

Special Issue

Evolution of Urban Patterns: Urban Morphology as an Open Reproducible Data Science

Martin Fleischmann¹ , Alessandra Feliciotti², William Kerr³

¹Department of Geography and Planning, University of Liverpool, Liverpool, UK, ²Department of Architecture, University of Strathclyde, Glasgow, UK, ³Gillespies, London, UK

The recent growth of geographic data science (GDS) fuelled by increasingly available open data and open source tools has influenced urban sciences across a multitude of fields. Yet there is limited application in urban morphology—a science of urban form. Although quantitative approaches to morphological research are finding momentum, existing tools for such analyses have limited scope and are predominantly implemented as plug-ins for standalone geographic information system software. This inherently restricts transparency and reproducibility of research. Simultaneously, the Python ecosystem for GDS is maturing to the point of fully supporting highly specialized morphological analysis. In this paper, we use the open source Python ecosystem in a workflow to illustrate its capabilities in a case study assessing the evolution of urban patterns over six historical periods on a sample of 42 locations. Results show a trajectory of change in the scale and structure of urban form from pre-industrial development to contemporary neighborhoods, with a peak of highest deviation during the post-World War II era of modernism, confirming previous findings. The wholly reproducible method is encapsulated in computational notebooks, illustrating how modern GDS can be applied to urban morphology research to promote open, collaborative, and transparent science, independent of proprietary or otherwise limited software.


Urban data science and morphology

Across all fields of science, the expanding availability of openly accessible data and constant development of new tools, technologies, and platforms is making possible the generation of new theories and testing of new analytical methods and processes. This is expanding our ability to interrogate and understand the world around us, delivering new evidence-based knowledge to guide action. Furthermore, it supports a shift towards quantitative geography, allowing its evolution in the direction of the fourth paradigm of science as well as its closer integration with critical geography. The former, based on the notion that science can be data driven to the point

Correspondence: Martin Fleischmann, Research Associate, Department of Geography and Planning, University of Liverpool, Liverpool, UK
e-mail: m.fleischmann@liverpool.ac.uk

Submitted: 9 March 2021; Revised version accepted: 24 June 2021

doi: 10.1111/gean.12302

© 2021 The Authors. *Geographical Analysis* published by Wiley Periodicals LLC on behalf of The Ohio State University. This is an open access article under the terms of the Creative Commons Attribution License, which permits use,  distribution and reproduction in any medium, provided the original work is properly cited.

where knowledge and theory can be retrieved from data only (Hey et al. 2009; Gahegan 2020), represents one direction of quantitative science. The latter goes in the opposite one, repeatedly suggested by Kwan and Schwanen (2009) or later Derudder and van Meeteren (2019) in their call for a “common language” and onboarding the critical insights stemming from quantitative approaches.

This trend manifests in the rapid growth of quantitative geography and geographic data science (GDS), fuelled by the development of new computational tools and availability of (big) open data. In this field, the quick emergence and maturation of a new generation of spatial data software ecosystems in both Python and R (Rey 2019; Bivand 2020), represented by *GeoPandas* (Jordahl et al. 2021), *PySAL* (Rey and Anselin 2007; Rey 2019), *xarray* (Hoyer and Hamman 2017), or *sf* (Pebesma 2018), has enabled highly specialized research applications—including the analysis and modelling of urban spatial systems.

Concerning cities, these new tools are finding fertile ground in the study of functional aspects of urban life, where we now rely on abundant data on, for example, urban populations coming from census records or social networks, or environmental performance, thanks to arrays of various sensors both on the orbit (Drusch et al. 2012) and on the ground (Mydlarz et al. 2017). This trend is extremely promising and is already bringing fundamental new knowledge on the performance and behavior of lived environments, with concrete applications in policy and planning (Kandt and Batty 2021). Conversely, it is leaving exposed the complementary study of physical aspects of cities—the spatial environment constituting the setting of all lived urban experience. Too often, data on buildings, streets, and open space are used limitedly to provide a hollow cartographic “backdrop” to a wide array of alternative urban dynamics, with little or no further data-driven insight. This is despite the fact that the built environment is not a simple passive component on which social and economic processes happen to occur: It is an active layer constantly influencing and affecting the quality and modes of our existence within cities and that can be directly or indirectly manipulated through design and planning to deliver long-term effects.

In recent years, GDS focusing on urban form has taken initial steps, (Boeing 2018; Oliveira and Medeiros 2016; Araldi and Fusco 2019; Dibble et al. 2019; Jochem et al. 2020) and yet this line of research is still at its infancy. Although most existing tools in quantitative geography are ill-suited for urban form analysis, tools specifically focused on analysis of urban form have a relatively narrow scope. The majority are geared towards the analysis of street networks (Hillier 1996; Porta, Latora and Strano 2010), leaving aside key considerations of geometry and composition of urban fabric. Capturing the complexity of urban form requires more than the characterization of street networks and a handful of other measurable characters.

Aware of this limitation, some researchers have attempted to develop new computational frameworks. These, however, present a range of limitations. Very often, they are built ad hoc and are rarely generalizable beyond the single case. The implemented code is infrequently made publicly available, and in many cases, researchers fail to properly document every design decision in their publications—drastically reducing reproducibility and unnecessarily complicating the process of tentatively rebuilding workflows multiple times. Collectively, this constitutes a substantial obstacle to a truly shared and evidence-based knowledge of urban form. In this sense, developing stronger foundations and bespoke tools for a data-driven science of urban form is key to reducing the existing gap and allowing a more comprehensive and actionable knowledge of the physical structure of the urban environment.

In this paper, we provide an overview of available tools for morphological analysis to understand better the severity of these issues and the potential to overcome them. We then argue that the field needs a shift from dominant traditional geographic information system (GIS) environments based on a graphical user interface (GUI; e.g., *QGIS* or *ArcMap*) towards reproducible open code-based workflows. That is further supported by an overview of the Python ecosystem and its ability to support research applications particularly in the area of urban morphology, a field of study concerned with the analysis of urban form. The suggested approach is then illustrated in the case study analysing alteration in structure and scale of urban patterns depending on their period of origin. We close up with a discussion on the future developments of quantitative urban morphology keeping up with open science and reproducibility.

Morphometric tools in the ecosystem: A state of the art

Available tools

Although the pool of advanced computing techniques for geospatial analysis is in rapid expansion, the current offer of tools geared towards the analysis of urban form is rather limited in scope and inconsistent in representativeness.

Currently, most researchers interested in urban form analysis rely on traditional “point-and-click” GIS software packages, such as *ArcGIS* or *QGIS*. Although more intuitive to use, these have three primary disadvantages. First, some require access to proprietary software, which comes with inherent barriers to accessibility, either related to affordability or to platform compatibility (i.e., *ArcGIS* is available only for Microsoft Windows). Second, even when free to use and multi-platform (i.e., *QGIS*), they are restricted by their underlying one-fits-all architecture. And although in some cases, these can be partially customized through user-developed single-purpose plug-ins, they tend to constrain users within predetermined software capabilities. Consequently, scientific methods are a direct function of the limitations imposed by the software, rather than by the underlying theory or the specific questions at hand (Harris et al. 2017; Poorthuis and Zook 2020). Lastly, as pointed out by, Boeing (2020b) toolkits relying on point-and-click interfaces are inefficient in the era of big data. Due to the limited scope for automation of tasks, not only is workflow efficiency reduced but also the reproducibility of the underlying research is compromised, because this largely depends on the (often undocumented) sequence of decisions manually operating the software.

This situation is particularly relevant in urban morphology, a field of study spreading from geography to architecture, focusing on the analysis of urban form and processes of its formation and transformation (Moudon 1997; Oliveira 2016; Kropf 2017). Unlike other strands of urban space research, urban morphology is concerned with a range of intermediate spatial scales—neighborhoods, blocks, streets, squares, plots, and buildings—and its study is meant to “identify the repeating patterns in the structure, formation and transformation of the built environment to help comprehend how the elements work together (...) to meet human needs and accommodate human culture” (Kropf 2014, p. 41). In this sense, urban morphology not only has contributed to the conceptualization of the spatial fabric of cities as a complex adaptive system and developed highly specialized methods for the study of its organizational structure but also has provided considerable insight on those intermediate spatial scales of central interest for urban designers, providing a valuable evidence base to contemporary urban design theory and practice.

Although traditionally geared towards qualitative approaches and “low-tech” methods, researchers in this field have recently shown an increasing interest for quantitative approaches

based on the use of GIS and related tools and the integration of digital cartographic products. Reviewing the most frequently used tools for urban form analysis (Table 1) highlights two trends. Initially, recalling earlier reference, most available tools are plug-ins associated with software packages such as *ArcMap* or *QGIS*. Furthermore, urban morphology is oriented around three fundamental spatial elements—plots, buildings, and streets (and their aggregations)—and their spatial formation. Accessibility to this data is underutilized due to the majority of current tools typically focusing on street network analysis only. The reason for this imbalance is the availability of data and tools at the time they were initially transferred between alternate disciplines and applied in urban morphology. The mathematical foundation for the majority of street network analysis is rooted in graph theory and physics of complex networks (Porta, Crucitti and Latora 2006), with methods applied to social and biological networks long before they were first applied to street formations. Furthermore, spatial networks—such as power grids, railways, or rivers, as well as streets—were among the first types of data to be available in a GIS environment. Even today, street network data remain the most abundant and widespread. For example, the crowd-mapping platform OpenStreetMap (OSM) reports an 83% network completeness worldwide (Barrington-Leigh and Millard-Ball 2017; and that is a single source of data), whereas building footprint data within the same platform are highly inconsistent, in terms of coverage, accuracy, and resolution (Brovelli and Zamboni 2018).

