

UC Berkeley

UC Berkeley Previously Published Works

Title

The Right Tools for the Job: The Case for Spatial Science Tool-Building

Permalink

<https://escholarship.org/uc/item/6zt4x7j8>

Author

Boeing, Geoff

Publication Date

2020

DOI

10.1111/tgis.12678

Peer reviewed

The right tools for the job: The case for spatial science tool-building

Geoff Boeing 

Department of Urban Planning and Spatial Analysis, Sol Price School of Public Policy, University of Southern California, Los Angeles, CA, USA

Correspondence

Geoff Boeing, University of Southern California, 301A Lewis Hall, Los Angeles, CA 90089-0626, USA.
Email: boeing@usc.edu

Abstract

This article was presented as the 8th annual *Transactions in GIS* plenary address at the American Association of Geographers annual meeting in Washington, DC. The spatial sciences have recently seen growing calls for more accessible software and tools that better embody geographic science and theory. Urban spatial network science offers one clear opportunity: from multiple perspectives, tools to model and analyze non-planar urban spatial networks have traditionally been inaccessible, atheoretical, or otherwise limiting. This article reflects on this state of the field. Then it discusses the motivation, experience, and outcomes of developing OSMnx, a tool intended to help address this. Next it reviews this tool's use in the recent multidisciplinary spatial network science literature to highlight upstream and downstream benefits of open-source software development. Tool-building is an essential but poorly incentivized component of academic geography and social science more broadly. To conduct better science, we need to build better tools. The article concludes with paths forward, emphasizing open-source software and reusable computational data science beyond mere reproducibility and replicability.

1 | THE WRONG TOOLS FOR THE JOB

Do I need to know the precise polygonal geometries of Los Angeles and the University of Southern California (USC) to assert that the latter is within the former? No. My mind contains no such precise geometric model of points and lines, yet I know that USC is in Los Angeles. When humans reason with the real world, they focus on its objects, relations, and processes—rather than starting with geometry—because these are the keys to

understanding and explaining the real world. Our GIS tools, however, usually do the opposite. Built from the geometry up, around the legacy logic of traditional cartography (geometries and layers), most GIS tools today are restricted by that legacy's limited ability to model objects, relations, and processes. A representational tension thus exists in GIScience between being a *geometric* information science and being an *ontological, relational, and processual* information science.

Computational tools help us reason with the world outside. Accordingly, their representations of reality should start with domain theory—well-substantiated systems of ideas to understand and explain phenomena—rather than the constraints of a certain technology or computing platform (Gahegan, 1999, 2018). For geographic research questions, the relevant domain theory often utilizes object-oriented relations and processes, rather than Cartesian abstractions of space and geometry, even if our computational tools cannot. We need tools that fundamentally embody appropriate scientific theory rather than twisting theory to fit within their representational and computational limitations (Harris et al., 2017; Poorthuis & Zook, 2019). Although new data sources and knowledge discovery systems can help us wrestle with tricky questions, we impoverish our ability to reason with computers if we do not center theory when we create computational representations of the real world—even if we must rethink or advance our technologies and tools to do so.

It is relatively easy to level such critiques, but if we want better GIS tools to study sociospatial objects, relations, or processes, we need to build them. Urgent tool-building opportunities exist today across many geographic subdomains. As but one interdisciplinary example, consider recent advances in spatially informed graph models. Modeling spatial dynamics, relations, and topology too often took a back seat historically to geometry, but graphs offer possible ways forward. For instance, geographic knowledge graphs allow us to build spatial information systems around objects and relations¹ rather than geometries, to better answer ontological spatial questions (Yan, Janowicz, Mai, & Zhu, 2019). Yet such tools remain in their infancy today. As another example—and the motivating example on which this article focuses—graph models of city transportation networks allow topological and dynamical inquiry into urban processes, flows, and structure (Barthelemy, 2011; Marshall, Gil, Kropf, Tomko, & Figueiredo, 2018). Yet such tools traditionally relied on network geometry rather than topology (due to data availability and computational constraints), incorporated domain theory poorly, and were usually ad hoc rather than generalizable, accessible, and reusable (Boeing, 2017).

To conduct better science, we need to build better tools. Such tool-building allows academics to better operationalize and hypothesis-test theory and therefore forms an essential—but poorly incentivized—pillar of scholarly research. In this article, I reflect on my own tool-building experiences in urban planning and geography: facing the need for a better tool to model and analyze urban street networks in a scalable, theoretically sound way, I developed a new open-source Python-based software package called OSMnx. This article considers its history, motivation, and purpose, then reviews its recent use in the empirical street network science literature. In turn, this illustrates the utility of academic tool-building and its downstream—and upstream—value. The article concludes by proposing better alignment of academic incentives with the positive externalities of conducting open science and developing open-source spatial research software.

2 | IF YOU WANT SOMETHING DONE RIGHT ...

Urban science, sitting at the intersection of city planning, geography, and computational data science, aims to advance our knowledge of cities' fundamental patterns and relationships by modeling spatial big data (Alberti, 2017; Acuto, Parnell, & Seto, 2018; Batty, 2013, 2019; Barthelemy, 2019; Kang et al., 2019; Kitchin, 2016; Kontokosta, 2018; Lobo et al., 2020; Mattern, 2013; Sallis et al., 2016; Solecki, Seto, & Marcotullio, 2013). Despite urban science's recent bold claims to a "new kind of science," urban geographers, sociologists, and planners have of course long investigated cities' patterns and processes through spatial data, mathematical models, and the scientific method (Batty, 1971, 1980; Behrend & Levin-Keitel, 2020; Bertuglia, Bianchi, & Mela, 1998; Branch, 1966;

Burgess, 1925; Derudder & van Meeteren, 2019; Hoyt, 1951; Johannesen, Olaisen, & Olsen, 1998; Lee, 1973, 1994; O'Sullivan & Manson, 2015). Computational geography itself now has a long history, yet, too often, geographic science and domain theory fail to fully permeate our computational tools (Arribas-Bel & Reades, 2018; Gahegan, 1999, 2018, 2020; Harris et al., 2017; Singleton & Arribas-Bel, 2019).

Why is that? Reflecting on these insufficient links between GIScience—the scholarly field—and GISystems—the software and tools, Gahegan (2018) highlights two themes of particular relevance here. First, he argues that the GIScience research community usually does not develop its own software tools because it is in nobody's short-term interests to do so. There are of course exceptions to this rule (some of which will be discussed later) but in general too little geographic science and theory make their way into reusable, accessible tools due to misaligned incentives, expectations, and training in academia. Second, and in turn, Gahegan argues that GIScientists must foster a more robust software development community to build and democratize better scientific research tools that are accessible and available to everyone.

Geography journals have witnessed a recent surge of attention to this under-appreciated importance of academic tool-building. For instance, Poorthuis and Zook (2019, p. 8) argue that “as a discipline we need to take charge of building and maintaining our own software platforms. These platforms should be open, accessible, and modifiable by the entire academic community and reflect the diversity and heterogeneity of our discipline.” Yet, to date, such tools rarely materialize in practice because, as Gahegan (2018, p. 24) puts it, we have an “academic culture that fails to reward those who build or maintain tools and software and encourages a short-sighted and individualistic approach to research.” Along similar lines, Rey (2019, p. 7) recalls as a junior scholar being told to stop developing tools because “You need to be writing papers.” He continues, “My colleagues were being brutally honest and trying to reign in my idealism so that my efforts were more aligned with the realities of promotion and tenure cases at the time.”

