Unmapped terrain and invisible communities: Analyzing topographic mapping disparities across settlements in the United States from 1885 to 2015

Johannes H. Uhl[1], Stefan Leyk[1], Dylan S. Connor[2], Yao-Yi Chiang[3], Craig A. Knoblock[4]

[1] Institute of Behavioral Science, University of Colorado Boulder, Boulder, CO, USA
[2] School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA
[3] Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA
[4] Information Sciences Institute, University of Southern California, Marina del Rey, CA, USA

Keywords: Topographic mapping, human settlements, historical GIS, history of cartography.

Mapping is an important and deeply political process. While much attention is now being devoted to the definition of boundaries (e.g., redlining, gerrymandering, redistricting), less is systematically known about where, when and at what scales maps are first created. This is an area of key concern because the creation of maps is key to generating spatial, topographic, demographic, or socio-economic data, resources which are of great strategic and economic importance. The absence of such information can, among other processes, impede strategic planning, political transparency, and sustainable development. There is thus much to learn about where and when maps are created, and which communities are either prioritized or "undermapped" within this decision-making process.

We know that while Europe and North America are well mapped today, many of the world's undermapped communities are located in the Global South. Similar disparities most certainly existed historically within Europe and North America too. To date, however, there has been very little quantitative examination of when communities within these regions were mapped, and what forces motivated these mapping decisions (e.g., economic expansion, resource extraction, sociocultural change). Herein, we use novel geospatial data sources to shed light on these processes by examining mapping in the United States between 1885 and 2015 from a quantitative, spatial-historical perspective.

Specifically, we employ two data sources: Metadata on the United States Geological Survey (USGS) Historical Topographic Map Collection (HTMC, Allord et al. 2014), and historical settlement data from the Historical settlement data compilation for the U.S. (HISDAC-US, Leyk & Uhl 2018). HTMC metadata allows for identifying areas that were topographically mapped, and those that were not mapped, in a given point in time, and HISDAC-US allows for identifying areas containing human settlements, or built-up areas at that time. Based on these historical depictions of mapped and built-up areas, we use spatial and statistical analysis to measure the interactions between these two spatial processes.

The HTMC is a collection of over 190,000 topographic map sheets from the 1884 to 2006 that were systematically scanned, georeferenced and made available to the public in a cloud-based data repository[1]. Moreover, the HTMC is equipped with a rich metadata set[2] containing information on the spatial coverage, the temporal reference (e.g., map edition year, map revision year, survey year), the map scale, and the color and print properties of each individual map sheet. These metadata are useful for conducting analyses of the dynamics of topographic mapping in the U.S. since the 1880s, such as the distribution of map scales over time (Fig. 1). As shown in Fig. 1, the most common map scales used since 1884 are scales 1:24,000 (used since the 1930s) and 1:62,500 (used until the 1970s).

[1] https://thor-f5.er.usgs.gov/ngtoc/metadata/htmc
[2] https://thor-f5.er.usgs.gov/ngtoc/metadata/misc/

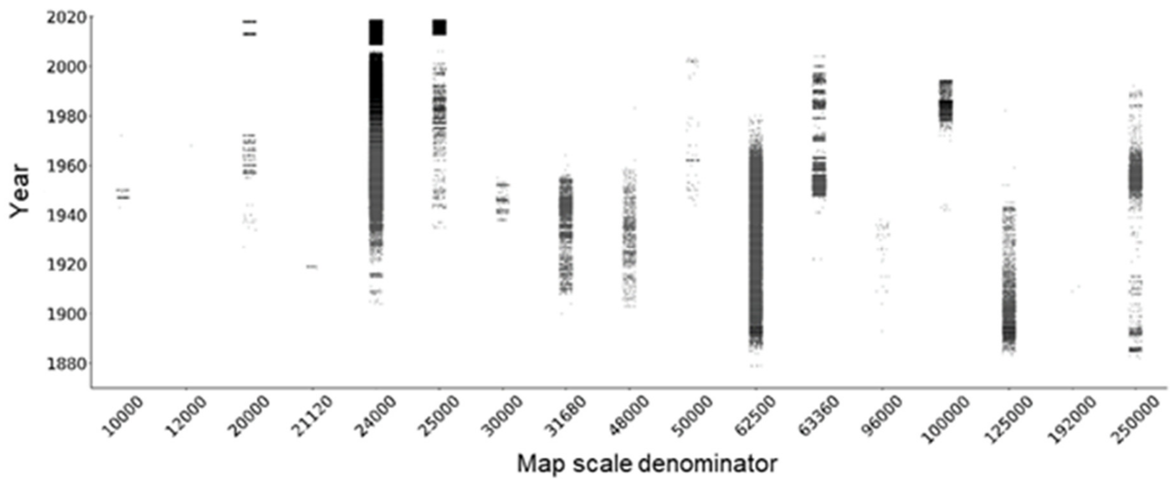

Fig. 1. USGS HTMC map scale distribution over time.