This imbalance is further reflected in “what” is being assessed, with network connectivity metrics being the only category of characters sufficiently covered. Following, Fleischmann et al. (2020a) there are five other categories (dimension, shape, spatial distribution, intensity, and diversity), which are severely underrepresented by available tools. Among tools for connectivity assessment, it can be further observed that the majority focus on the aspects of street centrality. Either building on the work “Multiple Centrality Assessment” by Porta, Crucitti and Latora (2006) or the school of “Space Syntax” (Hillier 1996), the first being based on a primal representation of the network of streets whereas the second being built on a dual approach. The former was initially implemented as a standalone tool (Gasser and Caillet 2013) and later reimplemented in different spatial analysis plug-ins/libraries. These include the *ArcMap* (and later *Rhino3D*) toolbox *Urban Network Analysis* (UNA) (Sevtsuk and Mekonnen 2012; Sevtsuk 2018), the open-source *Spatial Design Network Analysis* (Cooper and Chiaradia 2020), and recently as part of *OSMnx* (Boeing 2017) and in an expanded form in *momepy* (Fleischmann 2019). The latter was implemented by the Space Syntax group at the University College London in *depthmapX* (Turner 2001; depthmapX Development Team 2017) and *Place Syntax Toolkit* (PST; Stähle et al. 2005), both accessible as open-source softwares. In addition to centrality analysis, each tool can often measure other network-based variables (e.g., redundancy index in UNA, circuitry in *OSMnx*, and meshedness in *momepy*).

This leaves morphological elements such as buildings and plots, among others (i.e., blocks and street edges) considerably underrepresented.¹ Notable exceptions include *Metropolitan Form Analysis* toolbox (Amindarbari and Sevtsuk 2013) capturing seven metrics describing footprint and land-use pattern of a city, therefore operating on a metropolitan scale (of both grain and extent). *AwaP-IC* (Majic and Pafka 2019), a *QGIS* plug-in measuring two permeability-related metrics at the scale of blocks and buildings, *foot*, an R package describing gridded building footprints via a small number of metrics (Jochem and Tatem 2021), and *momepy* (Fleischmann 2019), which will be discussed separately.

It is to be remarked that the predominance of street-network and connectivity-based tools and the relative absence of tools for alternative morphological elements do not mean that researchers

Table 1. Overview of Available Open-Source Software for Urban Morphology

Tool	Focus	Type	Language	OS	License	Reference
<i>Urban Network Analysis</i> toolbox	Network analysis	<i>ArcMap</i> plug-in, <i>Rhino3D</i> plug-in	Python	Windows	CC BY-NC-SA 3.0	Sevtsuk (2018); Sevtsuk and Mekonnen (2012)
<i>Metropolitan Form Analysis</i> toolbox	City footprint, land use	<i>ArcMap</i> plug-in	Python	Windows	CC BY-ND 4.0	Amindarbari and Sevtsuk (2013)
<i>Multiple Centrality Assessment</i>	Network analysis	Standalone	Python	Cross-platform	unknown	Gasser and Caillet (2013)
<i>depthmapX</i>	Network analysis	Standalone, <i>QGIS</i> plug-in, CLI, R API	C++, R	Cross-platform	CC-GNU GPL	depthmapX Development Team (2017); Turner (2001)
<i>Place Syntax Tool</i>	Network analysis	<i>QGIS</i> plug-in, <i>MapInfo</i> plug-in	C++, Python	Cross-platform	GPL-3.0	Stähle Marcus and Karlström (2005)
<i>AwaP-IC</i>	Permeability	<i>QGIS</i> plug-in	Python	Cross-platform	GPL-3.0	Majic and Pafka (2019)
<i>Continuity in Street Networks</i>	Network analysis	<i>QGIS</i> plug-in, CLI	Python	Cross-platform	MIT	Tripathy et al. (2020)
<i>sDNA</i>	Network analysis	<i>QGIS</i> plug-in, <i>ArcGIS</i> plug-in, <i>AutoCAD</i> plug-in, CLI, Python API	C++, Python	Cross-platform/ Windows	GPL-3.0	Cooper and Chiaradia (2020)
<i>OSMnx</i>	Network analysis	Python package	Python	Cross-platform	MIT	Boeing (2017)
<i>momepy</i>	General purpose morphometrics	Python package	Python	Cross-platform	MIT	Fleischmann (2019)
<i>foot</i>	Building footprints	R package	R	Cross-platform	GPL-3.0	Jochem and Tatem (2021)

Abbreviations: CLI, command line interface; OS, operating system; sDNA, Spatial Design Network Analysis.

in this area have lacked engagement in measuring different elements of urban features (Araldi and Fusco 2019; Dibble et al. 2019). The issue is that these are hardly replicable and reproducible and come with no reusable tools or simply rely on ad hoc code. Dibble et al. (2019) for example, measured 207 morphological attributes at the level of blocks, buildings, and plots, as well as street networks in 45 urban neighborhoods (the author defines them as *sanctuary areas*). These measurements are all manually determined by Google satellite and Ordnance Survey data. An additional example is the work by Araldi and Fusco (2019), who measured 21 characters capturing the pedestrian point of view across the metropolitan area of Nice, France. Although this work does rely on an algorithmic approach using ad hoc code, this is largely undocumented in the published output, a decision which might be due to licensing restrictions. This situation is typical in partnerships between academia and the private sector or research carried out for the private sector alone.

One might argue that even in the absence of specific analytical tools, a traditional GIS environment (e.g., *QGIS*) using a GUI is more than sufficient for most applications, particularly considering the up-front investment (in time as well as learning effort) required for processing a method into a reusable code and lack of any academic reward for doing so. This overlooks several limitations affecting both research design and its applicability. The data science tools in Python are not restricted to geospatial analysis but offer a wide range of potential combinations from natural text processing to powerful artificial intelligence modelling, which can be intertwined with geographic data in scenarios that researchers require. The open-source code ensures transparency of methods as we can verify what each part of the process does, unlike in proprietary software, where the user has to believe the (often imprecise) documentation. Furthermore, code-based methods support reproducibility and replicability of the work by eliminating undocumented steps while avoiding the situation where methodological details and “rationales underpinning analytical decisions became obfuscated” (Boeing and Arribas-Bel 2021, p. 2) when relying on GUI.

Within the wider GIS research community, computational notebooks² and open-source software packages are increasingly seen as key solutions for research in the area and to be at the forefront of geographic open science (Boeing and Arribas-Bel 2021).

Python ecosystem for GDS

With the recent developments in both Python and R ecosystems for GDS, it is no longer the lack of fundamental building blocks that hinders the development and release of scientific software. Focusing on Python and vector-based analysis of urban form,³ we are witnessing a growing number of libraries and packages being released. These are quickly maturing to provide the required degree of stability, performance, and scalability for processing of large datasets.

Modern data science in Python is oriented around *pandas* (McKinney 2010), a package for tabular data analysis and manipulation. GDS follows this model with *GeoPandas* (Jordahl et al. 2021), extending *pandas* via support of geospatial features and operations, linking together various components of the ecosystem into a convenient form (Fig. 1). Its core depends on libraries written in C—*GEOS* (GEOS contributors 2021), *PROJ* (PROJ contributors 2021), and *GDAL* (GDAL/OGR contributors 2021). The geometry operations are handled by *shapely* (Gillies and others 2007), a Python interface to *GEOS*. Coordinate reference systems are managed by *pyproj* (Snow et al. 2021), which is interfacing *PROJ*. Capabilities for reading and writing geospatial data are using *fiona* (Gillies and others 2011), a module based on *GDAL*. The power of three

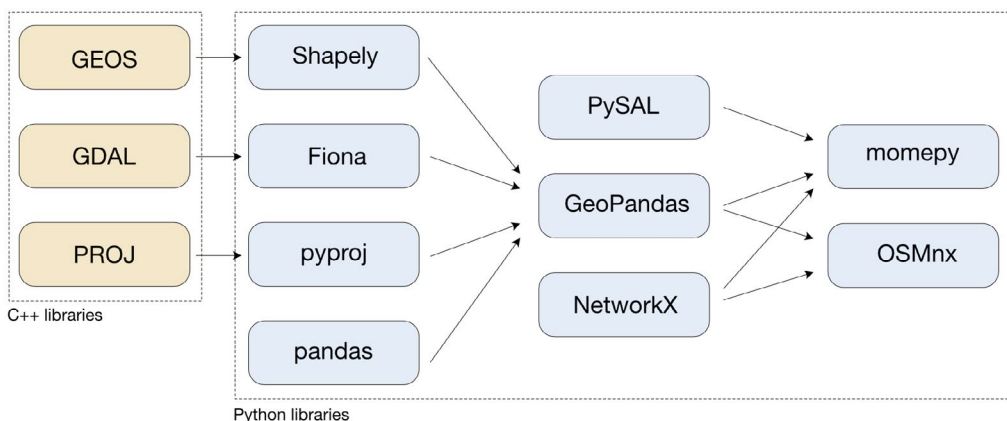


Figure 1. Schema of the software ecosystem supporting morphological research within this study.

performant C libraries (*GEOS*, *PROJ*, and *GDAL*) within a convenient pandas-like Python interface made *GeoPandas* a core tool for vector data manipulation.

