**Morphology of street networks in urban neighborhoods in Ghana**

A Special Study submitted to the Department of Planning, Kwame

Nkrumah University of Science and Technology, Kumasi

in partial fulfilment of the requirements for the

Degree of Bachelor of Science in Human Settlement Planning

By

JOSEPH NORKPLIM ATTAH

**CHAPTER THREE**

**RESEARCH METHODOLOGY AND PROFILE OF STUDY AREA**

**3.1 Introduction**

This chapter presents the modern analytical and open-science methods, tools, and resources used during the research. It continues to emphasize the importance of supporting and doing open and collaborative research using modern computational tools at our disposal as researchers and policymakers. The overarching aim of this study is to show how these tools make it easier to understand the intricate structure of street networks, making the argument that the only way to make things better especially in developing countries is by joining forces and doing mutually beneficial work that can be built upon by both policy and pedagogy. It also stresses that this study seeks to build upon work done by Dumedah & Garsonu (2021b) to advance and popularize the use of the modern open-science and computational urban informatics field and its importance in transportation planning, settlement design, and other fields involved geospatial analytics.

**3.2 Study Area and Data Sources**

The study area for this research encompasses six neighborhoods from two of the most populated districts in Ghana which comprise regional capitals in their respective regions, Accra and Kumasi. Accra (*see* Figure 1) the national capital of Ghana is by far the most populated and in a close second is, Kumasi (*see* Figure 2), which is the regional capital of the Ashanti Region. According to the provisional report from the population and housing census of Ghana conducted in 2021, one-third of persons living in Ghana live in either the Greater Accra Region with its capital Accra or Ashanti Region with its capital Kumasi.

All street network data is downloaded from OpenStreetMap, a collaborative open mapping project that provides spatial datasets covering every place on earth. As of 2016 it was reported to be 83% complete worldwide and of high spatial resolution (Boeing, 2020b; Haklay, 2010; Neis et al., 2011). Accessing the databases of OSM is free of any charges which is a huge motivator, especially for students wanting to conduct geospatial research. Also, for a country like Ghana where it is hard to obtain accurate and valuable geospatial datasets from any local agencies (Dumedah & Garsonu, 2021b), OSM is the best bet for obtaining data for any kind of geospatial analysis. However, it should be noted that, though the data from OSM is almost complete and of high quality, further preprocessing is needed to qualify the data for the kind of street network analysis described in this research.

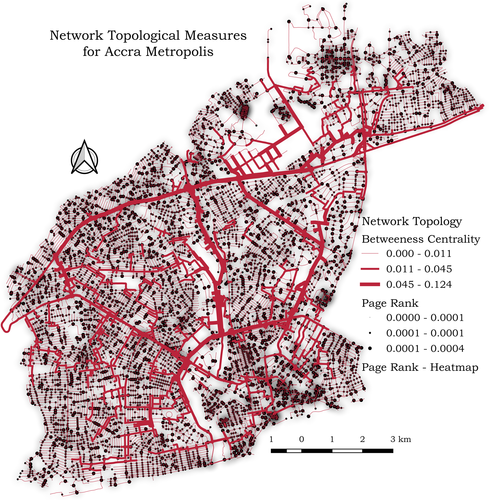


Figure 1. Graph theoretic model of Accra metropolis showing different topological and geometric features of the street network

*Source : Adapted from* (Dumedah & Garsonu, 2021b)

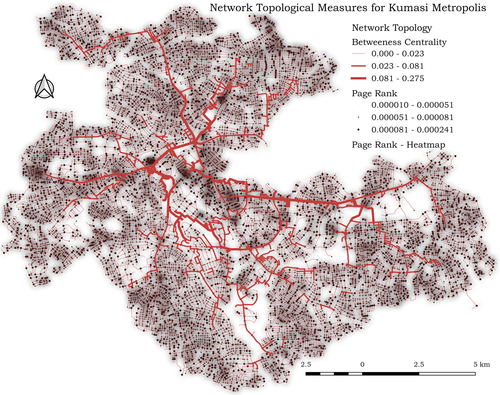


Figure 2. Graph theoretic model of Kumasi metropolis showing different topological and geometric features of the street network

*Source: Adapted from* (Dumedah & Garsonu, 2021b)