The widespread disincentivization of academic tool-building produces several negative outcomes. First, most of our tools rely on impoverished representations of geographic theory because our theoreticians have little incentive or training to build tools. Second, most scientific computational workflows exist only as ad hoc scripts to answer a specific research question before being shelved, rather than being generalizable, documented, shared, and accessible.² An enormous amount of scholarly effort is wasted as we endlessly reinvent each other's wheels. Third and accordingly, reproducibility and replicability remain an outstanding challenge (Brunsdon, 2016; Kedron, Frazier, Trgovac, Nelson, & Fotheringham, 2019; Koster & Rowe, 2020). This has become a key motivation for the open-source and open-science movements (Donoho, 2017; Rey, 2009). But as a scientific community we need to go beyond mere reproducibility and replicability to consider the public *reusability* of our tools and workflows. Otherwise we fail to unlock the broader benefits and spillover effects of tool-building.

These problems plague most geographic disciplines, including the study of cities. Today, a consensus is growing around the importance of harnessing GIScience to the open-source and open-science movements to make urban scientific research more tractable, replicable, theoretical, impactful, and approachable for non-computer scientists. Yet urban science too often lacks open data sources and reusable, accessible, theoretically sound tools. Incentivizing authors to share their data and computational workflows when submitting journal articles is one nascent step in the right direction. Incentivizing and building reusable, accessible, theoretically sound tools is another.

3 | TOOLS FOR STREET NETWORK SCIENCE

Many of these challenges hit close to home for me. A few years ago, I wanted to conduct a nationwide study of US street network form to better understand the fundamental characteristics and outcomes of different urban planning paradigms. How did network structure change as planners reoriented cities around the spatial logic of the automobile? And what do we see in recently built neighborhoods, considering the rise of neotraditional urban

design practice? I was interested in the geometry of these networks, but more important was their topology and the sociospatial dynamics they underlie and organize.

Initially I assumed that some tool must already exist to automatically construct non-planar directed graph models from ubiquitously available street network topology data. I was wrong on both counts. On the tool side, while some scholars had previously studied similar topics, no one's research software appeared to be publicly available, well documented for reuse, or sufficiently scalable. Meanwhile, on the data side, mapping platforms like Google Maps did not offer their spatial network data for download. What was ubiquitously available was the US Census Bureau's TIGER/Line roads shapefiles, but they place primacy on geometry and contain insufficient topological information to properly model non-planar networks (Boeing, 2020c). I eventually turned to OpenStreetMap.

First launched in 2004, OpenStreetMap is a wiki-style worldwide mapping project and geospatial data repository with good coverage and quality (Barron, Neis, & Zipf, 2014; Basiri et al., 2016; Corcoran, Mooney, & Bertolotto, 2013; Girres & Touya, 2010; Haklay, 2010; Maier, 2014; Sehra, Singh, Rai, & Anand, 2020; Zielstra, Hochmair, & Neis, 2013). To date, over 1 million different users have contributed content, including 6 billion nodes (i.e., geospatial points), 600 million ways (i.e., geospatial lines and boundaries), and related descriptive data. It is not a perfect data source: researchers estimate that more than 95% of OpenStreetMap contributors are male, suggesting the possibility of correlated biases in content creation (Graham, De Sabbata, & Zook, 2015; Schmidt & Klettner, 2013). Nevertheless, OpenStreetMap is public, free, and an Open Source Initiative affiliate. Volunteers provide some editorial oversight of edits, but anyone may edit the map using tools such as Esri's ArcGIS Editor for OpenStreetMap. OpenStreetMap contains data on streets and highways, transit systems, building footprints, parks and plazas, pedestrian and bicycle infrastructure, political boundaries, and more (though non-road coverage varies around the world). But best of all, unlike a shapefile, OpenStreetMap's data model centers spatial objects and their relations, including both geometric and non-planar topological information.

Researchers typically access OpenStreetMap data through its Overpass API or by downloading a prepackaged regional extract from third-party organizations like Geofabrik. These offer easily ingested street data, but the spatial networks' topological relationships require substantial processing to generate a useful graph model. Researchers often have to write hundreds or thousands of lines of ad hoc code to process the data into a graph and conduct algorithmic analyses for a one-off study. Dozens of small computational and modeling decisions inevitably go unreported in the subsequent peer-reviewed literature, yet these can drastically impact interpretation and replication when every research team codes its own models and analytics. Would it not be better to collectively contribute to and share a reusable, accessible, theoretically sound set of scientific tools?

One obstacle limiting this tool landscape is that high barriers to entry exist for all but those fluent in computer science and domain theory. When I first reviewed the urban spatial networks literature, I was struck by how many modeling methodologies made unexplained or even unjustified assumptions about either urban theory or spatial network theory. The landscape lacked tools to handle non-planar representations of space, which most spatial networks require due to overpasses and underpasses. For simplicity, many studies resorted to undirected graph models, which work fine for studies of form but poorly for studies of flows that obey directionality constraints. The precise handling of common street network features (such as self-loops, parallel edges, or culs-de-sac) often went undocumented. More fundamentally, the Overpass API was cumbersome to work with directly to pipe data into an appropriate model for spatial/network analysis.

In this context I began to develop what eventually became OSMnx. The tool itself is documented³ in detail elsewhere (Boeing, 2017), and tutorials and usage examples⁴ are available online, but I briefly summarize its functionality here. OSMnx is a Python package for collecting spatial network data from OpenStreetMap and then automatically constructing a directed non-planar graph model. It is built on top of the open-source geospatial Python stack (which I will return to later). OSMnx's key feature is its ease of use. With just one line of code, the researcher can download and model the street network of any study site in the world: cities, towns, neighborhoods, boroughs, counties, states, nations—any spatial boundary that OpenStreetMap has in its database. The researcher can specify the site's drivable, walkable, or bikeable network. OSMnx includes a suite of visualization

tools and graph-theoretic analytics (both geometric and topological) for common transportation planning, urban design, and network science research questions. It can also automatically collect and model elevation, building footprint, and points-of-interest data. The code is documented and open-source, so its formal representations of theory are not black boxes.

This project began as a few lines of Python code in a Jupyter notebook, before being collected into a module, and later refactored into a formal package distributed online. I am not a software engineer per se, but I knew how to code and was conversant in the relevant domain theory. But, in the midst of working on this project for months on end, I realized that the next person interested in similar empirical questions would face the exact same laborious tool-building process I was then struggling through. Accordingly, I found it useful to make OSMnx open-source for three primary reasons. First, it makes empirical work easier to review and reproduce. Second, it allows anyone else to contribute to the tool's ongoing development. As discussed later, other researchers desiring useful extensions to its functionality have been willing to add them to the codebase. Third, it empowers others working in urban science and planning to advance their empirical research on real-world spatial networks with a reusable, accessible, theoretically sound tool. I discuss these latter outcomes in the following section.

4 | EMPIRICAL STREET NETWORK SCIENCE WITH OSMnx

The teleology of tool-building suggests that the real value lies in the end use of the tool, rather than in its origins. The purpose of developing OSMnx was to conduct empirical research on urban form, travel dynamics, and the topological structure of transportation infrastructure. Since its public release three years ago, several such studies have been conducted—by myself and many others—using OSMnx for model generation, indicator calculation and visualization, and trip simulation. To illustrate the downstream benefits of tool-building, this section briefly reviews the recent empirical literature that uses OSMnx.

Across a variety of academic disciplines and study sites, researchers have recently used OSMnx to download data, generate models, and analyze real-world transportation networks. For example, Hofer, Jäger, and Füllsack (2018a, 2018b) model the street network of Graz, Austria, to simulate carbon dioxide emissions and traffic congestion and avoidance behavior using mobility data. Saha, Schweitzer, Scaglione, and Johnson (2019) model the streets of Mesa, Arizona, to generate a synthetic feeder network for electrical distribution. Wang, Gao, Feng, Murray, and Zeng (2018) model Washington, DC's street network to develop a geoprocessing framework for optimizing the meetup locations of multiple people under congested traffic conditions. Liu, Zhang, Jin, and Liu (2020) model Beijing's walkable street network to explore spatial patterns of residents' daily leisure activities. Natera Orozco, Battiston, Iñiguez, and Szell (2019) model "as-is" bicycle networks to demonstrate how cities can make small but targeted infrastructure investments to significantly increase their connectivity and directness. Dumedah and Eshun (2020) model Ghanaian street networks to investigate paratransit service coverage through GPS data. Riascos and Mateos (2020) model Manhattan's street network for their study of more than 1 billion taxi trips in New York.