Even prior to the initiation of the development of *GeoPandas* (June 2013), *PySAL* (Python Spatial Analysis Library) (Rey and Anselin 2007; Rey 2019) started being developed. Although originally independent, its relation to *GeoPandas* has become stronger over the years, and several *PySAL* modules now depend on *GeoPandas* and vice versa.⁴ *PySAL* brings a broad range of tools for spatial analysis, from implementation of spatial weights matrices to advanced spatial interpolation and multiscale geographically weighted regression models.

Parallel and largely disconnected has been the development of *NetworkX* (Hagberg, Schult and Swart 2008), a general purpose package for manipulation and analysis of networks, not necessarily spatial. However, in recent years, it has been incorporated into the several GDS applications, making it a fundamental component of spatial network analysis (typically streets).

These tools were crucial to the development of specialized software for morphological analysis. The first example is *OSMnx*, a library for modelling and analysing street networks obtained from OSM, depending on the capability of *NetworkX* and interfacing *GeoPandas*. Its ability to download and parse OSM directly from Python with a convenient interface opened new research possibilities democratizing access to data. However, the main analytical focus on *OSMnx* is still street network analysis and as such focuses on connectivity, similarly to the majority of existing tools outlined above.

The recent addition to the ecosystem is *momepy*, a library that builds on all *GeoPandas*, *PySAL*, *NetworkX*, and to a degree *OSMnx*, to develop an open repository of tools for morphometric assessment of built environments. This covers connectivity as well as the alternate five categories of measurable characters, which are reflected in different modules of the library. The close relationship to the rest of the ecosystem allows complex characterization of urban form within the frameworks of modern data science, which is reproducible as well as scalable.

The Python ecosystem centered around *momepy* has a potential to deliver great insights into the built-up patterns. Some initial applications include analysis of informal settlements (Mottelson and Venerandi 2020), prediction of building heights (Milojevic-Dupont et al. 2020), or classification of seashore streets in Portugal to evaluate local climate adaptation plans (Dal Cin et al. 2020). Combined with the parsing capability of *OSMnx*, we can now create open and

fully reproducible workflows analysing the structure and composition of urban form based on open data from OSM.

To prove this key point, we present below an application of such workflow. Specifically, to illustrate the potential of Python for urban morphometrics (i.e., the quantitative analysis of urban form). Different open-source Python tools are implemented to understand patterns of change in the structure and scale of urban form over time. To ensure reproducibility of the analysis, the whole method is delivered as a series of *Jupyter* notebooks executed within a containerized environment.

Case study: Alterations in structure and scale

Cities are in a state of continuous flow: They change in their economy, cultural landscape, societal norms, political discourse, and relationship with the natural environment. This global change is shaped over time through individuals and collective action. The form of cities also changes—buildings, plots, street fronts, blocks, and streets—each change at their own pace, moulded by new construction methods, technologies, resource availability, lifestyle preferences, and planning policy/theory.

And yet this change is not chaotic or random but reflects (and interacts with) accidents of history—key events of a social, political, cultural, and environmental nature—not only within the same city but also between cities that might be even considerably far apart in space.

For example, the technical innovations brought by the industrial revolution (i.e., prefabrication and serialization) have allowed forms and speed of building construction all over the world up to that point unthinkable; the freedom of movement associated to the shift to a car-dominated society has triggered the spread of low-density suburban lifestyles everywhere; the rapid population growth currently experienced in many developing regions of the world prompted a wave of unregulated informal urbanization characterized by extremely high density and compactness. Research in urban morphology has already started investigating the “laws” behind these recurrent urban form patterns through comparative studies.

Notably, Porta et al. (2014) have sought to show how different urban settlements all over the world all share what he calls a significant “alteration” in the scale of their street network, which corresponded to the macro-shift from a pre-industrial to an industrialized society. More specifically, they analyzed the intersection patterns of urban main streets from a pool of 100 case studies from 30 countries, characterized by diverse historic, sociocultural, and economic backgrounds divided into nine categories of urban form: ancient, medieval, renaissance, baroque, industrial, garden city, radiant city, new urbanism, and informal development. Albeit limited by sample size, the study uncovered the existence of a recurrent “400-m rule” in the intersection pattern of urban main streets in historic pre-industrial cities, a pattern that roughly doubled in all post-industrial samples up to the present day, with the notable exception of informal settlements, which followed the same “rule” observed in historical cases. Indeed, according to the author, this “alteration in scale” is on the one hand the product of concomitant socioeconomic shifts—the dominance of motorized vehicle, the engineerization of transportation planning, and the possibilities enabled by serial production—and, on the other, one of the prime culprits in the spread of “the unsustainable, car-dominated city of today” (p. 3,398).

The transformation of the network patterns over time was later studied by, Boeing (2020a) who carried out a similar study across the whole of the United States using *OSMnx* to quantify the link between configuration of the network and car ownership.

His big data approach confirmed Porta's manual measurement, showing that US cities and their street networks are on the journey from predominantly gridded configurations of pre-World War II (WW2) development through convoluted dendritic patterns peaking in the 1980s, back to more interconnected conventions in the 21st century. According to the author, the rise in car ownership was mirrored in the studied networks by a steady drift away from the connectivity and density of gridded patterns, and only very recently, with the awareness of the need of more walkable cities, this trend was partly reversed. In parallel, Barrington-Leigh and Millard-Ball (2020) carried out a similar analysis at a global level and developed a Street-Network Disconnectedness Index, finding that "in contrast to the corrective trend observed in the United States, where streets have become more connected since the late 20th century, we find that most of the world is building ever-more disconnected 'street-network sprawl'" (p. 1,941).

Broadly speaking, all these works share a similar conclusion: That certain spatial trends and "alterations," which are visible in the shape and configuration of the street networks in cities, are recurrent in cities regardless of geographic location as a result of high impact processes and events. Such hypothesis builds on a long strand of research in urban morphology, according to which changes in economy, technology, and culture drive phases of development, stagnation, and redevelopment across all constitutive elements of the urban form system (Conzen 1960; Feliciotti, Romice, and Porta 2017; Hallowell and Baran 2013).

But although this phenomenon has been studied from a quantitative perspective in relation to street networks only, it is reasonable to theorize that using Porta et al.'s (2014) own words, "similarly recursive spatial patterns within other elements of urban form" (p. 3,384).

To this regard, the proposed study builds further on this hypothesis by extending it to other components of urban form—namely, *buildings and tessellation cells*—and applies a rigorous quantitative approach to test whether patterns (and changes) similar to those observed for street networks across historical periods are also recognizable in structure and scale of these alternative elements of urban form.

Case study analysis method

For the purpose of this analysis, we defined six well-established historical periods to a degree replicating the subdivision in "urban design paradigms" adopted in the study by, Porta et al. (2014) although with a reduced number of classes: "pre-industrial," "industrial," "garden city," "modernist," "neo-traditional," and "informal." For each period, we sampled seven internally homogeneous areas, each defined by a 400-m buffer around a central location (Fig. 2), accumulating a total of 42 cases spread all over the world, covering different geographical and historical contexts (Fig. 3). Each of the 42 samples had to be internally homogeneous and highly representative of the relevant historical period.

Using *OSMnx*, we download building footprints and street networks from OSM as *GeoPandas* data structure as an input for *momepy*. To develop a more grounded understanding of the structure, we further generate morphological tessellation (Fig. 4), an analytical spatial unit derived from building footprints using Voronoi tessellation (Fleischmann et al. 2020b). Tessellation reflects the smallest spatial division, which at the same retains the information on contiguity, allowing identification of topological relationships between individual buildings. This property is used to measure various morphometric characters reflecting the spatial distribution of building footprints.



Figure 2. Examples of case studies (one per historical period) reflecting the variety of urban form patterns within set buffers.

Capturing the change in the structure of urban form is a complex task as there are endless possibilities on what could be analyzed. Within this work, we use a selection of 12 morphometric characters (Table 2) spanning across all six categories identified by Fleischmann et al. (2020a) and different spatial elements, that is, street networks, building footprints, and morphological tessellation. The set ranges from simple characters (i.e., *area of building footprint*), as well as more complex metrics reflecting the relationship between individual elements (i.e., *adjacency of buildings*) or capturing characters of the street profile (defined by a combination of streets and buildings). We treat this selection as a sample to illustrate the workflow while producing valuable insights. This can be expanded in a potential full-scale study. Statistical distributions of measured values within each period are analyzed using a Kruskal–Wallis one-way analysis of variance (Kruskal and Wallis 1952) and a pairwise Mann–Whitney *U*-test (Mann-Whitney 1947; following SciPy implementation [Virtanen et al. 2020]) to empirically test the hypothesis of change.

