1998). Broadly speaking, the applications mostly address the ways that remotely sensed data can be used to monitor and model environmental conditions related to the transmission or spread of diseases. Specific diseases and vectors are mentioned - cholera, malaria, dengue fever, mosquitos - Epstein highlights the role of current and future climate conditions, as measured by remotely sensed data, in disease spread and transmission. These applications highlighted in this chapter, the diseases and the role of environmental context are all relevant today. Moreover, the discussion of climate is clearly of great importance in many discussions of data, health and development.

Current trends in how remotely sensed data is being applied to address health research, have actually been pushed in a different direction than those pathways highlighted by Epstein, largely because of data limitations, especially in terms of geo-referenced, individual-level data and also because of an increased interest in fine-scale heterogeneity in outcomes (for some examples see Brown et al. 2014, Grace et al. 2014). While many of the ideas in Epstein's 1998 work could still be applied to health research today, greater interests in investigating variability in health outcomes within communities as well as between communities has led to a focus on individual-level health outcomes with attention to context (context often means something like village, town, neighborhood, activity-space or community). While, at the same time, concerns about maintaining confidentiality while also providing fine-scale spatial information are pressing. Epstein's ideas require accurate information about the population at risk for disease (the exposure population) but because a majority of the readily available individual-level data shifts the location of where individuals live[1] (often with no information on how long they have lived there, but this is for a different discussion), it is often not possible to match the individual to the condition of interest. For example, exposure to contaminated water or air pollution may be highly variable within a relatively small community or neighborhood, but because of relatively coarse spatial information on where an individual lives, it is not possible to accurately determine their exposure to the contaminated water or polluted air (for example see Balk et al. 2005).

In the interests of preserving participant confidentially and in the context of vastly increasing types of highly spatially and temporally detailed remotely sensed data, how have scholars responded? Researchers have modified their approaches when merging remotely sensed data with individual-level health data to take advantage of the highly detailed information that remotely sensed data provides on context. One way that remotely sensed data is used to describe a context is in investigations of food insecurity or in investigations of climate extremes. In cases where the focus is food insecurity, remotely sensed data is often used to estimate the amount of food available in a given community (Bahktsiyarava et al. 2018). The remotely sensed data, Normalized Difference Vegetation Index (NDVI) is often used, provides a way to estimate annual agricultural production at a fine scale (Landsat based NDVI data is at a resolution of <1km (USGS 2017)). While health information from a source like the DHS or the World Bank's LSMS does not allow for the identification of the specific spatial coordinates where people live or work, it does allow us to calculate the amount of vegetation within an area that contains the true community-location (Johnson and Brown 2014, Shively et al. 2014). Then, in comparing across time or space, we can investigate how individual outcomes vary during different years according to variability in estimated food availability. Similar approaches can be used to link climate extremes to health outcomes (Davenport et al. 2017, Isen et al. 2017). This approach similarly allows us to identify how individual characteristics interact with community characteristics related to food insecurity, heat waves, droughts, or other contextually relevant factors.

It is important when using remotely sensed data in this way to consider the ways that context relates to individual outcomes and how that can vary over space and over time. Using remotely sensed data to help control for variability due to contextual level factors can also help to develop targeted intervention policies that account for the uniqueness of the community. In other words, through accounting for variability in agricultural production, researchers can potentially develop a more focused analysis on the impacts to health of different household water sources.

---

[1] One major source of data on health and development is the USAID-funded The Demographic and Health Survey (DHS) Program. DHS is a major source of population and health data for the poorest countries in the world and provides high quality and detailed data on individual health outcomes – particularly outcomes that relate to maternal and child health. The primary sampling unit in the DHS are villages or village "clusters.". The DHS maintains confidentiality of the respondents by shifting the spatial coordinates of the cluster in the published data and does not permit access to the original data locations (Burgert et al. 2013). Coordinates for rural locations are displaced by 0-5 km in any direction and 1%, are randomly shifted up to 10 km. For urban locations, the displacement is up to 2 km only. Similar approaches have been adopted by other international organizations as well (the Gates Foundation and World Bank, for example).

Accounting for this kind of contextual level variation, potentially facilitates a more targeted investigation focused on a different aspect of health.