4.1 | Investigating new urban technologies

Studies such as these often investigate the frontier of new transportation and smart cities technologies, including autonomous vehicles, electric vehicles, ride-sharing, and bike-sharing. Beirigo, Schulte, and Negenborn (2018) use OSMnx to model service levels, operational and infrastructure costs, and fleet utilization in hybrid street networks with both autonomous-ready and not autonomous-ready zones. Lin, Deng, Sun, and Chen (2018) model Manhattan's street network alongside travel demand data to optimize ride-share routing. Luo et al. (2020) model

Shanghai's street network to predict demand for electric vehicle sharing systems, while Zhang, Lin, and Mi (2019) model Shanghai's bicycle network to propose a framework for planning dockless bike-sharing services' geofences.

4.2 | Network structure and urban centrality

Other studies look for fundamental relationships between topological structure—particularly network centrality and robustness—and travel patterns and land use. Wang, Chen, Mu, and Zhang (2020) use OSMnx to model Atlanta's street network to investigate ride-sharing accessibility as a function of network centrality and structure. Wang, Xu, and Guo (2018) model Shenzhen's street network to explore the relationship between street centrality and land use intensity. Hellervik, Nilsson, and Andersson (2019) develop a preferential centrality measure to predict urban activity based on street network structure. D'Angelo, Ferretti, and Ghini (2017) model the street network of Fano, Italy, to identify locations of high betweenness centrality. Masias, Hecking, Crespo, and Hoppe (2019) operationalize a spatial capture–recapture methodology to model social media users as a function of walkable street network centrality indicators. Dingil, Rupi, and Stasiskiene (2019) compute and visualize indicators of connectivity, centrality, and clustering across 86 urban areas worldwide to identify the role of network design in easing traffic congestion. Torres, González, García, and Fernández (2019) and Morelli and Cunha (2019) use centrality indicators to measure street networks' vulnerability to perturbation in Mexican and Brazilian cities, respectively (see also Baumann & Keupp, 2020; Sohounou, Christidis, Christodoulou, Neves, & Presti, 2020).

4.3 | Computer science methodological research

Computer scientists and statistical physicists often adopt urban transportation networks as tractable, real-world systems that can be well represented by graphs. OSMnx has been used accordingly to generate input graphs and feature sets for methodological research in machine learning and network algorithms (Yin et al., 2020; Young & Eccles, 2020). Ren, Cheng, and Zhang (2019) model Chengdu's street network then predict traffic flow with a deep spatiotemporal residual neural network. Law and Neira (2019) model the street networks at the centers of 100,000 cities worldwide to train a convolutional autoencoder to analyze network structure. O'Keefe, Santi, and Ratti (2019) train a recurrent neural network to reproduce vehicular mobility patterns, using taxi data and street network models. Martínez Mori and Samaranyake (2019) model several cities' street networks to empirically demonstrate heuristic approximation algorithms that make certain network analyses computationally tractable. Feng and Porter (2020) model cities around the world to explore their topology through persistent homology. Samson, Velez, Nobleza, Sanchez, and Milan (2018) model Filipino cities to develop a genetic algorithm for optimizing paratransit services in developing countries. Senturk and Kebe (2019) model Turkish cities to develop a heuristic solution to a clustered variant of the classic traveling salesman problem. Neukart et al. (2017) model Beijing's street network to develop a quantum annealer for traffic flow optimization on hybrid quantum computing hardware.

4.4 | Indicator calculation and visualization

Other researchers have used OSMnx for automated indicator calculation and visualization. Brandily and Rauch (2018) calculate street network indicators in 1,800 towns across sub-Saharan Africa to explore the relationship between street density and population growth. Holub (2017) calculates and visualizes indicators of bicycle network structure and connectivity in Austin, Charlotte, Columbus, and Minneapolis. Quistberg et al. (2019) calculate transportation network indicators to build a new platform for conducting cross-country urban health studies. Van

Etten, Lindenbaum, and Bacastow (2019) and Verendel and Yeh (2019) visualize street network characteristics across their respective studies' sites. Boeing (2020a) models the street networks of every US urbanized area, city/town, and Zillow-defined neighborhood to calculate dozens of indicators across tens of thousands of study sites at multiple spatial scales, then shares these models and indicators in a public repository. Wang, Yu, et al. (2020) calculate network structure indicators to compare American and Chinese cities. Natera Orozco, Deritei, Vancso, and Vasarhelyi (2020) compute quality-of-life indicators in Budapest by modeling its pedestrian network and local amenities.

4.5 | Urban morphology

Several morphological studies of urban form, sprawl, and density have used OSMnx as well. Gervasoni, Bosch, Fenet, and Sturm (2017) measure urban sprawl and network-constrained destination accessibility. Bristow (2019) compares alternative building density metrics in the context of urban morphology. Abdelkader, Boeing, Fasy, and Millman (2018) and Boeing (2020c) measure urban non-planarity to retheorize how planar assumptions impact street network analyses and to suggest better models. The World Bank collects building footprint data then trains a random forest model to identify OpenStreetMap coverage gaps that can inform future crowd-sourcing efforts and mapping campaigns (Jones, 2019). Boeing (2020b) explores the growing role of big data in computational urban morphology and visual analytics (Figure 1). Boeing (2019b) conducts a cluster analysis of urban street

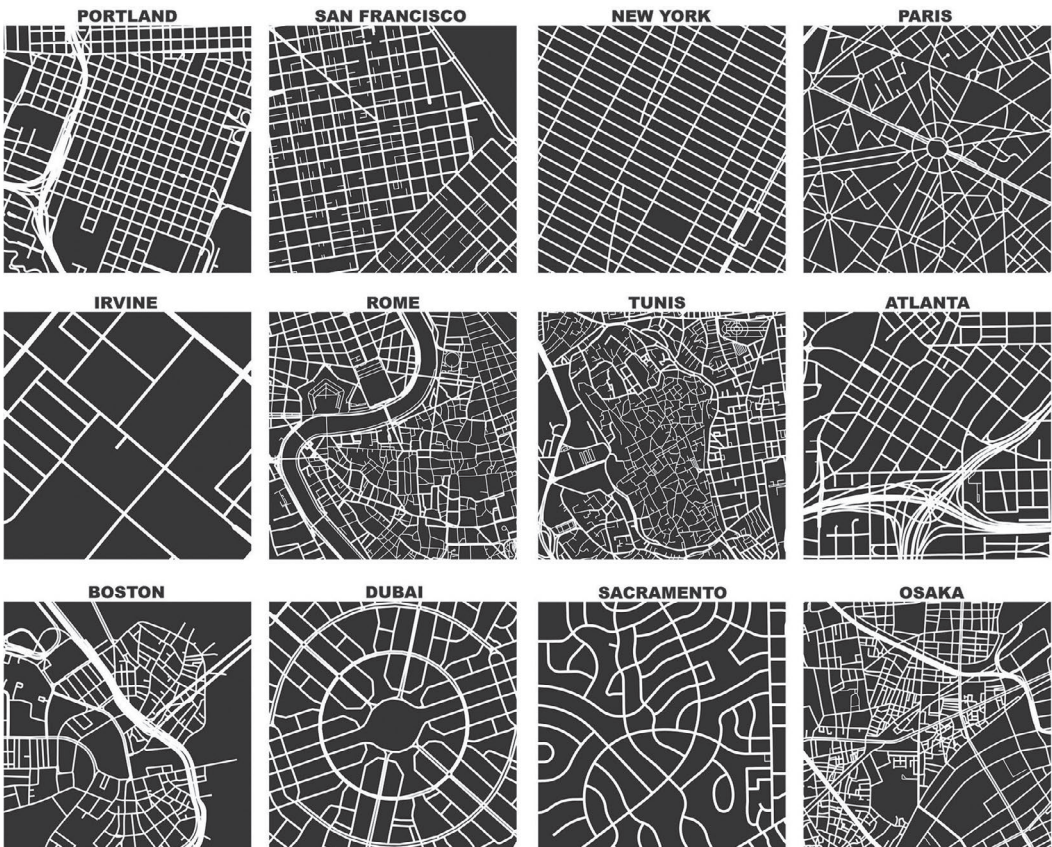


FIGURE 1 One square mile of different cities' street networks, held at the same scale to compare the urban form and grain (Boeing, 2020b)