**3.3 Opensource Analytical Framework**

Following the approach to produce research that qualifies as open in its entirety, all tools, methodologies, and resources used to create the analysis framework are open collaborative projects and free resources gathered from the internet. OSM data is obtained and preprocessed primarily with the OSMnx and NetworkX tools. This tool allowed us to acquire political boundaries and building footprints, and download, and construct street networks into multigraphs for further analysis. In the spirit of automating and documenting the workflow involved in carrying out the analysis, a separate python module, AutoGIS (Joseph Norkplim, 2022), was created to automate the processes involved in acquiring, visualizing, and obtaining topological and geometric measures from the data. This effort was done in direct response to the paper by Boeing (2020b) urging researchers to engage some of their efforts in building new tools and documenting existing ones to make the analytical landscape a more approachable one for the younglings in the field and break the dependence on firms who own most of the existing proprietary tools used in geospatial analysis.

All tools integral to the research include; 1. NetworkX, a python language package for exploration and analysis of networks and network algorithms, (Hagberg et al., 2008) was fundamental to a lot of the street network analysis that was undertaken and algorithms used. 2. Geopandas and Pandas, open source data analytics tools for fast and programmatic manipulation of data both geospatial or otherwise (Jordahl et al., 2019). 3. Jupyter notebook, another open source tool that provides a fully hosted in-browser python execution environment that facilitates the sharing of code snippets, workflows, data, and visualizations detailing the research process. It features a virtual lab environment for computational analysis and a community that actively develops, documents, and updates it to what is considered standard in the data analytics landscape (Boeing, 2019b; Randles et al., n.d.). 4. Matplotlib, a portable 2D plotting and imaging python package aimed primarily at the visualization of data (Barrett et al., 2005). At the heart of all these opensource tools is Python, a Turing Complete, general purpose, dynamically typed, interpreted, high-level programming language that for its expressiveness—due to its lux type system—is useful in the modern data analytics framework (Ayer et al., 2014; Van Rossum & Drake Jr, 1995).

For downloading and pre-processing of geospatial data, the autogis tool is responsible for taking coordinates or place names of the study area (mostly embedded in a CSV file) it geocodes the place names to coordinates or reverse-geocodes the coordinates to get the place names and then proceeds to use OSMnx to download and construct the street network graph of the specified areas. With the use of the Matplotlib tool, AutoGIS is capable of both interactive and static plotting of geographic coordinates in any CRS (Coordinate Referencing System). OSMnx uses the NetworkX tool to correct most of the anomalies that appear in representing geospatial data as multigraphs, it does these corrections by removing points along curves that separate single streets into multiple edges (Boeing, 2017a). All these are done under the hood and the process is not visible to the third-party user. Consequently, the output of all this work produces a graph-theoretic representation of the street network of the study areas, that we derive meaningful insights. All the processes, data, tools, and processes are completely documented, reproducible and open to the general public in the public repository specified in the reference (Joseph Norkplim, 2022).

**3.4 Measures of Network Topology and Geometry**

The morphological and design properties coupled with the network's topology and its design have a great effect on the functioning of the network and how efficient the network is in its performance. And since street networks are the backbone of things flowing through space, the entire urban infrastructure is affected if the street network is affected. It is, therefore, necessary that the topological configuration, connectedness, robustness of the network, and its geometry—which is a concern of its design and placement in space—are measured and insights drawn from such measurements to guide the continuous development of the entire urban infrastructure (Boeing, 2018; Sharifi, 2019).

From existing literature, the basic topological measures include the measures of density, connectedness, length, degrees of nodes, and edges in the graph-theoretic representation of the street network (Barthélemy, 2011a; Boeing, 2019c, 2020a). We measure for each network that we have, the total number of nodes and edges connecting those nodes or themselves—in the case of self-loops. The average of the node degree is calculated and indicates the connectedness of the network, higher values are indicative of a more connected graph with lots of options for turns (Yen et al., 2021). The optimal functioning of street networks hinges on the number and connectedness of nodes and edges (whether they are fine-grained or coarse-grained), their capacity, and how they are connected as Sharifi (2019) argues.

