**MORPHOLOGY OF STREET NETWORKS IN URBAN NEIGHBORHOODS IN GHANA – chapter three**

**Research Design and Materials—an Open-Science Approach**

**3.0 Introduction**

This chapter discusses the modern analytical and open-science methods, tools and resources used during the research. It continues to postulate on the importance of supporting and doing open collaborative research using the new computational tools at our disposal as settlement planners, 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 apart from this study, there is only one other research that uses the methods used in this study to analyse street networks in Ghana and how sad that is for all those involved in understanding the relationship between humans and the spaces they live in.

**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 the national capital of Ghana is by the far the most populated and in close second is the second most populated, Kumasi, 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, 2021), 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.

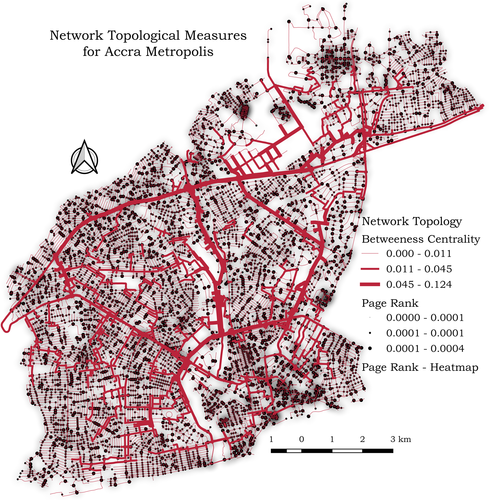


Fig 1

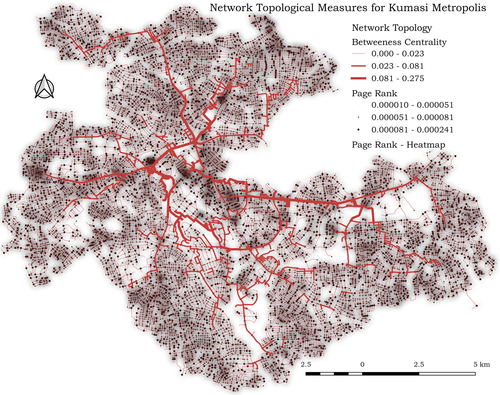


Fig 2

**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 tools allows 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 is 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, an 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 standards in data analytics (Boeing, 2019a; 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 name as of the study area—mostly 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 separate them 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 can 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 a 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 around in 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, 2019b, 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, 2017). 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, 2017; Sharifi, 2019). The closeness centrality of a node is indicative of the time and distance required to reach other nodes in the network. It is essential to consider closeness centrality when making decisions about location and accessibility of amenities. 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, 2017) 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 measures that will be taken into account are