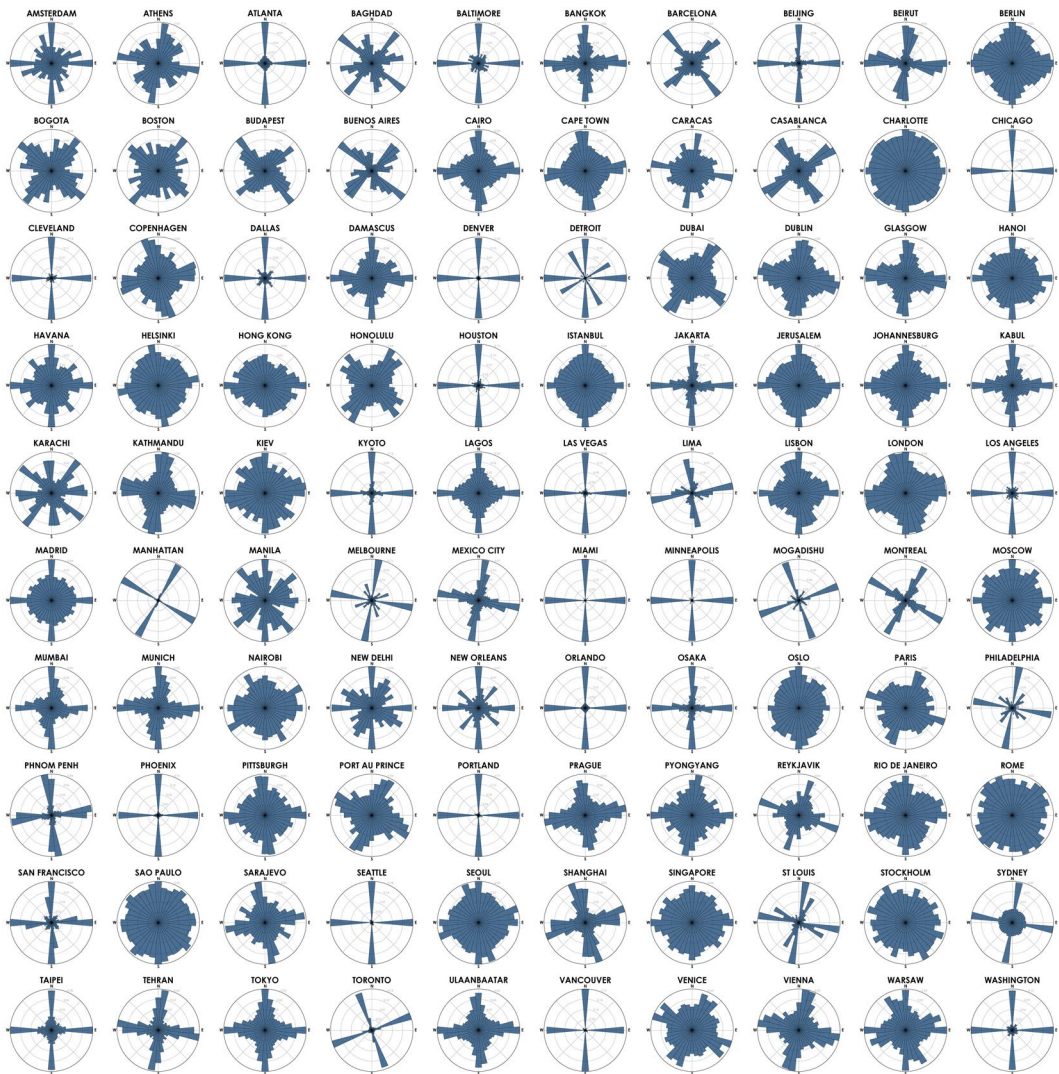


FIGURE 2 Polar histograms of street orientation in 100 world cities. Each histogram bin represents 10° around the compass, and each bar's length represents the relative frequency of streets with compass bearings falling in that bin (Boeing, 2019b)

networks around the world to theorize how planners produce different forms of geometric spatial order, illustrated in Figure 2. Coutrot et al. (2020) build on this theory to explore how spatial order impacts a city's residents' spatial navigation, finding that residents of more grid-like places demonstrate worse ability.

4.6 | Network-constrained trip simulation

Finally, several studies have used OSMnx for simulating trips along a network. Waddell et al. (2018) build regional planning models that integrate land use simulation, travel demand modeling, and traffic assignment. Hernández-Hernández, Siqueiros-García, Robles-Belmont, and Gershenson (2019) calculate commute routes and trip distances in a study of motorists' emotions and expressions of anger in Mexico City. Boeing (2019a) compares driving

and walking route circuitry by simulating millions of trips across 40 city street networks. Merchán, Winkenbach, and Snoeck (2020) and Merchán and Winkenbach (2019) investigate last-mile logistics circuitry in São Paulo and develop better approximation algorithms for urban route distance prediction. Padgham et al. (2019) examine hospital siting by simulating network-constrained stroke service center catchment basins, while Lin, Zhang, Zhu, and Meng (2019) identify the determinants of bicycle catchment basins around Shanghai's metro stations. Liao, Gil, Pereira, Yeh, and Verendel (2020) model street networks in several world cities to compare travel time by personal automobile and public transit. Finally, several recent studies have used OSMnx during the global Covid-19 pandemic for spatial epidemiology and health-care accessibility modeling (e.g., Adler et al., 2020; Kang et al., 2020).

4.7 | Summary

The preceding survey of recent studies covers a wide range of spatial topics and disciplines, from urban design to public health, transport engineering, and computer science. These research projects investigate various travel modes, from walking to cycling, driving, and ride-hailing. They study neighborhoods and cities in every inhabited continent. Most importantly, I could not have conducted most of these studies myself as they exceed my knowledge and skill: by making OSMnx open-source and accessible, it has percolated into others' research designs around the world. Academic tool-building thus entails upstream and downstream benefits for our field's wider endeavor of scientific discovery and theory building.

5 | TOOL-BUILDING IN ACADEMIA

A couple of years ago, I gave a talk which partly dove into OSMnx development and the benefits of academic tool-building and open science. At dinner that evening a senior faculty member asked if my development work continued. I said yes and listed a few features in development that would unlock exciting new spatial network analytics and research questions. The faculty member pushed back on this and—echoing Rey's (2019) recollections as a junior scholar—suggested that if I were to continue in academia, that I must give up tool-building. To paraphrase: “You will become known merely as a tool builder rather than a serious scholar. A serious scholar cannot waste time on anything but empirical research and advancing theory.”

Must scholars eschew tool-building in order to successfully further their field? I believe, rather, that we impoverish our field if we do not make our otherwise ad hoc, one-off research tools “open, accessible, and modifiable,” to return to Poorthuis and Zook's earlier words. In particular, there are three primary benefits of open-source tool-building for academics: unlocking your own empirical research; advancing the collective scientific and theoretical endeavor; and impacting a broader audience.

