**CHAPTER THREE**

**RESEARCH METHODOLOGY AND PROFILE OF STUDY AREA**

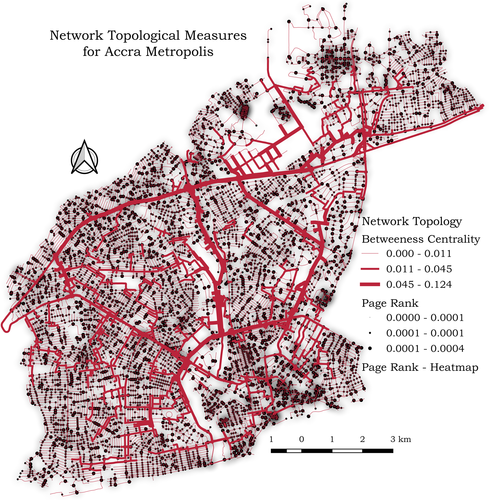
**3.0 Introduction**

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 collaborative research using the new computational tools at our disposal as people involved understanding how our settlements work, making the argument that the only way to make things better especially in developing countries is joining forces and doing mutually beneficial work that can be built upon by both practicing planners and those in 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 on transportation planning, settlement design and other fields involved geospatial analytics.

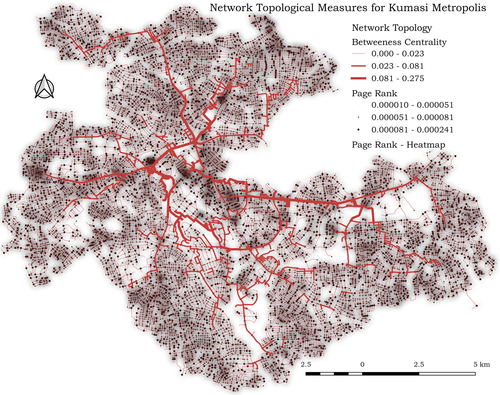
**3.1 Study Area and Data Sources**

The study area for this research encompasses six neigbourhoods from two of the most populated districts in Ghana which comprise regional capitals in their respective regions, Accra and Kumasi. Accra (*Fig 1*) the national capital of Ghana is by the far the most populated and in close second is the second most populated, Kumasi (*Fig 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 capital Accra or Ashanti Region with 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 at obtaining data for any kind of geospatial analysis. However, it should be noted that, though the data from OSM is almost complete and of a high quality, a further preprocessing is needed to qualify the data for the kind of street network analysis described in this research.



*Fig 1: Graph theoretic model of Accra metropolis showing different topological and geometric features of the street network (Adapted from* (Dumedah & Garsonu, 2021b)*)*



*Fig 2: Graph theoretic model of Kumasi metropolis showing different topological and geometric features of the street network (Adapted from* (Dumedah & Garsonu, 2021b)*)*

**2.2 Opensource Analytical Framework**

Following the approach to produce research that qualifies to be described as open in all of its entirety, all tools, methodologies and resources used to create the analysis framework are open collaborative projects and resources gathered from the internet. OSM data is obtained and preprocessed with the OSMnx tool. This tool allowed us to aquire 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](https://github.com/Joe-Degs/AutoGIS/tree/master/test-thesis/autogis), was constructed from scratch. This effort was in direct response to the the paper by Boeing (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.

Other tools integral to the development include: 1. NetworkX, a python language package for exploration and analysis of networks and network algorithms (Hagberg et al., 2008) which was fundamental to a lot of the street network analysis that was undertaken. 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 community that actively develops and updates it to the standards in data analytics (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 geocoordinates in any CRS (Coordinate Referencing System). OSMnx uses the NetworkX tool to correct most of the anomalies that appear in representing geospatial data as a multigraph, it does this corrections by removing points along curves that separates the street into multiple edges. All these are done under the hood and the process 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 from. All the processes, data, tools and processes are completely documented, reproducible and open to the general public in the public repository [here](https://github.com/Joe-Degs/AutoGIS/tree/master/test-thesis).

**2.3 Measures of Network Topology and Geometry**

The morphological and design properties coupled with the networks topology and its design have 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 the 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, 2011; Boeing, 2019c, 2020a). We measure for each network that we have, the total number of nodes and edges connecting those nodes to each other 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). Optimal functioning of street networks hinges on the number and connectedness of nodes and edges, their capacity and how they are situated with respect to one another as Sharifi argues (Sharifi, 2019).

When talking about the topology of a network, centrality and connectivity are major intertwined measures that are important to knowing how the network functions. Centrality is important because not all nodes or edges in a graph are the same, therefore there is the need to compute the importance of each node and 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 of a node is the number of nodes connected to it by edges. The more connected a node is to other nodes in the network, the higher its degree centrality in the network. Other measures of centrality include, betweeness 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 in the network from this node. It is essential to consider closeness centrality when making decisions about location and accessibility of amenities and services. Betweeness centrality on the other hand, is an indication of how many shortest paths pass through a certain node. The higher the betweenness centrality of a node, the higher the number of shortest paths passing through it. An unevenly distributed betweenness centrality is indicative of a fragile network, one that when the node(s) with high betweenness centrality fail (or are removed), the network breaks and things come to 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 emergency situations (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 totally 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 accessibility to services, employment, and utilities. People often have a perception of certain trip length thresholds when making decisions to walk or bike (Sharifi, 2019). Consequently, having a redundant connections are helpful to maintain the network state in the case of emergencies. Other measures of connectivity include intersection density—number of nodes per unit area, average distance between intersections and characteristic path lengths. It is to be noted that, street pattern has a significant bearing on how connected the network is (Sharifi, 2019).

Other impotant topological and geometric measures extracted from OSMnx using the autogis tool is presented is summarized in Table 1, which is adapted from Boeing and Dumedah & Garsonu (Boeing, 2017a; Dumedah & Garsonu, 2021b). Emphases is placed on network topogical measures like clustering, which measures how strongly connected a network is. Consequently, the averages of nodes and edge degrees, connectivity indices, intersection densities, PageRank and centrality and measure of street design intricacies like network patterns, area of network, 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 (Sharifi, 2019). It is to be noted that the all metrics and measures are extracted from a primal graph theoretic model of the street network in the selected study areas.

|  |  |
| --- | --- |
| **Metrics and Measures** | **Description** |
| Area | Total area that the network covers |
| n – number of nodes | Number of nodes in network |
| m – number of edges | Number of edges in 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 | Average length of edges in the network, which is a proxy for block sizes |
| Intersection density | Ratio of number of intersections to the total area of the network |
| Node/Edge densities | Ratio of total counts of nodes/edges to area of graph, which is indicative of whether the network is fine grained or coarse grained |
| Average street per node | Average of the number of street emanating from each node |
| Average circuity | Ratio of network distance to Euclidean distance (its inverse is directness) |
| Self-loop proportion | 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 | Average number of nodes that each node is connected to, used to rank the importance of each node in the network. |
| Node/Edge connectivity | Minimum number of nodes/egdes that have to be disconnected to disrupt flow in the network |
| Clustering Coefficient | Extent to which a node’s neighborhood (edges and nodes incident to it) form a complete graph |
| Betweenness centrality | Proportion of shortest paths passing through the node |
| Closeness centrality | Average distance from node to all other nodes in the network |
| Page rank | Ranking of nodes based on the structure of incoming edges |

*Table 1: Descriptive statistic measures of topological and geometric network features used to evaluate street networks (Adapted from* (Boeing, 2017b, 2019a; Dumedah & Garsonu, 2021a)*)*

**Variables and Data Collection and Analysis**

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

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

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

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

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

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