The assumption of the analysis is that results will present changes in the built-up patterns similar to what has been previously observed by, Porta et al. (2014), Boeing (2020a) and Barrington-Leigh and Millard-Ball (2020)—a significant alteration in the transition from pre-industrial to post-war development, with a slow tendency of the return to the pre-industrial patterns in particular contexts (i.e., informal settlements and early 21st century developments). The extent of such a change is not known yet.

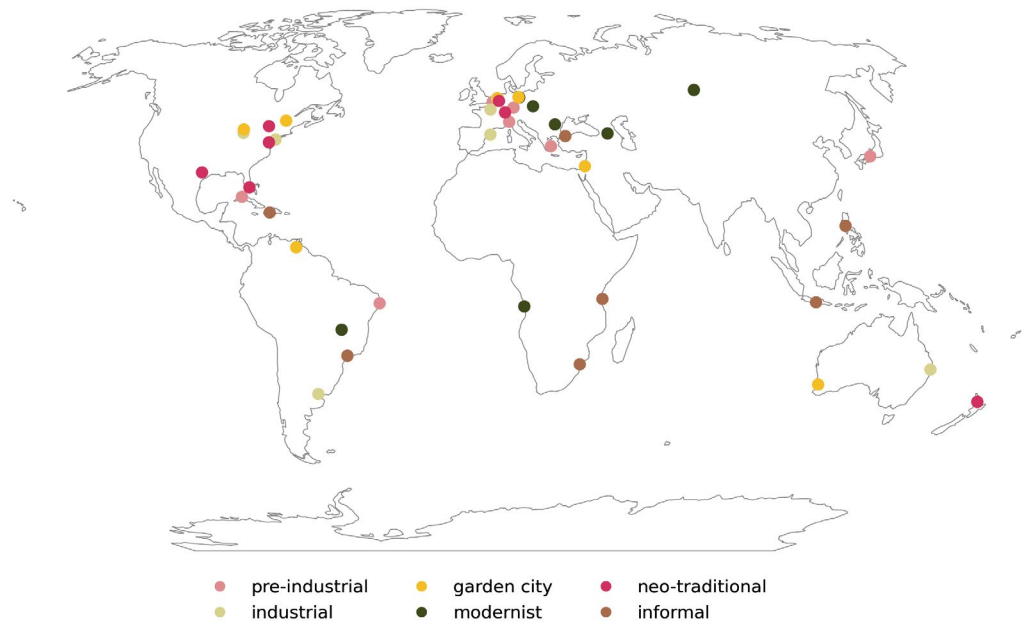


Figure 3. Selection of case studies and their allocation to historical periods. The tabular form is available in the code and data repository.

The whole method is encapsulated in three computational *Jupyter* notebooks covering each step from the beginning to the end, with the only external input being the list of case studies with the point of origin. We download and process data for a specific timestamp using *OSMnx* and *GeoPandas*, generate tessellation, and measure morphometric characters using *momepy*, generating statistical outputs and figures using *SciPy* (Virtanen et al. 2020) and *seaborn* (Waskom and team 2020). This is wholly processed within a reproducible containerized environment. The potential expansion of the list of case studies then depends only on their identification and data availability.

Results

Table 3 presents a summary statistic (median and interquartile range) of each morphometric character by historical period. The tendencies shown by the data are generally in line with the previous finding by Boeing (2020a) and, Porta et al. (2014) with the highest deviations occurring during the post-WW2 era of the modernist planning paradigm and slowly turning back afterwards. Notably, informal settlements tend to be structurally similar to pre-industrial and industrial development than to any other period, suggesting that the lack of the influence of planning and especially transportation technology (i.e., personal vehicles) have a strong tendency to generate similar walkable dependable patterns.

The apparent alteration in scale, observed by Porta et al. (2014) as a change in the distance between the main streets, is to a large degree present in the small-scale data also. The median *area of tessellation cells* and related *building areas* and *neighbor distance between buildings* all show very similar tendencies (Fig. 5), this being the change of the scale between industrial



Figure 4. Illustration of the behavior of morphological tessellation on the sample set of case studies.

and garden city periods, with a major peak during modernism. The change in scale is extensive. Although the median area of tessellation cells more than triples between the industrial era and garden city movement, it is more than 19× larger during modernism than in the industrial period. Although we can observe how neo-traditional development values are comparatively less extreme than in the modernist period, they are still almost triple in size compared with the (pre-)industrial era, despite their programmatic aim to develop patterns close to pre-WW2 fabric. Informal settlements are the most compact in this comparison, at roughly 65% the size of historical development. Although the difference in *mean neighbor distance* is not so radical, the tendency is the same. Simultaneously, the only radical deviation in building size is present in modernism (more than 3.5× larger in comparison with alternative forms) and in informal developments (almost half in comparison with alternative forms), indicating that the primary alteration change is in the pattern, rather than in the range—in other words, we have simply started building our houses further away from one another.

Such interpretation is further supported by the results of openness of a street profile, width of a street profile, and covered area ratio of tessellation cells. Openness refers to the presence of buildings along the street; that is, wider gaps between buildings will lead to a higher openness. As shown in Fig. 6a, the previously observed tendency is directly reflected in the change of openness. The openness of informal settlements is on par with (pre-)industrial cases, not deviated as the scale issue tends to be. The buildings are not further away only along the street but also across the street, leading to wider and more saturated streetscapes, illustrated in Fig. 6c. Lastly, the

Table 2. Morphometric Characters, Their Classification, and Definition

Character	Category	Definition
Area of a tessellation cell (m)	Dimension	Area of a tessellation cell polygon
Covered area ratio of tessellation cell	Intensity	Proportion of a tessellation cell covered by a building footprint
Area of a building footprint (m)	Dimension	Area of a building footprint polygon
Length of a perimeter wall (m)	Dimension	Length of an outer perimeter wall of a joined structure form of contiguous buildings
Adjacency of buildings	Spatial distribution	Ratio of joined built-up structures and buildings within set area
Mean distance between buildings (m)	Spatial distribution	Mean distance between buildings on adjacent tessellation cells
Length of a street segment (m)	Dimension	Length of a street segment line
Linearity of a street segment	Shape	Ratio of Euclidean distance between the first and the last point of a line and its length
Width of a street profile (m)	Dimension	Mean length of a street profile sections along street segment
Width deviation of a street profile (m)	Diversity	Standard deviation of length of street profile sections
Openness of a street profile	Spatial distribution	Ratio of street profile sections intersecting building
Meshedness of a street network	Connectivity	Ratio of the number of faces in the network to the maximum possible number of loops in an equivalent network with the same number nodes

Note: For implementation details, please refer to the documentation of *momepy* v0.4.2.

covered area ratio directly reflects the area of tessellation cells and area of buildings measured above and again placing the informal settlements very close to historical development.

An additional effect of the change of scale is reflected in the significant difference between the *length of a perceived contiguous perimeter wall* of (pre-)industrial development and that of the more recent origin. Similar is observed in building adjacency (Fig. 7). Both conclude that we have moved from building cities composed of adjacent buildings towards solitary development.

The final group of characters presents that not all aspects of built form follow the same trajectory. *Meshedness* (Fig. 8a), the proxy of street network connectivity, is highest in industrial cases, driven by a rigid grid, conventional of that era. The *linearity of a street segment* also reflects the effect of the grid while showing that informal settlements, often similar to historical developments, have the second most convoluted street network (just behind modernism). The effect of informality is also reflected in *the deviation of street profile width*, being the largest in this group of case studies.