5.1 | Unlocking individual research

First, tool-building unlocks your own research. Nearly all of us in the spatial sciences do *some* tool-building as we create simple macros, scripts, or libraries of code for routinizing mundane data-processing tasks or fitting models or generating graphical and tabular output for subsequent publication. In this era of big data and ubiquitous computation, it is inefficient and limiting to rely on point-and-click interfaces to conduct scientific research. Coding skills thus increasingly appear in spatial science curricula as necessary components of any GIS skill set. Fostering such abilities is crucial as data science and coding grow central to spatial analysis. But most importantly, as researchers motivated to ask new questions and develop new methods, we cannot assume that theory-rich tools

already exist to answer today's important questions. We must build the right tools to answer the right questions—and doing so can open up new personal research trajectories.

5.2 | Advancing the discipline

Second, once we have built these tools, they should be open-source and accessible to advance the wider discipline. Fortunately, this philosophy already has traction in the GIScience and urban science communities. Efforts are under way today to build a national geospatial software center (Geospatial Software Institute, 2020) and other examples abound. Luc Anselin's GeoDa and GeoDaSpace software for spatial analysis and econometric modeling are free and open-source (Anselin, Syabri, & Kho, 2006), as is Serge Rey's Python Spatial Analysis Library (PySAL: Rey, 2019; Rey et al., 2015). Paul Waddell's UrbanSim platform, a Python-based software library that supports urban land use and transportation modeling and simulation, is similarly free and open-source (Waddell, 2002). Beyond this sort of *hero model* of tool-building, many other community-driven open-source spatial software projects exist in the Python ecosystem and other programming languages (e.g., GDAL, geopandas, leaflet, PostGIS, QGIS, r-spatial, and even OpenStreetMap itself). Such projects often utilize a decentralized *many hands model* to grow a community around building better tools and collaborating synergistically around shared goals. Regardless of the exact model, academic-led projects can provide a theory-rich scaffolding on which research can be conducted and other tools (like OSMnx) can be built. These tools and their ongoing development are also important for empowering students and scholars who cannot afford expensive proprietary software licenses. We must reduce such barriers to entry.

Academics too often pull up the gangplanks by placing scientific findings behind paywalls, hoarding useful data sets, or concealing their research software. The intertwined open-science, open-data, and open-source movements address these three respectively by publicly sharing scientific findings, data, and software for the good of society. This in turn disseminates knowledge and empowers the wider scientific community. If our goal as academics is to produce empirical research and advance theory, our time *should* focus on research and writing—but we should also set aside time to build better theory-rich tools to answer difficult questions.

5.3 | Broader community impacts

Third, open-source tool-building impacts a broad community with bidirectional effects as the tools we build eventually underpin others' work downstream, while other researchers contribute back to our projects upstream. As two examples of downstream effects, the mobile crowd-sensing platform CrowdSenSim now uses OSMnx to simulate urban environments (Montori et al., 2019; Tomasoni et al., 2018) and the transportation planning company Remix developed its platform—now deployed in hundreds of cities worldwide—initially using OSMnx to model street networks. Reciprocally, regarding the upstream direction, as an open-source project OSMnx has received hundreds of code contributions from other scholars and members of the public. These contributors helped develop its points-of-interest module, its nearest-node and nearest-edge search algorithms, its building footprint functionality, and much more. I have reaped the benefits of dozens of others' contributions that enhanced the tool in ways I subsequently used to answer my own research questions.

Better tools and data models, spearheaded by academics, can help infuse theory into our field's quantitative work. But if we want better tools, we have to build them. It is not Esri's job to satisfy all the theoretical needs of the spatial sciences. Recent important scholarly critiques have highlighted how today's GIScience tool landscape is inaccessible, atheoretical, and ad hoc (e.g., Gahegan, 2018; Poorthuis & Zook, 2019). One clear path to better link this academic critique with tangible real-world action is to build and incentivize better tools for praxis.

Collectively, we need to spend more time fixing the lack of high-quality accessible tools rather than just writing yet another thinkpiece lamenting their dearth.

6 | TOWARD BETTER TOOLS

Software often feels inevitable because its backstory is often invisible. We click a download link, run an installer, and suddenly have a new tool to use. Yet this conceals years of human decisions, experiences, and constraints shaping software outcomes that are in no way predestined. In many ways, open-source software parallels other public infrastructure such as highways and bridges. Humans plan and engineer infrastructure in specific social, economic, and technological contexts. Like an individual highway's or bridge's broader network of connected roads, any single piece of software represents a complex network of entanglements with many other software projects on which it depends. And like a highway or bridge, software requires years of maintenance, updates, and retrofits after its initial development: although splashy new capital projects often receive funding and adulation, critical routine maintenance work usually receives far less of either.

Tool-building, in all its facets, remains an essential but poorly incentivized pillar of academia today. Computational geographic tools too often impoverish quantitative analytics through poor representations of scientific theory and squander precious time as everyone develops their own ad hoc scripts to solve similar problems. The case of urban street network science illustrated this. Building an open-source, reusable, accessible, theoretically sound tool as public infrastructure has generated various downstream and upstream benefits. So what can we do to foster a robust tool-building community and collectively reap more of these benefits?

Incentives are key. First, academic tenure, promotion, and annual review guidelines should explicitly reward the scholarly value of open-source research software and open-data contributions to better acknowledge their significant value. They should also account for contributions to pre-existing and decentralized projects to encourage collaborative progress and maintenance of the open-source commons. An ideal system would better balance the respective value of research, publication, and tool-building to advancing science and theory. Second, we should train the next generation of practitioners and scholars to be better tool creators and consumers. Curricula should include coding and informatics courses, and pedagogy should emphasize hands-on learning, such as using computational notebooks. Third, academic publication must continue its nascent steps toward open science: journal editors should require the submission and publication of data sets and computational workflows alongside quantitative manuscripts. Finally, we should increase funding opportunities for building and maintaining research software to foster the positive externalities and downstream benefits they generate throughout the research community.

The goal of academic tool-building—and the hope of every such tool-builder—is to construct some kind of useful infrastructure for your field. With better scaffolding in place to connect theory, science, and analytics, we can conduct better research to explore important geospatial questions and foreground objects, relations, and processes. OSMnx has not changed the world or reinvented urban science, but it has made empirical street network analysis a little easier and more reproducible for some urban scholars and practitioners. Hopefully it even unlocked a study or two along the way that otherwise could not have been conducted. Finally, there is one key lesson learned that I would like to share with other budding academic tool-builders: give your tool an easier name to remember and pronounce than “OSMnx.”

ACKNOWLEDGMENTS

This work was funded in part by a grant from The Public Good Projects. I wish to thank the editors of *Transactions in GIS* for the opportunity to present this plenary paper and two anonymous reviewers for their helpful suggestions in improving it. I also wish to thank the innumerable other developers of open-source software on which my

research and tool-building rely. Figure 1 is reprinted with permission from Boeing (2019a) and Figure 2 is reprinted from Boeing (2019b) under the terms of the Creative Commons Attribution 4.0 International License.

ORCID

Geoff Boeing  <https://orcid.org/0000-0003-1851-6411>

ENDNOTES

- ¹ A simple example of these objects and relations would be: Los Angeles → is a → city; USC → is a → university; USC → is in → Los Angeles.
- ² The “rule of three” in computer programming states that if similar code will be used in three or more places, it should be extracted for general reuse to avoid duplication. We might consider similar guidelines in making our research tools publicly reusable.
- ³ OSMnx installation instructions and user documentation are available online at <https://osmnx.readthedocs.org/> and the open-source project itself is hosted at <https://github.com/gboeing/osmnx>.
- ⁴ OSMnx tutorials and usage examples in Jupyter notebook format are available online at <https://github.com/gboeing/osmnx-examples>.

