Reducing Maternal Mortality by Improving Medical Facility Accessibility: A Methodology Demonstrated for the Democratic Republic of the Congo

by

Lindsay Lynette Aazami

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

December 2019

Copyright © 2019 by Lindsay Aazami

For my daughters, Tzipora Saoirse and Ellaria Shireen, may they become strong, independent, successful, and empowered women. My hard work and drive have always been to give them a better life than my own. I hope they see more, do more, learn more, and experience more than I could ever imagine for them. Thank you to my husband Kiarash, who sacrificed his time and especially energy to enable me to accomplish writing this while I worked more than full time all while keeping up with a three-year-old, a newborn, and then also moving across the world to New Zealand during my final semester. Without him taking care of the home, the cooking, play dates, bath times, bedtimes, grocery shopping, and an endless number of other tasks, there is no way I could have finished this. He always believed in me, and I want to thank him for being the only person that ever has.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my advisor, Dr Karen Kemp, for her patience, guidance, and for not giving up on me when she saw the first draft of this thesis (I’m sure she got a good laugh though).

List of Abbreviations

2SFCA	Two-Step Floating Catchment Area
AVHRR	Advanced Very High-Resolution Radiometer
CHD	Centre for Humanitarian Data
DRC	Democratic Republic of the Congo
GIS	Geographic Information System/Science
GIST	Geographic Information Science and Technology
GLCC	Global Land Cover Classification
GMTED2010	Global Multi-resolution Terrain Elevation Data, 2010
HDX	Humanitarian Data Exchange
MCSP	Global Maternal and Child Survival Program
NDVI	Normalized Difference Vegetation Index
NGA	National Geospatial-Intelligence Agency
OCHA	Office for the Coordination of Humanitarian Affairs
OSM	Open Street Map
UNFPA	The United Nations Population Fund (formerly: The United Nations Fund for Population Activities)
UNICEF	The United Nations Children's Fund (formerly: United Nations International Children's Emergency Fund)
USAID	United States Agency for International Development
USGS	United States Geological Survey
USC	University of Southern California
WHO	World Health Organization

Abstract

The Democratic Republic of the Congo (DRC) is the fourth most populated country in Africa, with approximately 87 million people, of which 44 million are female. Unfortunately, it also has the 10th highest maternal mortality rate of any country in the world at 693 deaths per 100,000 births in 2015. High maternal mortality in the DRC is due in large part to pregnant mothers being remotely located from medical facilities and routinely dying from preventable complications. Cars are not prevalent in the DRC, and the most common means of travel is by foot due to the destruction of the infrastructure caused by the First and Second Congo Wars in the late 1990s. Walking long distances during pregnancy or while in labor and especially at night is a significant barrier for women seeking medical care. This study's objective was to develop a simple methodology that could be used to identify ideal locations for new birth facilities where large populations are the furthest distance from existing facilities. The locations were identified through the generation of a tessellation grid over the areas of the DRC with low walking accessibility to medical facilities. For each tessellation grid cell, the distance to the nearest medical facility and the population within the cell were calculated. Based on a combination of population size and distance from a medical facility, a rank of locations for new facilities was created. Facilities built in these highest ranked locations would have the maximum impact by supporting the largest population that is the furthest distance from medical facilities. The resulting increase in medical accessibility could greatly decrease birth complications and preventable death.

Chapter 1 Introduction

Pregnancy and childbirth are a vulnerable time for women, and anyone could fall victim to minor complications, excessive bleeding, preterm birth, or infections. Mothers who have quick access to trained medical care have a better chance of surviving preventable complications. Maternal mortality prominently plagues Sub-Saharan Africa, as indicated by the United Nations Children's Fund (formerly known as the United Nations International Children's Emergency Fund, UNICEF) who have determined that the Sub-Saharan Africa region has the highest maternal mortality rate in the world (UNICEF 2019). According to UNICEF, there is evidence that almost all maternal deaths are preventable given the huge disparities in the mortality rates between wealthy and impoverished nations. Figure 1 shows the primary causes of maternal death in Sub-Saharan Africa, all of which are preventable if trained health services are accessible (Say et al. 2014). Indirect causes include anaemia, malaria, and heart disease (Nour 2008).

