Céline Van Migerode
celine.vanmigerode@kuleuven.be
+32 498634192

Identifying discrepancies in urban proxies for city delineation: comparing population, built-up area and night-time lights.

A wealth of urban definitions has been developed, attempted to capture the boundaries of diverse 'city' manifestations across the world. Such globally consistent city delineations are constructed around a variety of urban proxies. For example, residential population grids are often employed to identify cities as agglomerations of population. Other efforts rely on information regarding built-up surfaces to determine city boundaries through presence of human infrastructure, or night-time lights to detect concentrations of economic activity. However, because urbanisation patterns around the globe are highly heterogeneous, a certain proxy might only capture a specific subset of cities and consistently overlook others. For example, utilising night-time lights to identify city boundaries may underestimate the presence of settlements that are not electrified. Understanding the spatially heterogeneous effect of using different urban proxies to determine city boundaries is crucial, as it can have profound implications for subsequent urban and economic analyses.

There is some general understanding in the literature of how urban proxies incorporate biases towards specific world regions. However, explorations into the discrepancy between urban proxies at *city level* remain thin on the ground. Nonetheless, a focus on the city scale is crucial since two cities in different parts of the world (e.g., Paris, Fance and Chengdu, China) might be similarly identified across proxies, whereas two closely located cities (e.g., Cairo, Egypt and Tripoli, Libya) might exhibit vastly different patterns. Furthermore, utilising normative regionalisation such as 'South-East Asia', or the 'Middle East' does not only obscure heterogeneity, but might even reinforce the contemporary assumption that these regions are 'coherent' and 'relevant' units for global geographical analyses.

Against this backdrop, I focus on the city level, and identify which (type of) cities may be consistently overlooked when using a specific urban proxy. In a first step, I create globally consistent city delimitations based on three major urban proxies: population, built-up area and night-time lights. I utilise gridded data at a resolution of 1 km² and implemented a simple delineation algorithm which involves kernel smoothing and applying a statistical threshold. In a second step, I quantify the overlap between the three delineations for each city, and cluster cities with a similar overlap. As an example, cities identified by all three proxies will be clustered together, while those identified by only one or two proxies will be placed in separate clusters. Finally, I explore the spatial distribution of cities in these clusters to identify territorial imbalances in city definitions at a fine spatial granularity.

At first glance, results reveal the expected patters based on prior research: cities solely identified based on population data are predominantly situated in central parts of Africa and countries including India and China, while cities only identified by night-time lights are primarily found in the United States, countries in Europe and Gulf States in the Arabian Peninsula. However, looking at the results in more detail, we find that a large number of cities deviate from these broad regional patterns. For instance, small and medium-sized cities like Wuppertal (Germany), Hastings (UK) or Chernihiv (Ukraine) are not identified based on night-time lights. I will specifically focus on such 'deviating' cities to reveal where and why discrepancies between proxies exist beyond the 'expected' regional patterns and, as such, identify territorial imbalances at the fine granular city scale.