REFERENCES

- Abdelkader, A., Boeing, G., Fasy, B. T., & Millman, D. L. (2018). Topological distance between nonplanar transportation networks. In *Proceedings of the 28th Fall Workshop on Computational Geometry*, Flushing, NY. New York, NY: City University of New York. Retrieved from https://boole.cs.qc.cuny.edu/mgoswami/Proceedings_FWCG2018.pdf
- Acuto, M., Parnell, S., & Seto, K. C. (2018). Building a global urban science. *Nature Sustainability*, 1(1), 2–4.
- Adler, S. O., Bodeit, O., Bonn, L., Goldenbogen, B., Haffner, J. E. L., Karnetzki, M., ... Klipp, E. (2020). *Geospatially referenced demographic agent-based modeling of SARS-CoV2-infection (COVID-19) dynamics and mitigation effects in a real-world community* (MedRxiv preprint). Retrieved from <https://doi.org/10.1101/2020.05.03.20089235>
- Alberti, M. (2017). Grand challenges in urban science. *Frontiers in Built Environment*, 3(6), 1–5.
- Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical Analysis*, 38(1), 5–22.
- Arribas-Bel, D., & Reades, J. (2018). Geography and computers: Past, present, and future. *Geography Compass*, 12(10), e12403.
- Barron, C., Neis, P., & Zipf, A. (2014). A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Transactions in GIS*, 18(6), 877–895.
- Barthelemy, M. (2011). Spatial networks. *Physics Reports*, 499(1–3), 1–101.
- Barthelemy, M. (2019). The statistical physics of cities. *Nature Reviews Physics*, 1(6), 406–415.
- Basiri, A., Jackson, M., Amirian, P., Pourabdollah, A., Sester, M., Winstanley, A., ... Zhang, L. (2016). Quality assessment of OpenStreetMap data using trajectory mining. *Geospatial Information Science*, 19(1), 56–68.
- Batty, M. (1971). Modelling cities as dynamic systems. *Nature*, 231(5303), 425–428.
- Batty, M. (1980). Limits to prediction in science and design science. *Design Studies*, 1(3), 153–159.
- Batty, M. (2013). *The new science of cities*. Cambridge, MA: MIT Press.
- Batty, M. (2019). Urban analytics defined. *Environment & Planning B*, 46(3), 403–405.
- Baumann, P., & Keupp, M. M. (2020). Assessing the reliability of street networks: A case study based on the Swiss street network. In M. M. Keupp (Ed.), *The security of critical infrastructures: Risk, resilience and defense* (pp. 111–129). Cham, Switzerland: Springer.
- Behrend, L., & Levin-Keitel, M. (2020). Planning as scientific discipline? Digging deep toward the bottom line of the debate. *Planning Theory*, 19(3), 306–323.
- Beirigo, B., Schulte, F., & Negenborn, R. (2018). Dual-mode vehicle routing in mixed autonomous and non-autonomous zone networks. In *Proceedings of the 21st International Conference on Intelligent Transportation Systems*, Maui, HI (pp. 1325–1330). Piscataway, NJ: IEEE.
- Bertuglia, C. S., Bianchi, G., & Mela, A. (Eds.). (1998). *The city and its sciences*. Heidelberg, Germany: Physica-Verlag.
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment & Urban Systems*, 65, 126–139.
- Boeing, G. (2019a). The morphology and circuitry of walkable and drivable street networks. In L. D’Acci (Ed.), *The mathematics of urban morphology* (pp. 271–287). Basel, Switzerland: Birkhäuser.

- Boeing, G. (2019b). Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science*, 4(1), 67.
- Boeing, G. (2020a). A multi-scale analysis of 27,000 urban street networks: Every US city, town, urbanized area, and Zillow neighborhood. *Environment & Planning B*, 47(4), 590–608.
- Boeing, G. (2020b). Spatial information and the legibility of urban form: Big Data in urban morphology. *International Journal of Information Management*, , in press. <https://doi.org/10.1016/j.ijinfomgt.2019.09.009>
- Boeing, G. (2020c). Planarity and street network representation in urban form analysis. *Environment & Planning B*, 47(5), 855–869.
- Branch, M. C. (1966). Simulation, mathematical models, and comprehensive city planning. *Urban Affairs Quarterly*, 1(3), 15–38.
- Brandily, P., & Rauch, F. (2018). *Roads and urban growth* (Department of Economics Paper No. 859). Oxford, UK: University of Oxford.
- Bristow, N. J. O. (2019). *Measuring residential building density at the scale of the urban block* (Unpublished PhD dissertation). University of Westminster, London, UK.
- Brunsdon, C. (2016). Quantitative methods I: Reproducible research and quantitative geography. *Progress in Human Geography*, 40(5), 687–696.
- Burgess, E. W. (1925). Can neighborhood work have a scientific basis? In R. E. Park & E. W. Burgess (Eds.), *The city* (pp. 142–155). Chicago, IL: University of Chicago Press.
- Corcoran, P., Mooney, P., & Bertolotto, M. (2013). Analysing the growth of OpenStreetMap networks. *Spatial Statistics*, 3, 21–32.
- Coutrot, A., Manley, E., Yesiltepe, D., Dalton, R., Wiener, J. M., Hölscher, C., ... Spiers, H. J. (2020). *Cities have a negative impact on navigation ability: Evidence from 38 countries* (BioRxiv preprint). Retrieved from <https://doi.org/10.1101/2020.01.23.917211>
- D'Angelo, G., Ferretti, S., & Ghini, V. (2017). Modeling the Internet of Things: A simulation perspective. In *Proceedings of the 2017 International Conference on High Performance Computing and Simulation*, Genoa, Italy (pp. 18–27). Piscataway, NJ: IEEE.
- Derudder, B., & van Meeteren, M. (2019). Engaging with "urban science". *Urban Geography*, 40(4), 555–564.
- Dingil, A. E., Rupi, F., & Stasiskiene, Z. (2019). A macroscopic analysis of transport networks: The influence of network design on urban transportation performance. *International Journal of Transport Development & Integration*, 3(4), 331–343.
- Donoho, D. (2017). 50 years of data science. *Journal of Computational & Graphical Statistics*, 26(4), 745–766.
- Dumedah, G., & Eshun, G. (2020). The case of paratransit: Trotro service data as a credible location addressing of road networks in Ghana. *Journal of Transport Geography*, 84, 102688.
- Feng, M., & Porter, M. A. (2020). *Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence*. Preprint, arXiv:2001.01872.
- Gahegan, M. (1999). What is geocomputation? *Transactions in GIS*, 3(3), 203–206.
- Gahegan, M. (2018). Our GIS is too small. *Canadian Geographer*, 62(1), 15–26.
- Gahegan, M. (2020). Fourth paradigm GIScience? Prospects for automated discovery and explanation from data. *International Journal of Geographical Information Science*, 34(1), 1–21.
- Geospatial Software Institute. (2020). *GIS: Geospatial Software Institute*. Retrieved from <https://gsi.cigi.illinois.edu/>
- Gervasoni, L., Bosch, M., Fenet, S., & Sturm, P. (2017). Calculating spatial urban sprawl indices using open data. Presented at the 15th International Conference on Computers in Urban Planning and Urban Management. July 11-14, 2017. Adelaide, Australia.
- Girres, J.-F., & Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 14(4), 435–459.
- Graham, M., De Sabbata, S., & Zook, M. A. (2015). Towards a study of information geographies: (Im)mutable augmentations and a mapping of the geographies of information. *Geo: Geography and Environment*, 2(1), 88–105.
- Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment & Planning B*, 37(4), 682–703.
- Harris, R., O'Sullivan, D., Gahegan, M., Charlton, M., Comber, L., Longley, P., ... Evans, A. (2017). More bark than bytes? Reflections on 21+ years of geocomputation. *Environment & Planning B*, 44(4), 598–617.
- Hellervik, A., Nilsson, L., & Andersson, C. (2019). Preferential centrality: A new measure unifying urban activity, attraction and accessibility. *Environment & Planning B*, 46(7), 1331–1346.
- Hernández-Hernández, A. M., Siqueiros-García, J. M., Robles-Belmont, E., & Gershenson, C. (2019). Anger while driving in Mexico City. *PLoS ONE*, 14(9), e0223048.
- Hofer, C., Jäger, G., & Füllsack, M. (2018a). Including traffic jam avoidance in an agent-based network model. *Computational Social Networks*, 5(1), 5.
- Hofer, C., Jäger, G., & Füllsack, M. (2018b). Large scale simulation of CO₂ emissions caused by urban car traffic: An agent-based network approach. *Journal of Cleaner Production*, 183, 1–10.