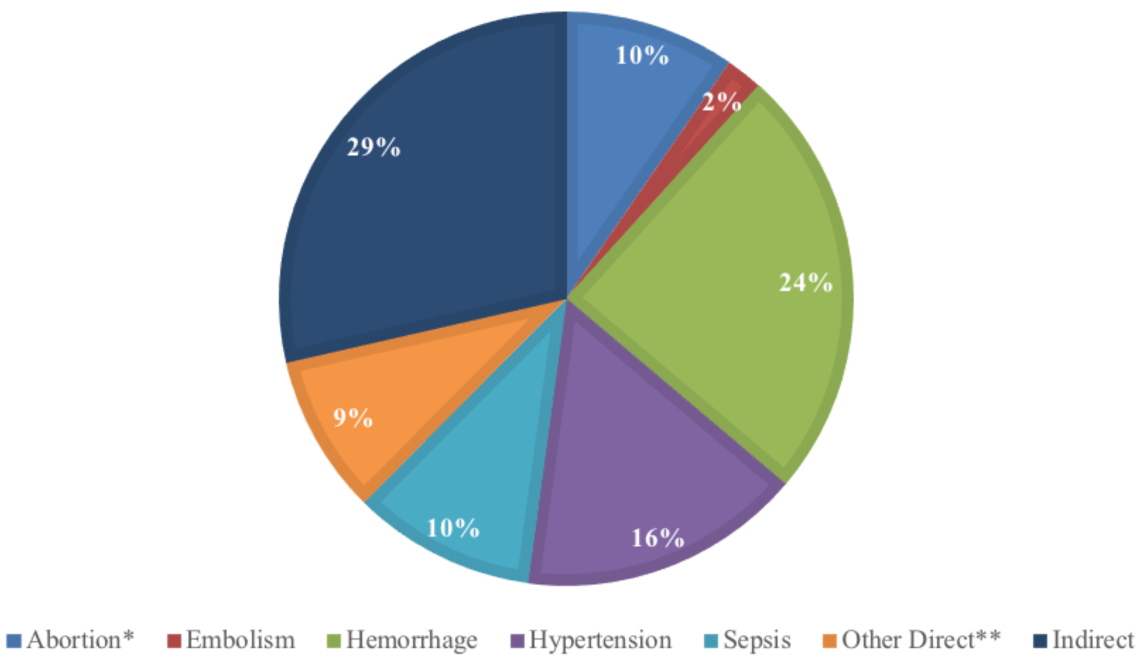

Figure 1 The Causes of Maternal Mortality in Sub-Saharan Africa. *Nearly all (99 per cent) of abortion deaths are due to unsafe abortions. **This category includes deaths due to obstructed labor or anemia. Source: Say et al. 2014

In 2016, the United Nations Population Fund, formerly the United Nations Fund for Population Activities (UNFPA), reported that women who do not receive medical care during pregnancy or at the time of birth are three to four times more likely to die during childbirth. Understanding how to optimally implement new medical facilities throughout Sub-Saharan Africa could prevent or reduce maternal mortality. In 2015, a meeting was conducted by the United States Agency for International Development's (USAID) global Maternal and Child Survival Program (MCSP) where 72 participants from over 25 global health organizations, government agencies, donors, universities, and other groups suggested that GIS mapping and data visualization could play a significant role in ending preventable maternal mortality. Examples suggested were mapping of hospital service areas of coverage based on distance or identifying locations more at risk for maternal mortality. Mapping populations in areas not practically serviced by a medical facility could enable policymakers to establish birthing centers in these locations for women in need.

The Democratic Republic of the Congo (DRC), the focus of this research, is the largest country within Sub-Saharan Africa and the fourth most populated country in Africa with a population of approximately 87 million people, of which 44 million are female (United Nations 2019). Unfortunately, in 2015, the DRC also had the 10th highest maternal mortality rate of any country in the world at 693 deaths per 100,000 births (World Bank 2015). This rate had dropped from 879 deaths per 100,000 births in 1990, but it is still unnecessarily high.