We used these metadata to reconstruct mapped, and unmapped, areas at annual resolution for the conterminous United States from 1885 to 2015, focusing on the largest available, and most commonly used map scales 1:24,000 and 1:62,500 (Fig. 2 top part). These depictions use the map quadrangles (7.5’x7.5’ for scale 1:24,000, 15’x15’ for scale 1:62,500) as analytical unit (https://doi.org/10.6084/m9.figshare.17209433.v2).

In prior work, we described a large data production effort that uses a large, industry-generated property database (Zillow Transaction and Assessment Dataset, ZTRAX[3], to generate gridded, historical settlement layers for the United States from 1810 to 2015 (Leyk & Uhl 2018). These datasets measure the number of built-up properties per 250x250m grid cell, in intervals of 5 years (Uhl et al. 2021, see Fig. 2 bottom part). Herein, we consider a grid cell as “built-up” if it contains at least one built-up structure.

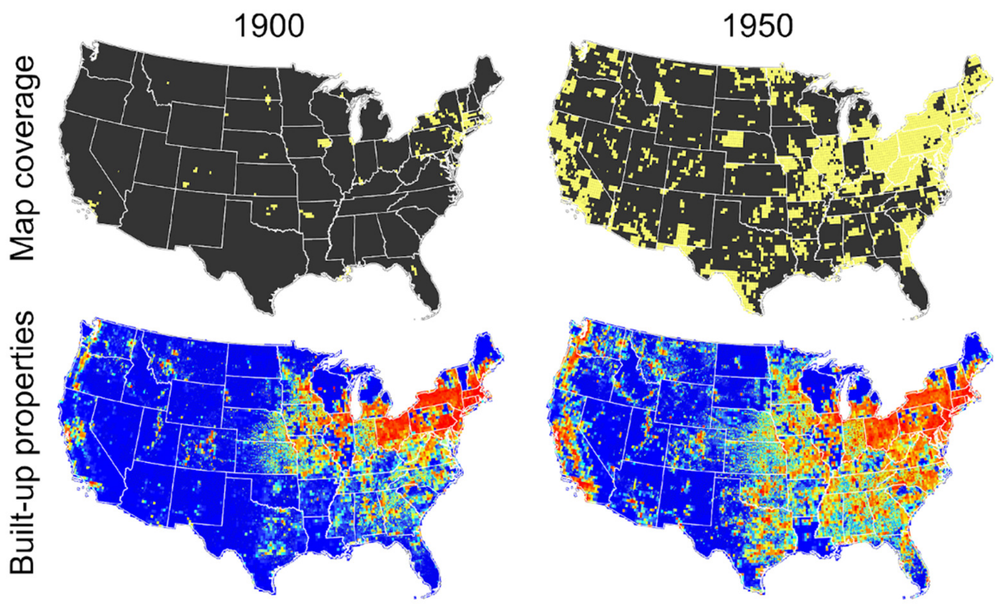

Fig 2. Historical depictions of mapped areas and human settlement distribution. Top row: mapped terrain (scale 1:62,500 or larger) in 1900 and in 1950, in yellow; bottom row: number of built-up properties per map quadrangle in 1900 and in 1950, on a continuous scale from not built-up (blue) to high settlement density (red).

[3] https://www.zillow.com/research/ztrax/

HISDAC-US enables the creation of unique depictions of the historical settlement distribution in the United States and have been employed in a variety of studies, such as on the evolution of urban structure (Uhl et al. 2021) and urban road networks (Burghardt et al. 2022), or on long-term land development (Leyk et al. 2020).

We integrated HTMC metadata-based historical map quadrangle polygons with the gridded, historical settlement layers from HISDAC-US. Specifically, for each year from 1885 to 2015, we calculated zonal sums of built-up properties and of built-up areas contained within each map quadrangle. This integrated dataset allows both, cross-sectional and longitudinal assessments of land development within mapped and unmapped areas. For example, Fig. 3 shows the percentage of the built-up areas and built-up properties located in terrain mapped at different map scales. In 1940, 80% of the built-up properties were located outside of areas mapped at large scales. Interestingly, it was only in 1975 when all built-up areas existing at that point in time were mapped at large scales.