- Holub, T. (2017). *The invisible messenger bag: Cycling advocacy and identity politics in selected U.S. cities* (Unpublished BA thesis). Berkeley, CA: University of California, Berkeley.
- Hoyt, H. (1951). Is city growth controlled by mathematics or physical laws? *Land Economics*, 27(3), 259.
- Johannesen, J.-A., Olaisen, J., & Olsen, R. (1998). The philosophy of science, planning and decision theories. *Built Environment*, 24(2/3), 155–168.
- Jones, N. (2019). *Identifying gaps in OpenStreetMap coverage through machine learning*. Retrieved from <https://towardsdatascience.com/identifying-gaps-in-openstreetmap-coverage-through-machine-learning-257545c04330>
- Kang, J.-Y., Michels, A. C., Lyu, F., Wang, S., Agbodo, N., Freeman, V. L., & Wang, S. (2020). *Rapidly measuring spatial accessibility of COVID19 healthcare resources: A case study of Illinois, USA* (MedRxiv preprint). Retrieved from <https://doi.org/10.1101/2020.05.06.20093534>
- Kang, W., Oshan, T., Wolf, L. J., Boeing, G., Frias-Martinez, V., Gao, S., ... Xu, W. (2019). A roundtable discussion: Defining urban data science. *Environment & Planning B*, 46(9), 1756–1768.
- Kedron, P., Frazier, A. E., Trgovac, A. B., Nelson, T., & Fotheringham, A. S. (2019). Reproducibility and replicability in geographical analysis. *Geographical Analysis*, in press. <https://doi.org/10.1111/gean.12221>
- Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 374(2083), 20160115.
- Kontokosta, C. E. (2018). Urban informatics in the science and practice of planning. *Journal of Planning Education and Research*, in press. <https://doi.org/10.1177/0739456X18793716>
- Koster, S., & Rowe, F. (2020). Fueling research transparency: Computational notebooks and the discussion section. *Region*, 6(3), E1–E2.
- Law, S., & Neira, M. (2019). *An unsupervised approach to geographical knowledge discovery using street level and street network images*. Preprint, arXiv:1906.11907.
- Lee, D. B. (1973). Requiem for large-scale models. *Journal of the American Institute of Planners*, 39(3), 163–178.
- Lee, D. B. (1994). Retrospective on large-scale urban models. *Journal of the American Planning Association*, 60(1), 35–40.
- Liao, Y., Gil, J., Pereira, R. H. M., Yeh, S., & Verendel, V. (2020). Disparities in travel times between car and transit: Spatiotemporal patterns in cities. *Scientific Reports*, 10(1), 4056.
- Lin, D., Zhang, Y., Zhu, R., & Meng, L. (2019). The analysis of catchment areas of metro stations using trajectory data generated by dockless shared bikes. *Sustainable Cities and Society*, 49, 101598.
- Lin, Q., Deng, L., Sun, J., & Chen, M. (2018). Optimal demand-aware ride-sharing routing. In *Proceedings of the IEEE International Conference on Computer Communications*, Honolulu, HI (pp. 2699–2707). Piscataway, NJ: IEEE.
- Liu, Y., Zhang, Y., Jin, S. T., & Liu, Y. (2020). Spatial pattern of leisure activities among residents in Beijing, China: Exploring the impacts of urban environment. *Sustainable Cities & Society*, 52, 101806.
- Lobo, J., Alberti, M., Allen-Dumas, M., Arcaute, E., Barthelemy, M., Tapia, B., ... Youn, H. (2020). *Urban science: Integrated theory from the first cities to sustainable metropolises* (Mansueto Institute for Urban Innovation Research Paper). Chicago, IL: University of Chicago.
- Luo, M., Du, B., Klemmer, K., Zhu, H., Ferhatosmanoglu, H., & Wen, H. (2020). D3P: Data-driven demand prediction for fast expanding electric vehicle sharing systems. *Proceedings of the ACM on Interactive, Mobile, Wearable & Ubiquitous Technologies*, 4(1), 21.
- Maier, G. (2014). OpenStreetMap, the Wikipedia map. *Region*, 1(1), R3–R10.
- Marshall, S., Gil, J., Kropf, K., Tomko, M., & Figueiredo, L. (2018). Street network studies: From networks to models and their representations. *Networks & Spatial Economics*, 18, 735–749.
- Martínez Mori, J. C., & Samaranyake, S. (2019). Bounded asymmetry in road networks. *Scientific Reports*, 9(1), 11951.
- Masias, V. H., Hecking, T., Crespo, F., & Hoppe, H. U. (2019). Detecting social media users based on pedestrian networks and neighborhood attributes: An observational study. *Applied Network Science*, 4(1), 96.
- Mattern, S. (2013). Methodolatry and the art of measure. *Places Journal* (November). <https://doi.org/10.22269/131105>
- Merchán, D., & Winkenbach, M. (2019). An empirical validation and data-driven extension of continuum approximation approaches for urban route distances. *Networks*, 73(4), 418–433.
- Merchán, D., Winkenbach, M., & Snoeck, A. (2020). Quantifying the impact of urban road networks on the efficiency of local trips. *Transportation Research Part A: Policy & Practice*, 135, 38–62.
- Montori, F., Cortesi, E., Bedogni, L., Capponi, A., Fiandrino, C., & Bononi, L. (2019). CrowdSenSim 2.0: A stateful simulation platform for mobile crowdsensing in smart cities. In *Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Miami, FL (pp. 289–296). New York, NY: ACM.
- Morelli, A. B., & Cunha, A. L. (2019). Verificação de vulnerabilidades em redes de transporte. In *Congresso de Pesquisa e Ensino em Transportes da ANPET*, Balneário Camboriú, Brazil.
- Natera Orozco, L. G., Battiston, F., Iñiguez, G., & Szell, M. (2019). *Data-driven strategies for optimal bicycle network growth*. Preprint, arXiv:1907.07080.
- Natera Orozco, L. G., Deritei, D., Vancso, A., & Vasarhelyi, O. (2020). Quantifying life quality as walkability on urban networks: The case of Budapest. In H. Cherifi, S. Gaito, J. F. Mendes, E. Moro, & L. M. Rocha (Eds.), *Complex networks and their applications VIII* (pp. 905–918). Cham, Switzerland: Springer.