When talking about the topology of a network, centrality, and connectivity are major interwoven measures that are important to understanding how the network functions (or at the very least speculating it). Centrality is important because each node/edge in the network differs from the others, which can be a factor of its placement in space and its interaction with other elements in space, therefore there is the need to compute the importance (centrality) of each node/edge to the overall functioning of the network. A node’s degree is the number of edges incident to it (Boeing, 2017a). Therefore, the degree of centrality is used to rank nodes based on the number of edges/nodes connected (incident) to it. The more connected a node is to other nodes in the network, the higher its degree of centrality in the network. Other measures of centrality include betweenness centrality, closeness centrality, information centrality, and straightness centrality (Boeing, 2017a; Sharifi, 2019). The closeness centrality of a node is indicative of the time and distance required to reach other nodes (destinations) in the network assuming it is the source. It is essential to consider closeness centrality when making decisions about the location and accessibility of amenities and services (Sharifi, 2019). Betweenness centrality, on the other hand, is an indication of how many shortest paths pass through a certain node, this measure is particularly indicative of how central the nodes to the network. The higher the betweenness centrality of a node, the higher the number of shortest paths passing through it, which implies how central it is to commute behaviors as most rational commuters will pick routes that get them to their destinations in less time. An unevenly distributed betweenness centrality is indicative of a fragile network, one that when node(s) with high betweenness centrality fail (or are removed), the network breaks, and things come to a halt in the system (Boeing, 2020a; Sharifi, 2019).

Connectivity measures are used to examine the functionality of the street network under normal and/or in emergencies (Sharifi, 2019). The node connectivity of the network is indicative of how resilient the network is, especially concerning the minimum number of nodes to remove from the network to disconnect the network (Boeing, 2017a) and the same goes for the edge connectivity, it is the minimum number of edges to remove from the network to disconnect it and render it useless. A well-connected street network is expected to facilitate smooth origin-destination flow, reduce travel distance, and improve access to services, employment, and utilities; people often have a perception of certain trip length thresholds when making decisions to walk, bike, or drive (Sharifi, 2019). Consequently, having redundant connections (alternative routes) in the network is helpful to maintain the state of the network in the case of emergencies. Other measures of connectivity include intersection density—the number of nodes per unit area, the average distance between intersections, and characteristic path lengths. It is to be noted that, street patterns have a significant bearing on how connected the network is (Barthélemy & Flammini, 2008; Boeing, n.d.; Sharifi, 2019).

Other important topological and geometric measures extracted from OSMnx using the autogis tool are presented and summarized in Table 1, which is adapted from Boeing and Dumedah & Garsonu (2017a; 2021b). Emphasis is placed on network topological measures like clustering, which measures how strongly connected a network is (Boeing, 2017a). Consequently, the averages of nodes and edge degrees, connectivity indices, intersection densities, PageRank and centrality, and measures of street design intricacies like network patterns, area of the network, and block sizes are taken into consideration in the analyses because of how important they are to understanding the street network in all of its totality (Boeing, 2017a; Sharifi, 2019). It is to be noted that all metrics and measures are extracted from a planar graph model of the street network in the selected study areas.

Table 1. Descriptive statistic measures of topological and geometric network features used to evaluate street networks

|  |  |
| --- | --- |
| **Metrics and Measures** | **Description** |
| Area | The total area that the network covers |
| n – number of nodes | Number of nodes in the network |
| m – number of edges | Number of edges in the network |
| Proportion and count of streets per node | Number of streets incident to each node and the ratio of edges to each node |
| Average Street length | The average length of edges in the network, which is a proxy for block size |
| Intersection density | The ratio of the number of intersections to the total area of the network |
| Node/Edge densities | The ratio of total counts of nodes/edges to the area of the graph; is indicative of whether the network is fine-grained or coarse-grained |
| Average street per node | Average of the number of streets emanating from each node |
| Average circuity | The ratio of network distance to Euclidean distance (its inverse is directness); is a proxy for how long and complicated routes in the network are. |
| Self-loop proportion | The proportion of edges that have a single incident node. |
| Diameter/Radius | Maximum/minimum distance from a node to all other nodes in the network. |
| Degree Centrality | The average number of nodes that a node is connected to; is used to rank the importance of each node in the network. |
| Node/Edge connectivity | The minimum number of nodes/edges that have to be disconnected to disrupt the flow of information in the network. |
| Clustering Coefficient | The extent to which a node’s neighborhood (edges and nodes incident to it) forms a complete graph; high values are indicative of a self-sufficient and strong network structure. |
| Betweenness centrality | The proportion of shortest paths passing through the node; is indicative of node importance in the network. |
| Closeness centrality | The average distance from a node to all other nodes in the network. |
| Page rank | Ranking of nodes based on the structure of incoming edges. |

*Source: Adapted from* Boeing (2017b, 2019a) and Dumedah & Garsonu (2021a)