The interpretation of median values and box plots leading to a conclusion that there is a significant change of the built-up patterns between different periods is further supported by the results of Kruskal–Wallis test, which indicate that the distributions of morphometric values

Table 3. Summary (Median [Interquartile Range]) Values of Morphometric Characters Per Historical Period

Character	Pre-industrial	Industrial	Garden city	Modernist	Neo-traditional	Informal
Area of a tessellation cell (m)	177.03 (262.39)	171.4 (276.77)	589.16 (642.93)	3,377.37 (3,742.74)	504.6 (639.15)	109.79 (122.36)
Covered area ratio	0.62 (0.28)	0.53 (0.2)	0.22 (0.15)	0.16 (0.1)	0.27 (0.17)	0.51 (0.27)
Area of a building footprint (m)	98.74 (136.78)	80.32 (113.72)	141.49 (136.75)	545.36 (592.86)	143.27 (137.69)	52.87 (56.23)
Length of a perimeter wall (m)	282.85 (351.53)	335.94 (399.06)	60.26 (42.09)	169.43 (101.07)	68.33 (39.54)	32.48 (22.24)
Building adjacency	0.2 (0.29)	0.2 (0.2)	0.84 (0.34)	0.92 (0.42)	1.0 (0.44)	1.0 (0.05)
Mean neighbor distance between buildings (m)	3.44 (3.57)	6.56 (4.4)	14.67 (10.45)	32.56 (19.4)	12.94 (9.54)	3.0 (3.32)
Length of a street segment (m)	91.81 (71.42)	136.51 (51.1)	154.63 (121.79)	196.23 (169.47)	131.6 (100.27)	133.96 (124.83)
Linearity of a street segment	1.0 (0.01)	1.0 (0.0)	1.0 (0.04)	1.0 (0.17)	1.0 (0.03)	0.99 (0.06)
Width of a street profile (m)	10.16 (5.81)	15.16 (10.96)	29.12 (6.93)	33.25 (13.07)	24.7 (9.86)	12.44 (6.86)
Width deviation of a street profile (m)	3.16 (2.44)	2.31 (1.88)	3.59 (1.88)	3.3 (2.82)	2.87 (1.99)	4.75 (1.74)
Openness of a street profile	0.15 (0.13)	0.18 (0.13)	0.38 (0.23)	0.61 (0.22)	0.35 (0.23)	0.13 (0.14)
Meshedness of a street network	0.17 (0.07)	0.26 (0.05)	0.15 (0.09)	0.14 (0.05)	0.2 (0.06)	0.12 (0.08)

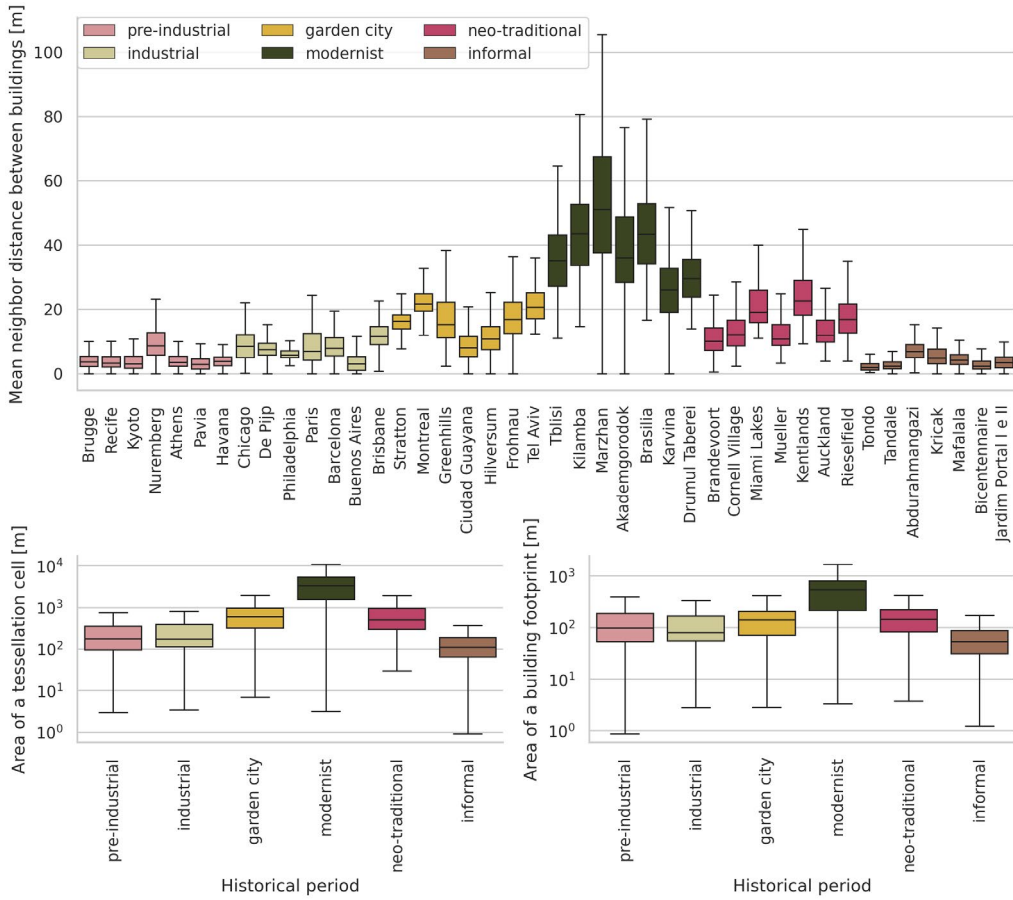


Figure 5. Box plots of the (a) mean neighbor distance between buildings, (b) area of tessellation cells grouped by historical period, and (c) area of building footprints grouped by historical period.

obtained from samples from different historical periods cannot be considered the same ($p < .05$). The consequent analysis using Mann–Whitney U -test comparing distributions of every pair of periods for each measurable character indicates that with three exceptions, the values obtained from any two pairs of periods significantly differ ($p < .05$). The only distributions that cannot be considered different are *building adjacency* of garden city and modernist periods ($p = .4$), *linearity of a street segment* of garden city and modernist periods ($p = .15$), and *width deviation of a street profile* of pre-industrial and modernist periods ($p = .17$). Complete results are reported in the complimentary *Jupyter* notebook.

Discussion

The alteration in patterns observed in the 42 sample cases highlights once more that profound transformations have occurred over time in the way we construct the built environment. These changes are manifested in every aspect of urban form, both in terms of structures and scale, mainly confirming previous findings by Porta et al. (2014) and Boeing (2020a) regarding street

Geographical Analysis

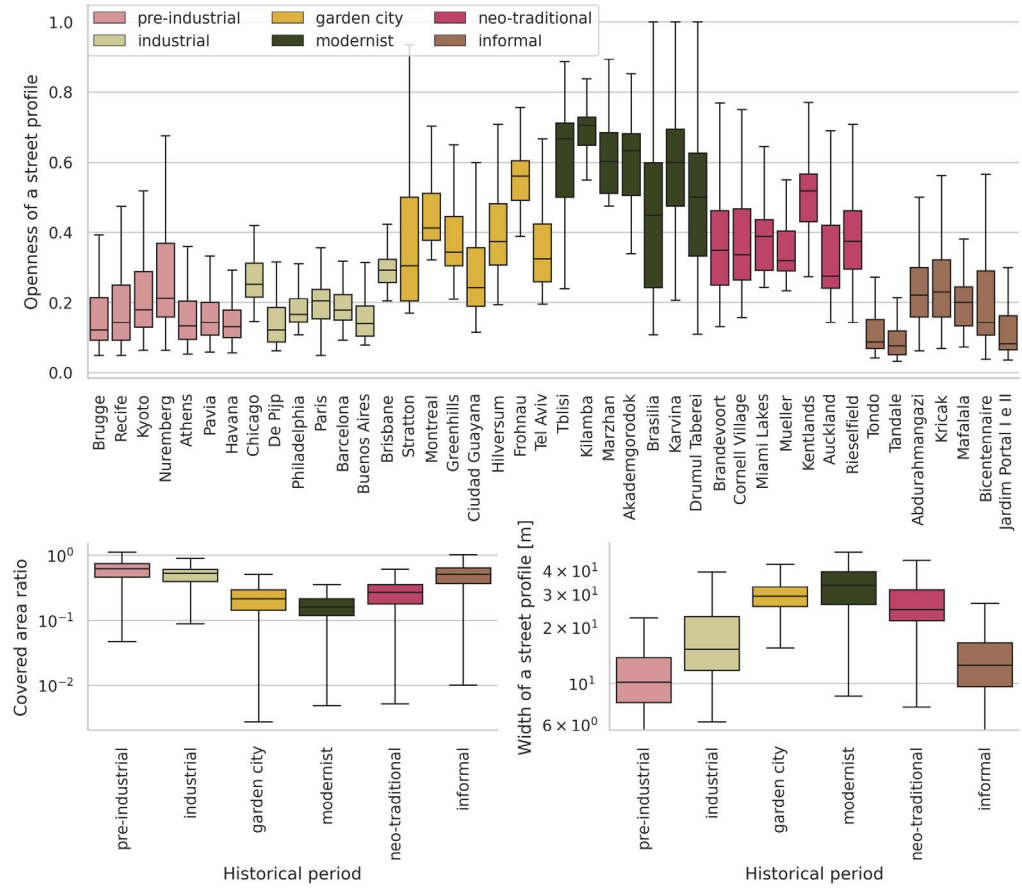


Figure 6. Box plots of the (a) openness of a street profile, (b) covered area ratio grouped by historical period, and (c) width of a street profile grouped by historical period.