High maternal mortality in the DRC is due in large part to pregnant mothers being remotely located from medical facilities and routinely dying from preventable complications such as hemorrhage, unsafe abortion, hypertension, and sepsis. Medical facilities in the DRC are sparse and only located in major cities, leaving much of the population throughout the country to choose between traveling by foot over a long distance or to deliver at home and hope everything goes perfectly. Cars are not prevalent in the DRC, and the most common means of travel is by foot due to the destruction of the infrastructure caused by the First and Second Congo Wars in the late 1990s. Walking long distances during pregnancy or while in labor and especially at night is a significant barrier for women seeking medical care. The lack of access to trained medical care facilities is the direct link to maternal mortality death from preventable causes.

1.1. Objective

This objective of this study was to develop a simple GIS-based methodology for Sub-Saharan Africa using readily accessible data that can be used to identify ideal locations for new medical centers in order to increase accessibility to medical care. An increase in accessibility will reduce incidences of preventable maternal death. The methodology was developed with a focus on the Democratic Republic of the Congo (DRC) given its high maternal mortality rate related to a lack of access to medical facilities. This research identified and ranked by need, areas with the highest populations that are also the furthest distance from existing medical facilities within the DRC. The ranked list could be used to target and support the populations in the most need of medical services within the DRC. Further, the project sought to create a reproducible methodology that can be applied in other nations with limited medical care accessibility.

1.2. Overview of Methodology

This study identified the isolated areas of the DRC unserved by present medical facilitates, then identified and ranked by need, those isolated locations with the highest populations that are also the furthest distance from existing medical facilities. The isolated areas were identified by first determining regions of the country that are located within a maximum three-hour walking distance from present medical facilities. This three-hour walking distance service area takes into account straight-line distance as well as how land use type and elevation affect travel time. A traditional method utilizing the Generate Service Area tool could not be used as the infrastructure network had been destroyed and currently lacks functionality. As a result, the Path Distance tool, along with a Cost Surface, was used to determine the service area.

After the part of the DRC that fell within the three-hour medical facilities service area was identified, the remaining part of the country became the study area. With the unserved areas of the DRC identified, a tessellation grid was generated to break up the country's unserved area into 210.44 sq km hexagons. The generated grid was intersected with a polygon layer of health zones that had been joined with a health zones population table, the lowest level of aggregation of DRC population data publicly available. Finally, the grid cells' distances from existing medical facilities were calculated utilizing the Near function.

Ranks in ascending order were created for each hexagon for both the grid distance from existing medical facilities and the population that was assigned to each the grid polygon. The combination of both of these ranks in descending order creates the needs-based rank for the entire unserved parts of the country.

1.3. Structure of the Thesis

The remainder of this thesis is composed of 4 additional chapters. Chapter 2 discusses background research on related topics such as causes of maternal mortality, spatial analysis of maternal and newborn health, spatial distribution of medical facilities, and research related to the methodology used such as cost-distance. Chapter 3 details the data utilized and the source from which they were obtained. Chapter 3 also explains in detail the methodology used. Chapter 4 contains the results of the applied methodology on the designated study area. Chapter 5 is a discussion of the limitations, potential improvements, and opportunities for further expansion of this research.

Chapter 2 Related Work

This chapter explores research on related subject matters as well the methodology that was used in this study. Subject matter related to this study includes the causes of maternal mortality, spatial analysis of maternal and newborn health, and the spatial distribution of medical facilities. Methodologies discussed that are related to this study are service areas, catchment area, and cost distance.

2.1. Causes of Maternal Mortality

Maternal mortality is defined by the World Health Organization (WHO) as the death of a woman while pregnant or in the postnatal period within 42 days of the end of pregnancy from any cause related to or exacerbated by the pregnancy or its management, but not from accidental or incidental causes (WHO 2019). Maternal mortality is a global issue but primarily affects more impoverished nations due to the lack of access to medical facilities. In most cases, the reasons are entirely preventable or recoverable with proper medical care.

The leading causes of maternal mortality in the DRC are primarily poverty in the country and a general lack of the ability to pay for prenatal or childbirth medical expenses (The Guardian 2012). According to The Guardian article, even in cases where a woman was able to arrive at a medical facility in time to obtain care for the birth, if she lacked funds, she may be repeatedly transferred long distances between facilities. This report also noted that even if women have the funds and can be admitted for care, preventable incidences of maternal death still occur.