**3.5 Variables and Data Collection**

The research starts out seeking to understand the current tool landscape as used in similar literature, one of its most salient objectives is to identify ways in which to secure free spatial data and tools for analytics. As put forward by Boeing (2020b) and Fleischmann et al., (2021), the current tool landscape of geospatial data analyses is dominated by point-and-click GIS tools (e.g. QGIS and ArcGIS) that do not take full advantage of the computational power that current computer systems come with, this is not to say they are not important. But they require more time to operate, are proprietary and difficult to document, and lack any form of comprehensive automation capabilities (especially for inexperienced people who are just starting out in the field). This raises the bar of entry for interested people both in policy and in research and limits the usability of these tools especially in this era of big data. Consequently, these limitations led the author of this work to contribute to developments in the field by creating and documenting tools that can be built upon by researchers and policymakers in the pursuit of studying urban street network form and structure. Graph theory as is employed by most studies involved in acquiring a modern intuitive view of spatial networks—of which street networks are an integral part—is the backbone of the current literature (Barthélemy, 2011b; Boeing, 2017b, 2019b, 2020b, 2021; Brede, 2012; Corcoran et al., 2013; Dumedah & Eshun, 2020; Dumedah & Garsonu, 2021b).

Constructing a graph model of street networks is no small feat and the use of tools that make it hard to automate the processes involved make it harder for interested people who are not well vexed in the mathematical and computational rigor required to do such work. The criteria for selecting the tools used in this research was 1. Tools that are free and open-source 2. Tools that do not require domain specific knowledge in computer science or mathematics; because tools like this are harder to work with. 3. Tools that are programmatic and easy to automate. 4. Tools that do not require any kind of special hardware to use; there are some analysis tools that require the use of GPUs (Graphic Processing Units) instead of CPUs (Central Processing Unit) to ran any kind of efficient analysis. 5. Tools that are more suitable for spatio-temporal network analysis. Considering the above criteria, only QGIS (Development Team, 2009) (and its network analysis tools) and Python—specifically OSMnx and NetworkX—make the cut as free and open-source tools (or ecosystems) used in the geospatial network analytics landscape. Based on google search indexing, searching for the keyword “geospatial network analysis tools” almost always yields a front page result listing ArcGIS Pro, a proprietary software package for geospatial analysis and ESRI, the company that creates, documents, updates and owns the software. It is almost as if the term geospatial analysis is synonymous with ArcGIS or other GIS related tooling. Point-and-click GIS tools were disqualified based on the premise that they are harder to automate and documentation and usability does not nearly reach the level that Python tools can boast of (Ayer et al., 2014; Boeing, 2020b). And because python is an open-source project, voluntary contributions to its usability and documentation make it an easy tool to learn and use. According to the TIOBE index, Python is the most popular programming language in the world, the criteria for this ranking is based on assessing the results of searches on the language, tooling, skilled engineers, courses (and documentation) and third-party vendors. It is also relatively easier to learn the python programming language as resources are numerous and available in various languages and formats, most of which are also free and open-source.

**3.5.1 Spatial Data Collection and Analysis**

The study sites total six urban neighborhoods located in Accra and Kumasi, three from each district. To collect the data, 0.7 km2 bounding boxes are defined from randomly picked points in the various districts and the street network extracted from within the bounding boxes (see Figure 3). It is noted that these study sites are small and do not conform to local boundaries in their respective districts, but on a small scale are useful for visual comparisons of spatial variables inherent in network structure and configuration (Boeing, 2017a). Next, the autogis tool using the OSMnx python package downloads and constructs the directed street network graph of each selected site, projects them in the correct CRS (which is automatically determined at runtime) and plots them. It then proceeds with the NetworkX python package to calculate network topological and geometrical measures summarized in Table 1 for each graph model generated for every study site and aggregates them for further analysis and processing. These network variables are used to compare and contrast how different planning regimes and geographic features affect the structure of networks in different locations. The study identifies which districts possess more fine grained networks and which possess more coarse grained networks and how network patterns affect the functioning and resilience of the network (Sharifi, 2019).