The second use of remotely sensed data that facilitates health investigations while also preserving confidentiality, is using remotely sensed data to identify areas of vulnerability and then conducting specific analyses in those communities. This approach in part has grown from the early warning systems of organizations like the Famine Early Warning System Network's (FEWS NET) (fews.net). In the case of FEWS NET, researchers use differently remotely sensed based measures of rainfall, surface water, temperature and others to identify areas of concern. Their focus is on food production among subsistence producers and so areas where rainfall is lower than usual during the growing season may indicate a food system failure in the future. People who live in these areas may face both short- and long-term adverse health outcomes after a food system failure (Brown and Funk 2008, Funk et al. 2008). Using vulnerable areas, as designated by FEWS NET, as communities to focus analyses or interventions targeting health outcomes, is an important way of investigating health using remotely sensed data. Using some of the cases as described by Epstein in the original chapter (1998), could also be used to target investigations – areas where nearby sources of water show signs of cholera contaminated water, for example. Similarly, communities under threat of drought or severe cold, as determined by remotely sensed data, could be investigated prospectively and retrospectively.

Epstein highlighted five cases where remotely sensed data could be used to support health research. These examples are still highly relevant to contemporary health topics. The component of analysis that Epstein did not account for and that is partially the cause of a shift in directions away from some of the described examples, is the actual logistical and theoretical challenges of merging remotely sensed data with health data. It is possible that small scale health studies can be undertaken and health survey data can be spatially and temporally connected to remotely sensed data. However, the majority of health data lacks specific information on where exactly a person lives or works and how long they have lived/worked in those places, therefore measuring exposure and risk at an individual-level becomes impossible. Remotely sensed data provides a wealth of quantitative-based information, and this is especially important in the poorest communities in the world where there is no alternative survey to provide information on water quality, agricultural production, temperature, rainfall and other factors. Because of humanitarian concerns combined with a commitment to maintain confidentiality, the ways that remotely sensed data can be used to support health research is limited in important ways, but improvements in remotely sensed data and advancements in the theory and methods related to activity space (Perchoux et al. 2013) provide a promising future for researchers interested in the use of remotely sensed data to support health research.

References

Bakhtsiyarava, M., Grace, K., & Nawrotzki, R. J. (2018). Climate, Birth Weight, and Agricultural Livelihoods in Kenya and Mali. *American journal of public health*, (0), e1-e7.

Brown, M. E., & Funk, C. C. (2008). Food Security Under Climate Change. *Science*, *319*(5863), 580-581.

Brown, M. E., Grace, K., Shively, G., Johnson, K. B., & Carroll, M. (2014). Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change. *Population and environment*, *36*(1), 48-72.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). Geographic Displacement Procedure and Georeferenced Data Release Policy for the Demographic and Health Surveys. DHS Spatial Analysis Reports No. 7. Calverton, Maryland, USA: ICF International.

Davenport, F., Grace, K., Funk, C., & Shukla, S. (2017). Child health outcomes in sub-Saharan Africa: A comparison of changes in climate and socio-economic factors. *Global Environmental Change*, *46*, 72-87.

Epstein, P. R. (1998). Health applications of remote sensing and climate modeling. *People and pixels: Linking remote sensing and social science*.

Funk, C., Dettinger, M. D., Michaelsen, J. C., Verdin, J. P., Brown, M. E., Barlow, M., & Hoell, A. (2008). Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. *Proceedings of the national academy of sciences*, *105*(32), 11081-11086.

Grace, K. (2017). Considering climate in studies of fertility and reproductive health in poor countries. *Nature Climate Change*, *7*(7), 479.

Grace, K., Brown, M., & McNally, A. (2014). Examining the link between food prices and food insecurity: A multi-level analysis of maize price and birthweight in Kenya. *Food Policy*, *46*, 56-65.

Isen, A., Rossin-Slater, M., & Walker, R. (2017). Relationship between season of birth, temperature exposure, and later life wellbeing. *Proceedings of the National Academy of Sciences*, *114*(51), 13447-13452.

Johnson, K., & Brown, M. E. (2014). Environmental risk factors and child nutritional status and survival in a context of climate variability and change. *Applied Geography*, *54*, 209-221.

Perchoux, C., Chaix, B., Cummins, S., & Kestens, Y. (2013). Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. *Health & place*, *21*, 86-93.

Shively, G., Sununtnasuk, C., & Brown, M. (2015). Environmental variability and child growth in Nepal. *Health & place*, *35*, 37-51.

United States Geological Survey (USGS). Product Guide: Landsat Surface Reflectance-Derived Spectral Indices. Version 3.5. April, 2017. https://landsat.usgs.gov/sites/default/files/documents/si_product_guide.pdf

VanWey, L. K., Rindfuss, R. R., Gutmann, M. P., Entwisle, B., & Balk, D. L. (2005). Confidentiality and spatially explicit data: Concerns and challenges. *Proceedings of the National Academy of Sciences*, *102*(43), 15337-15342.