Women in the DRC have six births on average in their lifetime, and the chance of complications increases with each delivery (World Bank 2015). The pregnancy rate is primarily due to lack of contraception and the lack education surrounding family planning. Due to the limited use of contraception, multiple pregnancies that are closely spaced contribute to the high maternal mortality rates in the country.

Additionally, according to the authors of “Socio-Cultural Factors in Maternal Morbidity and Mortality,” women are likely to forgo utilizing medical services during pregnancy and childbirth because of the cultural preference to engage the services of local traditional healers (Okolocha et al. 1998). Local traditional healers offer cheaper and more accessible care than modern medical facilities but do not have the medical training or supplies needed to save lives when complications arise. A general lack of education regarding health care within the DRC leads the majority of the population to believe traditional healers are more knowledgeable and capable than medical doctors. The Okolocha et al. study showed that lack of education is a primary contributing factor in maternal mortality. The more educated the mother is, the more she can distinguish medical fact from superstition. Mothers with more education also tend to be located in areas of the country with higher accessibility to healthcare.

Graham and Ronsmans (2006) point out that impoverished countries have had the least success in the last 25 years in the global attempt to reduce preventable maternal mortality. The impoverished nations have been unable to reduce the rate due to weak health care systems, high fertility rates, and poor availability of data. The authors explain that poor and rurally located mothers are most susceptible to maternal mortality. They found that the highest cause of maternal death in Sub-Saharan Africa is severe bleeding or hemorrhaging along with hypertensive diseases and infections, which are all preventable with access to trained medical care. Access to nearby medical care increases the mother’s likelihood of utilizing care and relates to a faster response time when complications arise. It was observed by Graham and Ronsmans (2006) that women who have higher access to health care due to proximity have a lower maternal mortality rate. This suggests that implementing birth centers in remotely populated locations will increase the utilization of medical services and decrease maternal mortality rates.

Further research adds strength to the argument that maternal mortality amongst poor rural women in the DRC is directly related to lack of access to medical facilities. An analysis of World Health Organization data on global maternal mortality trends by Khan et al. (2006) reveals that while hemorrhage is the major contributor to maternal mortality, there are variations in causes at the regional level. In less developed African nations, hemorrhage, sepsis, and hypertension are the leading factors.

Finally, as reported by Ellison (2017), a lack of access to health care even in developed nations is a contributing factor of maternal mortality. In more developed nations, the lack of access to care is derived primarily from lack of financial means or insurance rather than being a long distance from medical facilities.

2.2. Spatial Analysis of Maternal and Newborn Health

Researchers have begun to use GIS to analyze and map out maternal and newborn health. One article by Ebner (2015) goes in-depth on spatial techniques used for analyses and visualization such as thematic mapping, spatial analysis to assess accessibility, spatial modeling, and small area estimation. These are all forms of analysis that could be used to both analyze the data and display the results. The author noted that the biggest issue in maternal health analysis is finding data that is aggregated at the lowest level possible as most data is typically published at the national level only.

Spatial analysis has been used previously to identify communities that are at a higher risk of a particular health issue. Chong et al. (2013) used spatial analysis to identify communities at a higher risk of both smoking while pregnant and delaying the initial antenatal visit. The authors reported that mothers that delay their initial prenatal visit are likely to experience adverse pregnancy-related outcomes. Primiparous women (first-time mothers), young mothers, and mothers with low socioeconomic status are likely to have delayed their initial antenatal visit. Chong at al. began by georeferencing reported home locations of at-risk mothers as recorded at the time of birth. After that, the authors utilized SaTScan to identify any spatial clusters that showed significant correlation between smoking mothers and mothers with late initial antenatal visits. The spatial clustering analysis revealed four areas that show correlated clustering. From here, the authors were able to come to the conclusion that a higher proportion of at-risk mothers lived in the most disadvantaged areas and a lower proportion were first time mothers.

A study by Molla et al. (2017) mapped the distribution of live births, pregnancies, and population data in order to find a cause behind maternal mortality. Additionally, the authors explained how mapping non-mortality factors that could support at-risk mothers, such as the distribution of midwifery services, and mapping indicators such as the proportion of women who need emergency obstetric care versus those that actually receive it, could reveal areas that underutilize medical services and those that lack access due to low proximity. Mapping factors and indicators such as these could enable governments to redistribute medical services to better support a population.