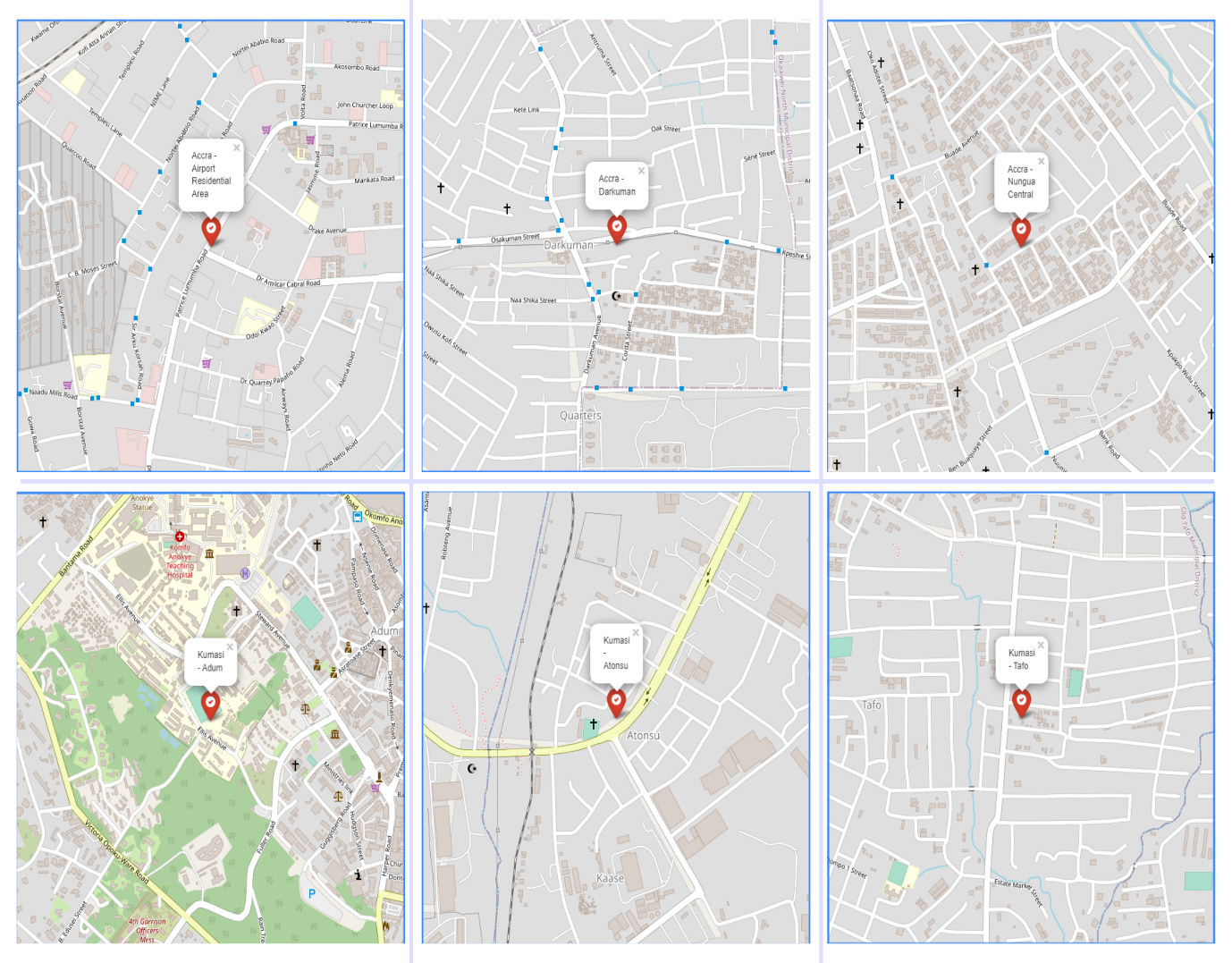


Figure 3. Six 0.7 km2 sections of street network from selected study regions

*Source: Author’s Construct (2022)*

**References**

Ayer, V., Miguez, S., & Toby, B. (2014). Why scientists should learn to program in Python. *Powder Diffraction*, *29*, S48-D64. https://doi.org/10.1017/S0885715614000931]

Barrett, P., Hunter, J., Miller, J. T., Hsu, J.-C., & Greenfield, P. (2005). *matplotlib -- A Portable Python Plotting Package*.