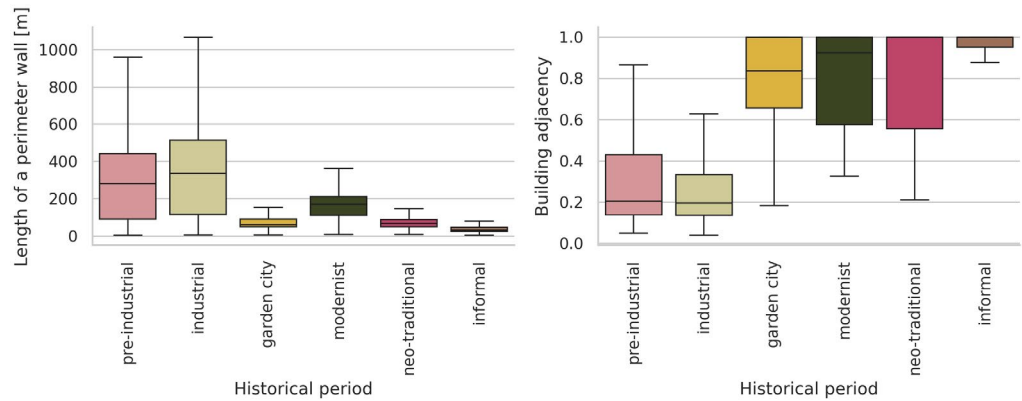


Figure 7. Box plots of (a) length of a perimeter wall and (b) building adjacency grouped by historical period.

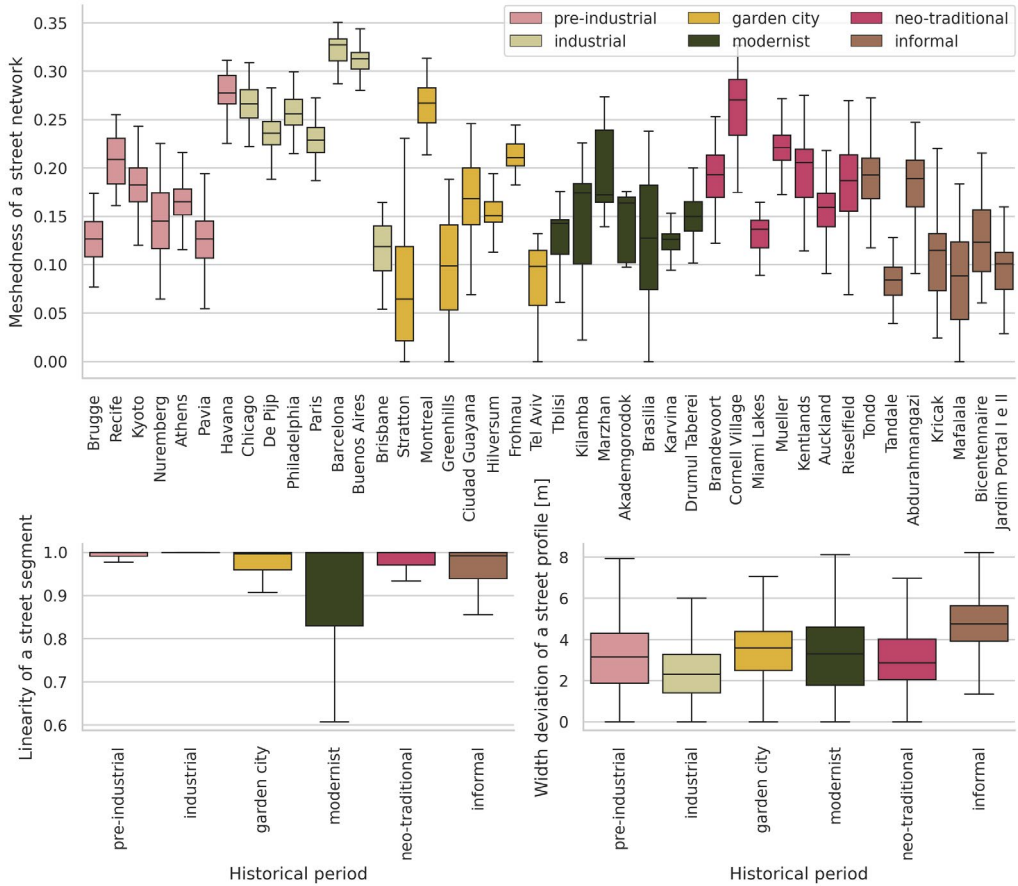


Figure 8. Box plots of the (a) mean meshedness of a street network, (b) mean linearity of street segments grouped by historical period, and (c) mean width deviation of a street profile grouped by historical period.

networks. Although the explanation of such (often radical) differences is left to further research, it is clear how changes influencing the structure of a particular morphological complex (i.e., network of streets) are intertwined to changes in alternative complexes due to the tangled interdependencies of the urban form system. This is undoubtedly a serious pursuit, which will allow the deeper uncovering of structural tendencies in the environments around us and ultimately advance our current understanding of different performances of cities. Although the selection of case studies used in this example aims to be representative, a wider set would provide more robust results. A similar situation applies to the set of morphometric characters used, where a larger number of employed characters would result in a more comprehensive picture. However, the primary purpose of the case is to illustrate the abilities of the Python ecosystem in the study of urban form, and as such, it is less affected by these limitations.

From the perspective of the Python GDS ecosystem and its ability to support and deliver an analysis of urban form, this case study illustrates its achieved maturity and reliability. Every step of the procedure is fully contained and processed within *Jupyter* notebooks without ever requiring the analyst to switch between environments. Furthermore, as the whole method is written

only in Python, it reduces the burden on researchers to learn a broad range of tools to process different steps of the analysis.

The proposed method is fully replicable, reproducible, and expandable because it only requires open data from OSM (moreover linked to a specific point in time) and relies on an entirely automated workflow. It is replicable because running the code within the provided Docker container (lightweight executable environment) should always lead to the exact same results. It is reproducible as the code can be simply applied to different case studies of choice or can be run on different data sources capturing urban form besides OSM, potentially exploiting resources provided by open data portals at municipal, national, and global levels. Finally, it is expandable insofar the code can be optimized and extended to include further metrics in the analysis—either existing or created from scratch by other researchers. On the other hand, it also requires a basic knowledge of Python to use the workflow, which may be initially seen as limiting.

Switching to a code-based analysis may be associated with a steep learning curve. However, not everyone needs to reach the developer level as the data science ecosystem aims to provide a middle ground user level. That is a bit like Lego—the researcher learns how to put pieces together and then find pieces they need to *build a house*. Other researchers already packaged complex scripts into straightforward functions (like in the case of *PySAL*). Furthermore, truly reproducible workflows should *just run*, with minimal knowledge needed, as is illustrated by the presented case study. The user only needs to prepare an environment (either manually or via Docker image) and optionally edit the table listing individual cases. As the coding proficiency advances, it is easier to build reproducible research and share it with the community, which results in increasing impact.

The more this mindset becomes widespread among researchers, the more highly specialized tools addressing alternative aspects of urban form analysis will become available. In this way, the addition of newly developed tools that contribute to the existing ecosystem becomes a standard practice. The new methods can result in dedicated packages (where the scope of the work does not fit in any of the existing) while retaining compatibility with the ecosystem, allowing direct exchange of data and consistency of workflows. In other cases, they can become contributions to existing packages in a similar fashion in which *PySAL* is being developed (e.g., implementations of work proposed by Jiang 2013; Arribas-Bel, Garcia-López and Viladecans-Marsal 2019; Wolf, Knaap and Rey 2019). The majority of the infrastructural work overlaps between different applications. Thus, we should not spend time reimplementing it over and over again, as there is no requirement to constantly reinvent the wheel.

Running a morphometric assessment of the sort just presented, scaled from small pieces of urban tissue to broad metropolitan areas, can become computationally demanding if not overwhelming for traditional GIS environments. With more steps involved, the point-and-click workflow becomes obfuscated and its processing toolkits inefficient. Although *GeoPandas* performs all operations as a single-core process, Python's ecosystem can support its parallelization and eventually out-of-core computation for larger than memory data. It, alongside code optimization, allows researchers to handle vast amounts of data and very demanding computations. Depending on *Dask* (Rocklin 2015; a library for scalable Python computation) minimizes the requirements to learn additional frameworks (as *Apache Sedona* [Yu, Zhang, and Sarwat 2019] depending on *Apache Spark*) due to close relation to pandas ecosystem, API, and data manipulation logic. The work on the scalability of *GeoPandas*-based computation based on *Dask* is already under development both directly as a *dask-geopandas* extension (Signell, Van den Bossche and Fleischmann 2021) and indirectly leveraging *Dask* directly in *momepy*. It can be expected that the support will

evolve into a seamless implementation, which in turn allows straightforward scalability of urban morphometrics to regional or national extents and beyond.

Disregarding the technical aspects, relying on open-source data and operating in an open-source environment dramatically widen opportunities for a fully open research agenda in urban morphology and open new possibilities in terms of compatibility and cooperation. New quantitative methods can derive rich data, enabling explorations of the applicability of the fourth paradigm of science in urban morphology. Even the ability to think about fully data-driven research is new to the field and would not be possible without the inclusion of GDS. At the same time, we can strengthen links between traditional and quantitative methods, with the former providing a theoretical component and the latter descriptive one. Such mixed methods can link the detail and profound insights of traditional urban morphology with descriptive power and scalability of urban morphometrics. Both of these options are becoming more prominent and will require time and critical assessment to properly mature. Nevertheless, both already enrich the portfolio of urban morphologists, making it more open and more reproducible. The open research paradigm, based on open platforms and transparent community-led governance, has the potential to democratize science and remove unnecessary friction caused by the lack of cooperation between research groups while bringing additional transparency to research methods and outputs.