- Neukart, F., Compostella, G., Seidel, C., von Dollen, D., Yarkoni, S., & Parney, B. (2017). Traffic flow optimization using a Quantum Annealer. *Frontiers in ICT*, 4, 29.
- O'Keefe, K., Santi, P., & Ratti, C. (2019). *Modeling vehicular mobility patterns using recurrent neural networks*. Preprint, arXiv:1910.11851.
- O'Sullivan, D., & Manson, S. M. (2015). Do physicists have geography envy? And what can geographers learn from it? *Annals of the Association of American Geographers*, 105(4), 704–722.
- Padgham, M., Boeing, G., Cooley, D., Tierney, N., Sumner, M., Phan, T. G., & Beare, R. (2019). An introduction to software tools, data, and services for geospatial analysis of stroke services. *Frontiers in Neurology*, 10, 743.
- Poorthuis, A., & Zook, M. (2019). Being smarter about space: Drawing lessons from spatial science. *Annals of the American Association of Geographers*, 110(2), 349–359.
- Quistberg, D. A., Diez Roux, A. V., Bilal, U., Moore, K., Ortigoza, A., Rodriguez, D. A., ... Miranda, J. J. (2019). Building a data platform for cross-country urban health studies: The SALURBAL study. *Journal of Urban Health*, 96(2), 311–337.
- Ren, Y., Cheng, T., & Zhang, Y. (2019). Deep spatio-temporal residual neural networks for road-network-based data modeling. *International Journal of Geographical Information Science*, 33(9), 1894–1912.
- Rey, S. J. (2009). Show me the code: Spatial analysis and open source. *Journal of Geographical Systems*, 11(2), 191–207.
- Rey, S. J. (2019). PySAL: The first 10 years. *Spatial Economic Analysis*, 14(3), 273–282.
- Rey, S. J., Anselin, L., Li, X., Pahle, R., Laura, J., Li, W., & Koschinsky, J. (2015). Open geospatial analytics with PySAL. *ISPRS International Journal of Geo-Information*, 4(2), 815–836.
- Riascos, A. P., & Mateos, J. L. (2020). Networks and long-range mobility in cities: A study of more than one billion taxi trips in New York City. *Scientific Reports*, 10(1), 4022.
- Saha, S. S., Schweitzer, E., Scaglione, A., & Johnson, N. G. (2019). *A framework for generating synthetic distribution feeders using OpenStreetMap*. Preprint, arXiv:1910.07673.
- Sallis, J. F., Bull, F., Burdett, R., Frank, L. D., Griffiths, P., Giles-Corti, B., & Stevenson, M. (2016). Use of science to guide city planning, policy, and practice: How to achieve healthy and sustainable future cities. *Lancet*, 388(10062), 2936–2947.
- Samson, B. P. V., Velez, G. A. T., Nobleza, J. R., Sanchez, D., & Milan, J. T. (2018). Optimizing the efficiency, vulnerability and robustness of road-based para-transit networks using genetic algorithm. In Y. Shi, H. Fu, Y. Tian, V. V. Krzhizhanovskaya, M. H. Lees, J. Dongarra, & P. Slood (Eds.), *Computational Science - ICCS 2018* (Lecture Notes in Computer Science, Vol. 10860, pp. 3–14). Cham, Switzerland: Springer.
- Schmidt, M., & Klettner, S. (2013). Gender and experience-related motivators for contributing to OpenStreetMap. In *Proceedings of the 16th AGILE International Conference on Geographic Information Science*, Leuven, Belgium (pp. 13–18). AGILE.
- Sehra, S. S., Singh, J., Rai, H. S., & Anand, S. S. (2020). Extending processing toolbox for assessing the logical consistency of OpenStreetMap data. *Transactions in GIS*, 24(1), 44–71.
- Senturk, I. F., & Kebe, G. Y. (2019). A novel shortest path routing algorithm for wireless data collection in transportation networks. In *Proceedings of the Fourth International Conference on Computer Science and Engineering*, Samsun, Turkey. Piscataway, NJ: IEEE.
- Singleton, A., & Arribas-Bel, D. (2019). Geographic data science. *Geographical Analysis*, in press. <https://doi.org/10.1111/gean.12194>
- Sohouenou, P. Y. R., Christidis, P., Christodoulou, A., Neves, L. A. C., & Presti, D. L. (2020). Using a random road graph model to understand road networks robustness to link failures. *International Journal of Critical Infrastructure Protection*, 29, 100353.
- Solecki, W., Seto, K. C., & Marcotullio, P. J. (2013). It's time for an urbanization science. *Environment: Science and Policy for Sustainable Development*, 55(1), 12–17.
- Tomasoni, M., Capponi, A., Fiandrino, C., Kliazovich, D., Granelli, F., & Bouvry, P. (2018). Profiling energy efficiency of mobile crowdsensing data collection frameworks for smart city applications. In *Proceedings of the Sixth IEEE International Conference on Mobile Cloud Computing, Services, and Engineering*, Bamberg, Germany (pp. 1–8). Piscataway, NJ: IEEE.
- Torres, J. E. H., González, S. H., García, J. A. J., & Fernández, V. F. (2019). Análisis de vulnerabilidad de la infraestructura de transporte aplicando redes complejas: Red de avenidas de la ciudad de Celaya, Guanajuato. *Pistas Educativas*, 41, 133.
- Van Etten, A., Lindenbaum, D., & Bacastow, T. M. (2019). *SpaceNet: A remote sensing dataset and challenge series*. Preprint, arXiv:1807.01232.
- Verendel, V., & Yeh, S. (2019). Measuring traffic in cities through a large-scale online platform. *Journal of Big Data Analytics in Transportation*, 1(2), 161–173.
- Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, 68(3), 297–314.
- Waddell, P., Garcia-Dorado, I., Maurer, S. M., Boeing, G., Gardner, M., Porter, E., & Aliaga, D. (2018). Architecture for modular microsimulation of real estate markets and transportation. Presented at the Applied Urban Modelling Symposium. June 27–29, 2018. Cambridge, England.

- Wang, M., Chen, Z., Mu, L., & Zhang, X. (2020). Road network structure and ride-sharing accessibility: A network science perspective. *Computers, Environment & Urban Systems*, 80, 101430.
- Wang, S., Gao, S., Feng, X., Murray, A. T., & Zeng, Y. (2018). A context-based geoprocessing framework for optimizing meetup location of multiple moving objects along road networks. *International Journal of Geographical Information Science*, 32(7), 1368–1390.
- Wang, S., Xu, G., & Guo, Q. (2018). Street centralities and land use intensities based on points of interest (POI) in Shenzhen, China. *ISPRS International Journal of Geo-Information*, 7(11), 425.
- Wang, S., Yu, D., Kwan, M.-P., Zheng, L., Miao, H., & Li, Y. (2020). The impacts of road network density on motor vehicle travel: An empirical study of Chinese cities based on network theory. *Transportation Research Part A: Policy & Practice*, 132, 144–156.
- Yan, B., Janowicz, K., Mai, G., & Zhu, R. (2019). A spatially explicit reinforcement learning model for geographic knowledge graph summarization. *Transactions in GIS*, 23(3), 620–640.
- Yin, Y., Varadarajan, J., Wang, G., Wang, X., Sahrawat, D., Zimmermann, R., & Ng, S.-K. (2020). A multi-task learning framework for road attribute updating via joint analysis of map data and GPS traces. In *Proceedings of the 2020 Web Conference*, Taipei, Taiwan (pp. 2662–2668). New York, NY: ACM.
- Young, D. L., & Eccles, C. (2020). Automatic construction of Markov decision process models for multi-agent reinforcement learning. In *Proceedings of SPIE 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, 114130Y. Bellingham, WA: SPIE.
- Zhang, Y., Lin, D., & Mi, Z. (2019). Electric fence planning for dockless bike-sharing services. *Journal of Cleaner Production*, 206, 383–393.
- Zielstra, D., Hochmair, H. H., & Neis, P. (2013). Assessing the effect of data imports on the completeness of OpenStreetMap: A United States case study. *Transactions in GIS*, 17(3), 315–334.

How to cite this article: Boeing G. The right tools for the job: The case for spatial science tool-building. *Transactions in GIS*. 2020;00:1–16. <https://doi.org/10.1111/tgis.12678>