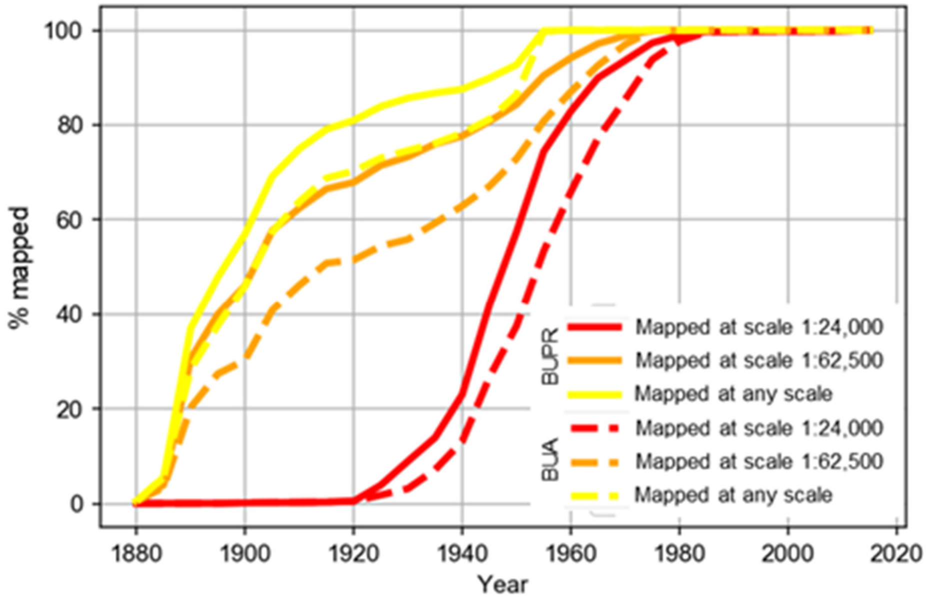

Fig. 3. Percentages of the historical built-up areas (BUA) and built-up properties (BUPR) mapped in each year from 1880 to 2012, for maps of any scale (yellow), of scale 1:24,000 (red), and of scale 1:62,500 (orange).

These preliminary results tentatively reveal the relationship between presumed mapping priorities and historical settlement dynamics. Our ongoing work focuses on describing how the undermapped communities revealed in our analysis differ from that were prioritized for mapping. We are using historical census data to generate estimates of the population living in unmapped areas. We are analyzing the socio-economic characteristics (i.e., working sector, income levels, racial and cultural composition) of populations living in mapped and unmapped areas to examine the social and political determinants of mapping. Moreover, the created, integrated dataset of historical mapping coverage and settlement dynamics can be used to identify if and to what degree mapping was driven by urbanization processes. Were newly and more frequently mapped areas predominantly characterized by recent increases in built-up area? Were these decisions economically motivated, as is implied by the early mapping of mining towns? We are addressing these questions in our ongoing work through a justice- and equity-oriented framework. The analyses presented here were executed in Python, using open source tools such as pandas, geopandas, and GDAL.

References:

Allord, G. J., Walter, J. L., Fishburn, K. A., & Shea, G. A. (2014). *Specification for the US Geological Survey Historical Topographic Map Collection.* U.S. Geological Survey Techniques and Methods, book 6, chap. B11, 65 p., http://dx.doi.org/10.3133/tm11B6.

Burghardt, K., Uhl, J., Lerman, K., & Leyk, S. (2022). Road Network Evolution in the Urban and Rural United States Since 1900. *Computers, Environment and Urban Systems, 95*, 101803.

Leyk, S., & Uhl, J. H. (2018). HISDAC-US, historical settlement data compilation for the conterminous United States over 200 years. *Scientific data*, *5*(1), 1-14.

Leyk, S., Uhl, J. H., Connor, D. S., Braswell, A. E., Mietkiewicz, N., Balch, J. K., & Gutmann, M. (2020). Two centuries of settlement and urban development in the United States. *Science advances*, *6*(23), eaba2937.7

Uhl, J. H., Connor, D. S., Leyk, S., & Braswell, A. E. (2021). A century of decoupling size and structure of urban spaces in the United States. *Communications earth & environment*, *2*(1), 1-14.

Uhl, J. H., Leyk, S., McShane, C. M., Braswell, A. E., Connor, D. S., & Balk, D. (2021). Fine-grained, spatiotemporal datasets measuring 200 years of land development in the United States. *Earth system science data*, *13*(1), 119-153.