2.3. Spatial Distribution of Medical Facilities

Understanding the spatial distribution of medical facilities is just as important as understanding the distribution of the population being studied. The specific locations, the distributions, and the density of medical facilities as it is related to population density all affect the quality of healthcare available. Nwakeze (2011) compared the distribution of medical facilities as well as the ratio of available physicians in each region to the population distribution in Niger. The article argues that the sheer number of medical facilities is not what matters most. More important is the condition and size of the medical facility as well as the quality and quantity of health care professionals within it. Nwakeze blames a lack of organized leadership and inefficiencies in implementing the developed medical policies for the lack of health care availability. This is a similar issue in the DRC where three decades of civil war have left the country with no public funding for medical care, weak national leadership, and zero regulations enforced on the health sector leaving the population dependent on external aid (Ntembwa and Lerberghe 2015).

Previous research has been conducted on the distance a mother is living from existing medical facilities. Gabrysch et al. (2011) analyzed this relationship to find what influence the travel distance had on the quality of health care these mothers received. The article reported that the distance a mother is from medical facilities influences whether or not she is more likely to choose to give birth at home unassisted. This is an issue that is at the core of the research in this paper. Decreasing the travel distance to medical facilities and thus increasing the accessibility will help increase a mother's likelihood that she will seek out services when in need. Gabrysch et al. also analyzed the household's ability to pay for medical services based on assets owned. They concluded that the inability to pay may influence expectant mothers' willingness to travel for medical assistance during birth.

It is optimal to have new locations for services be ideally located so that the most remotely located childbearing aged female populations will have access. Ruiz (2010) discussed the implementation of maternity waiting homes and their effectiveness in Mozambique, Africa. Maternity waiting homes were created to supplement areas that were beyond a medical facility's service area. The maternity waiting homes were specifically created in order to increase the number of births in medical facilities and decrease maternal mortality. Ruiz noted that while these maternal waiting homes helped the populations that previously lacked access to medical services, there are still women located in remote regions that are unserved and at risk for complications.

GIS had been used previously to identify optimal locations for additional medical facilities. Massey (2011) wrote that by just looking at the numbers, there may appear to be enough health workers for a population, but this relationship is likely skewed due to distribution. The author notes that the health workers and medical facilities are likely clustered in the most urban areas of the country, leading to a disparity in coverage. Massey used spatial analysis with a priority index to identify areas in need of trained midwives. His methodology calculated the current number of midwives by region in Senegal and divided it by the WHO-recommended number of midwives per population which is 1 midwife per every 300 women. Then that result was subtracted from 1 and multiplied by 100 to result in the percent of shortage of midwives.

Another methodology Massey implemented in his article identified the priority area for midwife service expansion. Massey did a hot spot analysis based on percent shortage of midwives (at regional level), percent of childbirths occurring unassisted (at regional level), percent of women that received no antenatal care (at regional level), and percent of childbirths that take place at a health care facility. In his result, he was able to identify the eastern region of Senegal as the area most in need of midwifery services.

2.4. Service Areas, Catchment Areas & Cost-Distance

A network service area is a region that encompasses all streets accessible from a point in a designated amount of time. This analysis can be completed using a tool within ArcPro called

Generate Service Areas. Given a layer of points to start from, it can generate a map of the service areas of these points on the network based on travel time selected. Catchment areas, on the other hand, identify the locations from which individuals are attracted to access service or institutions. Catchments are defined by multiple factors like distance or population size, or they may simply partition the entire space into nearest neighbors. Becker (2016) utilized catchment areas to understand the accessibility and availability of veteran healthcare services utilizing a two-step floating catchment area (2SFCA) model. He elaborates on previous research done in this area by incorporating variables such as wait times, patient satisfaction, and acceptability of care. In his research, he was able to identify locations that do not meet the primary care service standards designated by the federal government.