Data availability statement

The repository, containing reproducible code in *Jupyter* notebooks, data, and a complete set of figures, is available at <https://github.com/martinfleis/evolution-gean>, <https://doi.org/10.5281/zenodo.5034747>. The research has been executed within a Docker container `martinfleis/momepy:0.4.2`.

Notes

- 1 It is to be noted that although UNA and PST can link other features to street segments (i.e., buildings and plots), the gist of the analysis is still network analysis.
- 2 Boeing and Arribas-Bel (2021) describe a computational notebook as “a computer file that contains code, output, images, and narrative text woven together. Notebooks allow users to consolidate their analytics workflows, blending code, documentation, and results into a single reproducible and distributable file.” (p. 1).
- 3 We do not address remote sensing and rasters in this article due to their limited scope of application in urban morphology at the time of writing. However, the current underutilization of raster data is likely to change, opening additional avenues of research.
- 4 *GeoPandas* is using *PySAL*’s *mapclassify* library in its choropleth mapping.

References

- Amindarbari, R., and A. Sevtsuk. (2013). “Measuring Growth and Change in Metropolitan Form.” Urban Affairs Association Conference, San Francisco, 2013.
- Araldi, A., and G. Fusco. (2019). “From the Street to the Metropolitan Region: Pedestrian Perspective in Urban Fabric Analysis.” *Environment and Planning B: Urban Analytics and City Science* 46(7), 1243–63. <https://doi.org/10.1177/2399808319832612>.
- Arribas-Bel, D., M.-À. Garcia-López, and E. Viladecans-Marsal. (2019). “Building (s and) Cities: Delineating Urban Areas with a Machine Learning Algorithm.” *Journal of Urban Economics*. 103217. <https://doi.org/10.1016/j.jue.2019.103217>.
- Barrington-Leigh, C., and A. Millard-Ball. (2017). “The World’s Open-Source Street Map is More Than 80% Complete.” *PLoS One* 12(8), 1–20. <https://doi.org/10.1371/journal.pone.0180698>.

- Barrington-Leigh, C., and A. Millard-Ball. (2020). “Global Trends Toward Urban Street-Network Sprawl.” *Proceedings of the National Academy of Sciences* 117(4), 1941–50. <https://doi.org/10.1073/pnas.1905232116>.
- Bivand, R. S. (2020). “Progress in the R Ecosystem for Representing and Handling Spatial Data.” *Journal of Geographical Systems* Springer 1–32. <https://doi.org/10.1007/s10109-020-00336-0>.
- Boeing, G. (2017). “OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks.” *Computers, Environment and Urban Systems* 65, 126–39. <https://doi.org/10.1016/j.compenvurbsys.2017.05.004>.
- Boeing, G. (2018). “A Multi-Scale Analysis of 27,000 Urban Street Networks: Every US City, Town, Urbanized Area, and Zillow Neighborhood.” *Environment and Planning B: Urban Analytics and City Science* 219(4), 239980831878459. <https://doi.org/10.1177/2399808318784595>.
- Boeing, G. (2020a). “Off the Grid... and Back Again? The Recent Evolution of American Street Network Planning and Design.” *Journal of the American Planning Association*, 87(1), 123–137. <https://doi.org/10.2139/ssrn.3665445>.
- Boeing, G. (2020b). “The Right Tools for the Job: The Case for Spatial Science Tool-Building.” *Transactions in GIS* 24(5), 1299–314. <https://doi.org/10.1111/tgis.12678>.
- Boeing, G., and D. Arribas-Bel. (2021). “GIS and Computational Notebooks.” In *The Geographic Information Science & Technology Body of Knowledge*, edited by J. P. Wilson. Ithaca, NY: University Consortium for Geographic Information Science.
- Brovelli, M. A., and G. Zamboni. (2018). “A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints.” *ISPRS International Journal of Geo-Information* 7(8), 289. <https://doi.org/10.3390/ijgi7080289>.
- Conzen, M. (1960). *Alnwick, Northumberland: A Study in Town-Plan Analysis*. London: George Philip & Son, Available at: <http://www.jstor.org/stable/pdf/621094.pdf>.
- Cooper, C. H., and A. J. Chiaradia. (2020). “sDNA: 3-d Spatial Network Analysis for GIS, CAD, Command Line & Python.” *SoftwareX* 12. 100525. <https://doi.org/10.1016/j.softx.2020.100525>.
- Dal Cin, F., M. Fleischmann, O. Romice, and J. P. Costa. (2020). “Climate Adaptation Plans in the Context of Coastal Settlements: The Case of Portugal.” *Sustainability* 12(20), 19. <https://doi.org/10.3390/su12208559>.
- depthmapX development team. (2017). *DepthmapX*. Available at: <https://github.com/SpaceGroupUCL/depthmapX/>.
- Derudder, B., and M. van Meeteren. (2019). “Engaging with “Urban Science”.” *Urban Geography* 40(4), 555–64. <https://doi.org/10.1016/j.softx.2020.100525>.
- Dibble, J., A. Prelorendjos, O. Romice, M. Zanella, E. Strano, M. Pagel, and S. Porta. (2019). “On the Origin of Spaces: Morphometric Foundations of Urban Form Evolution.” *Environment and Planning B: Urban Analytics and City Science* 46(4), 707–30. <https://doi.org/10.1177/2399808317725075>.
- Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini. (2012). “Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services.” *Remote sensing of Environment* 120, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>.
- Feliciotti, A., O. Romice, and S. Porta. (2017). “Urban Regeneration, Masterplans and Resilience: The Case of Gorbals, Glasgow.” *Urban Morphology* 21(1), 61–79.
- Fleischmann, M. (2019). “momepy: Urban Morphology Measuring Toolkit.” *Journal of Open Source Software* 4(43), 1807. <https://doi.org/10.21105/joss.01807>.
- Fleischmann, M., A. Feliciotti, O. Romice, and S. Porta. (2020a). “Morphological Tessellation as a Way of Partitioning Space: Improving Consistency in Urban Morphology at the Plot Scale.” *Computers, Environment and Urban Systems* 80, 101441. <https://doi.org/10.1016/j.compenvurbsys.2019.101441>.
- Fleischmann, M., O. Romice, and S. Porta. (2020). “Measuring Urban Form: Overcoming Terminological Inconsistencies for a Quantitative and Comprehensive Morphologic Analysis of Cities.” *Environment and Planning B: Urban Analytics and City Science* 239980832091044. <https://doi.org/10.1177/2399808320910444>.
- Gahegan, M. (2020). “Fourth Paradigm GIScience? Prospects for Automated Discovery and Explanation from Data.” *International Journal of Geographical Information Science* 34(1), 1–21. <https://doi.org/10.1080/13658816.2019.1652304>.