Barthélemy, M. (2011a). Spatial Networks. *Physics Reports*, *499*(1–3), 1–101. https://doi.org/10.1016/j.physrep.2010.11.002

Barthélemy, M. (2011b). Spatial networks. *Physics Reports*, *499*(1–3), 1–101. https://doi.org/10.1016/J.PHYSREP.2010.11.002

Barthélemy, M., & Flammini, A. (2008). Modeling Urban street patterns. *Physical Review Letters*, *100*(13). https://doi.org/10.1103/PHYSREVLETT.100.138702

Boeing, G. (n.d.). *Planarity and street network representation in urban form analysis*. https://doi.org/10.1177/2399808318802941

Boeing, G. (2017a). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Boeing, G. (2017b). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. https://doi.org/10.1016/J.COMPENVURBSYS.2017.05.004

Boeing, G. (2018). Planarity and Street Network Representation in Urban Form Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3191236

Boeing, G. (2019a). *The Morphology and Circuity of Walkable and Drivable Street Networks*. https://doi.org/10.31235/osf.io/edj2s

Boeing, G. (2019b). Urban street network analysis in a computational notebook. *Region*, *6*(3), 39–51. https://doi.org/10.18335/region.v6i3.278

Boeing, G. (2019c). Urban spatial order: street network orientation, configuration, and entropy. *Applied Network Science*, *4*(1). https://doi.org/10.1007/s41109-019-0189-1

Boeing, G. (2020a). Planarity and street network representation in urban form analysis. *Environment and Planning B: Urban Analytics and City Science*, *47*(5). https://doi.org/10.1177/2399808318802941

Boeing, G. (2020b). The right tools for the job: The case for spatial science tool-building. *Transactions in GIS*, *24*(5), 1299–1314. https://doi.org/10.1111/TGIS.12678

Boeing, G. (2021). Spatial information and the legibility of urban form: Big data in urban morphology. *International Journal of Information Management*, *56*. https://doi.org/10.1016/j.ijinfomgt.2019.09.009

Brede, M. (2012). Networks—An Introduction . Mark E. J. Newman. (2010, Oxford University Press.) $65.38, £35.96 (hardcover), 772 pages. ISBN-978-0-19-920665-0. . *Artificial Life*, *18*(2), 241–242. https://doi.org/10.1162/ARTL\_R\_00062

Corcoran, P., Mooney, P., & Bertolotto, M. (2013). Analysing the growth of OpenStreetMap networks. *Spatial Statistics*, *3*, 21–32. https://doi.org/10.1016/j.spasta.2013.01.002

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. https://doi.org/10.1016/j.jtrangeo.2020.102688

Dumedah, G., & Garsonu, E. K. (2021a). Characterising the structural pattern of urban road networks in Ghana using geometric and topological measures. *Geo: Geography and Environment*, *8*(1). https://doi.org/10.1002/geo2.95

Dumedah, G., & Garsonu, E. K. (2021b). Characterising the structural pattern of urban road networks in Ghana using geometric and topological measures. *Geo: Geography and Environment*, *8*(1), e00095. https://doi.org/10.1002/GEO2.95

Fleischmann, M., Feliciotti, A., & Kerr, W. (2021). Evolution of Urban Patterns: Urban Morphology as an Open Reproducible Data Science. *Geographical Analysis*. https://doi.org/10.1111/GEAN.12302

Hagberg, A., Schult, D., & Swart, P. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. *Undefined*.

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. https://doi.org/10.1068/B35097

Jordahl, K., Bossche, J. Van den, Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Farmer, C., Cochran, M., Gillies, S., Bartos, M., Culbertson, L., Eubank, N., maxalbert, Fleischmann, M., Hjelle, G. A., Arribas-Bel, D., Ren, C., Rey, S., … Trengrove, J. (2019). *geopandas/geopandas: v0.4.1*. https://doi.org/10.5281/ZENODO.2585849

Joseph Norkplim, A. (2022). *AutoGIS Street Network Analytics Package*. https://github.com/Joe-Degs/AutoGIS/tree/master/test-thesis

Neis, P., Zielstra, D., & Zipf, A. (2011). The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, *4*(1), 1–21. https://doi.org/10.3390/FI4010001

QGIS Development Team. (2009). *QGIS Geographic Information System*. http://qgis.org

Randles, B. M., Golshan, M. S., Pasquetto, I. V, & Borgman, C. L. (n.d.). *Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study*. https://doi.org/10.1016/j.future.2011.08.004

Sharifi, A. (2019). Resilient urban forms: A review of literature on streets and street networks. *Building and Environment*, *147*, 171–187. https://doi.org/10.1016/J.BUILDENV.2018.09.040

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Yen, Y., Zhao, P., & Sohail, M. T. (2021). The morphology and circuity of walkable, bikeable, and drivable street networks in Phnom Penh, Cambodia. *Environment and Planning B: Urban Analytics and City Science*, *48*(1), 169–185. https://doi.org/10.1177/2399808319857726