- Gasser, L., and J. Caillet. (2013). mca. GitHub Repository. GitHub. Available at: <https://github.com/jcaillet/mca>.
- GDAL/OGR contributors. (2021). GDAL/OGR Geospatial Data Abstraction Software Library. Manual. Available at: <https://gdal.org>.
- GEOS contributors. (2021). GEOS Geometry Engine, Open Source. Manual. Available at: <https://trac.osgeo.org/geos/>.
- Gillies, S. (2007). Shapely: Manipulation and Analysis of Geometric Objects. toblery.org. Available at: <https://github.com/Toblerity/Shapely>.
- Gillies, S., and others. (2011). Fiona is OGR's Neat, Nimble, No-Nonsense API. Available at: <https://github.com/Toblerity/Fiona>.
- Hagberg, A. A., D. A. Schult, and P. J. Swart. (2008). "Exploring Network Structure, Dynamics, and Function using NetworkX." In Proceedings of the 7th Python in Science Conference, 11–15, edited by G. el Varoquaux, T. Vaught, and J. Millman. Pasadena, CA, 1 January 2008.
- Hallowell, G. D., and P. K. Baran. (2013). "Suburban Change: A Time Series Approach to Measuring Form and Spatial Configuration." *The Journal of Space Syntax* 4(1), 74–91.
- Harris, R., D. O'Sullivan, M. Gahegan, M. Charlton, L. Comber, P. Longley, C. Brunson, N. Malleon, A. Heppenstall, A. Singleton, D. Arribas-Bel, and A. Evans. (2017). "More Bark Than Bytes? Reflections on 21+ Years of Geocomputation." *Environment and Planning B: Urban Analytics and City Science* 44(4), 598–617. <https://doi.org/10.1177/2399808317710132>.
- Hey, A. J., S. Tansley, and K. M. Tolle (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Redmond, WA: Microsoft Research.
- Hillier, B. (1996). Space is the Machine: A Configurational Theory of Architecture. Cambridge: Cambridge University Press.
- Hoyer, S., and J. Hamman. (2017). "xarray: N-D labeled Arrays and Datasets in Python." *Journal of Open Research Software* 5(1), 1–6. <https://doi.org/10.5334/jors.148>.
- Jiang, B. (2013). "Head/Tail Breaks: A New Classification Scheme for Data with a Heavy-Tailed Distribution." *The Professional Geographer* 65(3), 482–94. <https://doi.org/10.1080/00330124.2012.700499>.
- Jochem, W. C., D. R. Leasure, O. Pannell, H. R. Chamberlain, P. Jones, and A. J. Tatem. (2020). "Classifying Settlement Types from Multi-Scale Spatial Patterns of Building Footprints." *Environment and Planning B: Urban Analytics and City Science* 48(5), 1161–1179. <https://doi.org/10.1177/2399808320921208>.
- Jochem, W. C., and A. J. Tatem. (2021). "Tools for Mapping Multi-Scale Settlement Patterns of Building Footprints: An Introduction to the R Package Foot." *PLOS ONE* 16(2), e0247535. <https://doi.org/10.1371/journal.pone.0247535>.
- Jordahl, K., J. van den Bossche, M. Fleischmann, J. McBride, J. Wasserman, J. Gerard, A. Garcia Badaracco, A. D. Snow, J. Tratner, M. Perry, C. Farmer, G. A. Hjelle, M. Cochran, S. Gillies, L. Culbertson, M. Bartos, G. Caria, N. Eubank, S. Veera, S. Rey, M. Albert, A. Bilogur, B. Ward, C. Ren, D. Arribas-Bel, F. Lavin, L. Wasser, L. J. Wolf, and M. Journois (2021). Geopandas/Geopandas: V0.9.0. Zenodo. <https://doi.org/10.5281/zenodo.4569086>.
- Kandt, J., and M. Batty. (2021). "Smart Cities, Big Data and Urban Policy: Towards Urban Analytics for the Long Run." *Cities* 109, 102992. <https://doi.org/10.1016/j.cities.2020.102992>.
- Kropf, K. (2014). "Ambiguity in the definition of built form." *Urban Morphology* 18(1), 41–57.
- Kropf, K. (2017). The Handbook of Urban Morphology. Chichester: John Wiley & Sons. Available at: <http://cds.cern.ch/record/2316422>.
- Kruskal, W. H., and W. W. Wallis. (1952). "Use of Ranks in One-Criterion Variance Analysis." *Journal of the American Statistical Association* 47(260), 583–621.
- Kwan, M.-P., and T. Schwanen. (2009). "Quantitative Revolution 2: The Critical (Re)Turn." *The Professional Geographer* 61(3), 283–91. <https://doi.org/10.1080/00330120902931903>.
- Majic, I., and E. Pafka. (2019). "AwaP-IC—An Open-Source GIS Tool for Measuring Walkable Access." *Urban Science* 3(2), 48. <https://doi.org/10.3390/urbansci3020048>.
- Mann, H. B., and D. R. Whitney. (1947). "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other." *The Annals of Mathematical Statistics* 18(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>.

- McKinney, W. (2010). "Data Structures for Statistical Computing in Python." In Proceedings of the 9th Python in Science Conference, 56–61, edited by S. van der Walt and Jarrod Millman. <https://doi.org/10.25080/Majora-92bf1922-00a>.
- Milojevic-Dupont, N., N. Hans, L. H. Kaack, M. Zumwald, F. Andrieux, D. de Barros Soares, S. Lohrey, P.-P. Pichler, and F. Creutzig. (2020). "Learning from Urban Form to Predict Building Heights." *PLOS ONE* 15(12), e0242010. <https://doi.org/10.1371/journal.pone.0242010>.
- Mottelson, J., and A. Venerandi. (2020). "A Fine-Grain Multi-Indicator Analysis of the Urban Form of Five Informal Settlements in East Africa." *Urban Science* 4(3), 31. <https://doi.org/10.3390/urbansci4030031>.
- Moudon, A. V. (1997). "Urban Morphology as an Emerging Interdisciplinary Field." *Urban Morphology* 1(1), 3–10.
- Mydlarz, C., J. Salamon, and J. P. Bello. (2017). "The Implementation of Low-Cost Urban Acoustic Monitoring Devices." *Applied Acoustics* 117, 207–218. <https://doi.org/10.1016/j.apacoust.2016.06.010>.
- Oliveira, V. (2016). *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. Cham: Springer International Publishing.
- Oliveira, V., and V. Medeiros. (2016). "Morpho: Combining Morphological Measures." *Environment and Planning B: Planning and Design* 43(5), 805–25. <https://doi.org/10.1177/0265813515596529>.
- Pebesma, E. J. (2018). "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10(1), 439. <https://doi.org/10.32614/RJ-2018-009>.
- Poorthuis, A., and M. Zook. Poorthuis, A., and M. Zook. (2020). "Being Smarter about Space: Drawing Lessons from Spatial Science." *Annals of the American Association of Geographers* 110(2), 349–59. <https://doi.org/10.1080/24694452.2019.1674630>.
- Porta, S., P. Crucitti, and V. Latora. (2006). "The Network Analysis of Urban Streets: A Primal Approach." *Environment and Planning B: Planning and Design* 33(5), 705–25. <https://doi.org/10.1068/b32045>.
- Porta, S., V. Latora, and E. Strano. (2010). "Networks in Urban Design. Six Years of Research in Multiple Centrality Assessment." In *Network Science*, 107–129. London: Springer London. Available at: http://link.springer.com/10.1007/978-1-84996-396-1_6.
- Porta, S., O. Romice, J. A. Maxwell, P. Russell, and D. Baird. (2014). "Alterations in Scale: Patterns of Change in Main Street Networks Across Time and Space." *Urban Studies* 51(16), 3383–400. <https://doi.org/10.1177/0042098013519833>.
- PROJ contributors. (2021). PROJ Coordinate Transformation Software Library. Manual. Available at: <https://proj.org/>.
- Rey, S. J. (2019). "PySAL: The First 10 Years." *Spatial Economic Analysis* 14(3), 273–82. <https://doi.org/10.1080/17421772.2019.1593495>.
- Rey, S. J., and L. Anselin. (2007). "PySAL: A Python Library of Spatial Analytical Methods." *The Review of Regional Studies* 37(1), 5–27.
- Rocklin, M. (2015). "Dask: Parallel Computation with Blocked Algorithms and Task Scheduling." In Proceedings of the 14th Python in Science Conference, Austin, TX: SciPy, 130–6, edited by K. Huff and J. Bergstra.
- Sevtsuk, A. (2018). *Urban Network Analysis for Rhinoceros 3D: Tools for Modeling Pedestrian and Bicycle Trips in Cities*. Boston, MA: City Form Lab.
- Sevtsuk, A., and M. Mekonnen. (2012). "Urban Network Analysis. A New Toolbox for ArcGIS." *Revue internationale de géomatique* 22(2), 287–305. <https://doi.org/10.3166/ig.22.287-305>.
- Signell, J., J. Van den Bossche, and M. Fleischmann. (2021). Dask-Geopandas. Available at: <https://github.com/geopandas/dask-geopandas/>.
- Snow, A. D., J. Whitaker, M. Cochran, J. Van den Bossche, C. Mayo, J. de Kloe, C. Karney, B. Couwenberg, G. Lostis, J. Dearing, G. Ouzounoudis, C. Gohlke, M. Itkin, R. May, B. M. Wiedemann, C. Barker, C. Willoughby, D. Haberthür, E. Popov, J. de Maeyer, J. Ranalli, K. Evers, M. A. da Costa, M. Raspaud, M. Brett, M. Dunphy, and M. Taves (2021). Pyproj4/Pyproj: 3.0.1 Release. Zenodo.
- Ståhle, A., L. H. Marcus, and A. Karlström. (2005). "Place Syntax : Geographic Accessibility with Axial Lines in GIS." In Proceedings, Fifth International Space Syntax Symposium, 131–44, 1 January 2005.

- Tripathy, P., P. Rao, K. Balakrishnan, and T. Malladi. (2020). “An Open-Source Tool to Extract Natural Continuity and Hierarchy of Urban Street Networks.” *Environment and Planning B: Urban Analytics and City Science*. <https://doi.org/10.1177/2399808320967680>.
- Turner, A. (2001). “Depthmap: A Program to Perform Visibility Graph Analysis.” In 3rd International Space Syntax Symposium, Atlanta, 2001.
- Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. j. Carey, Í. Polat, Y. u. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. a. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt. (2020). “SciPy 1.0: Fundamental algorithms for scientific computing in python.” *Nature Methods* 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
- Waskom, M., and team the seaborn development. (2020). Mwaskom/Seaborn. Zenodo. <https://doi.org/10.5281/zenodo.592845>.
- Wolf, L. J., E. Knaap, and S. Rey. (2019). “Geosilhouettes: Geographical Measures of Cluster Fit.” *Environment and Planning B: Urban Analytics and City Science* 48(3), 521–539. <https://doi.org/10.1177/2399808319875752>.
- Yu, J., Z. Zhang, and M. Sarwat. (2019). “Spatial Data Management in Apache Spark: The Geospatial Perspective and Beyond.” *GeoInformatica* 23(1), 37–78. <https://doi.org/10.1007/s10707-018-0330-9>.