Accessibility surface models using cost-distance algorithms have also been used to create catchment areas. Blanford et al. (2012) specifically addressed the ease of physical access to healthcare in Niger in both the wet season and the dry season. It was stated that due to walking being the primary form of transportation, access to medical facilities can be grossly overestimated based upon the season. Blanford et al. discovered that in Niger, women in rural areas walk on average 26 miles (about 42 km) to seek medical assistance. Walking 26 miles would take a woman about 8 hours. They noted that the utilization of medical facilities diminishes with both the travel distance, the quality of transportation, as well as road conditions. This issue is similar to the situation within the DRC due to a lack of infrastructure from damage caused by civil wars. Blanford et al. created an accessibility surface model based on distance using a cost distance algorithm. They also created a friction surface for this calculation to account for travel times of various travel methods across various surfaces. After that, using a subnational subdivision of population, they determined the distribution of the population that fell outside a four-hour travel distance. Blanford et al.'s study provided the framework for the development of the methodology used in this study.

2.5. Summary

The related research in this chapter provided a brief overview on maternal mortality and its causes and the uses and benefits of mapping maternal and newborn data. The complexity of mapping the distribution of facilities increases when taking into account whether the amount of trained medical personnel can support the population they are obligated to serve. There was also a brief overview of service areas, catchment areas, and cost-distance to find service areas. This study was unable to use network service areas as there is no traditional transportation infrastructure network to generate it. It was also unable to utilize catchment areas, which are designed for a population that is fully covered by services such that the catchment areas indicate which service facility each individual belongs to (in the way that school zones are assigned to house addresses). Blanford et al.'s (2012) study did provide some valuable guidance that is discussed further below.

Chapter 3 Data and Methods

This study sought to identify optimal locations to implement additional medical facilities for pregnancy and childbirth to support unserved areas of the DRC. This chapter provides details on the data and methods used to determine these optimal locations.

3.1. Data

The data needed to conduct the analysis in the study area of the DRC was a point file of the medical facilities, digital elevation data, land cover data, a table containing population data by health zone, and health zone boundaries of the DRC. Table 1 summarizes the content of each dataset used and lists its source.

Table 1 Data used in this study and sources

Data	Format	Content	Source	URL
Study Area	Polygon	Boundary of the DRC	OCHA CHD	https://data.humdata.org/dataset/drc-administrative-boundaries-levels-0-2
Health Zones	Polygon	Boundaries of the 519 Health Zones	OCHA CHA	https://data.humdata.org/dataset/dr-congo-health-0
Population Data	Table	Adult Female population data by health zone	OCHA CHD	https://data.humdata.org/dataset/rdc-statistiques-des-populations
Medical Facilities	Point	XY locations of all 1,086 medical facilities in the DRC	Healthsites	https://www.healthsites.io/#country-data
Elevation	Raster	30-arc second (1km) elevation data	USGS	https://www.usgs.gov/land-resources/eros/coastal-changes-and-impacts/gmted2010?qt-science_support_page_related_con=0#qt-science_support_page_related_con

Data	Format	Content	Source	URL
Land Cover	Raster	Classified land use raster	GLCC USGS	https://www.usgs.gov/centers/eros/science/usgs-eros-archive-land-cover-products-global-land-cover-characterization-glcc?qt-science_center_objects=0#qt-science_center_objects.

All of the data obtained from United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Centre for Humanitarian Data (CHD) (centre.humdata.org) was downloaded from their Humanitarian Data Exchange (HDX) website (data.humdata.org). While the HDX is an open data sharing site and data quality varies, the CHD strives for data of sufficient quality that is useful for humanitarian purposes. Importantly, OCHA intends that this site will increasingly become a source for the often difficult to find socio-demographic data for developing countries, such as sub-national level population data. However, for the purposes of this project intended to demonstrate a methodology, data quality is not critical, so it is assumed that the data retrieved from this site are of sufficient quality to develop the methodology.

3.1.1. Study Area

The study area was the country of the Democratic Republic of the Congo. The DRC is located in central Africa and by area is the largest country in Sub-Saharan Africa. The country boundary polygon for the DRC was obtained from the United Nations OCHA in West & Central Africa. The country boundary was used primarily to narrow the scope of data obtained to the DRC area. This boundary is shown in several of the maps below.

3.1.2. Health Zone Boundaries

The health system in the DRC is organized into 519 health zones which typically cover a population of 100,000 to 200,000 people. Medical facilities are not equally distributed between the zones or the population of the country as a whole. Like the study area boundary, the health zones polygon data were obtained from OCHA in West & Central Africa. These data are updated annually and were last updated on June 13, 2019. Figure 2 shows the 519 health zones within the DRC.

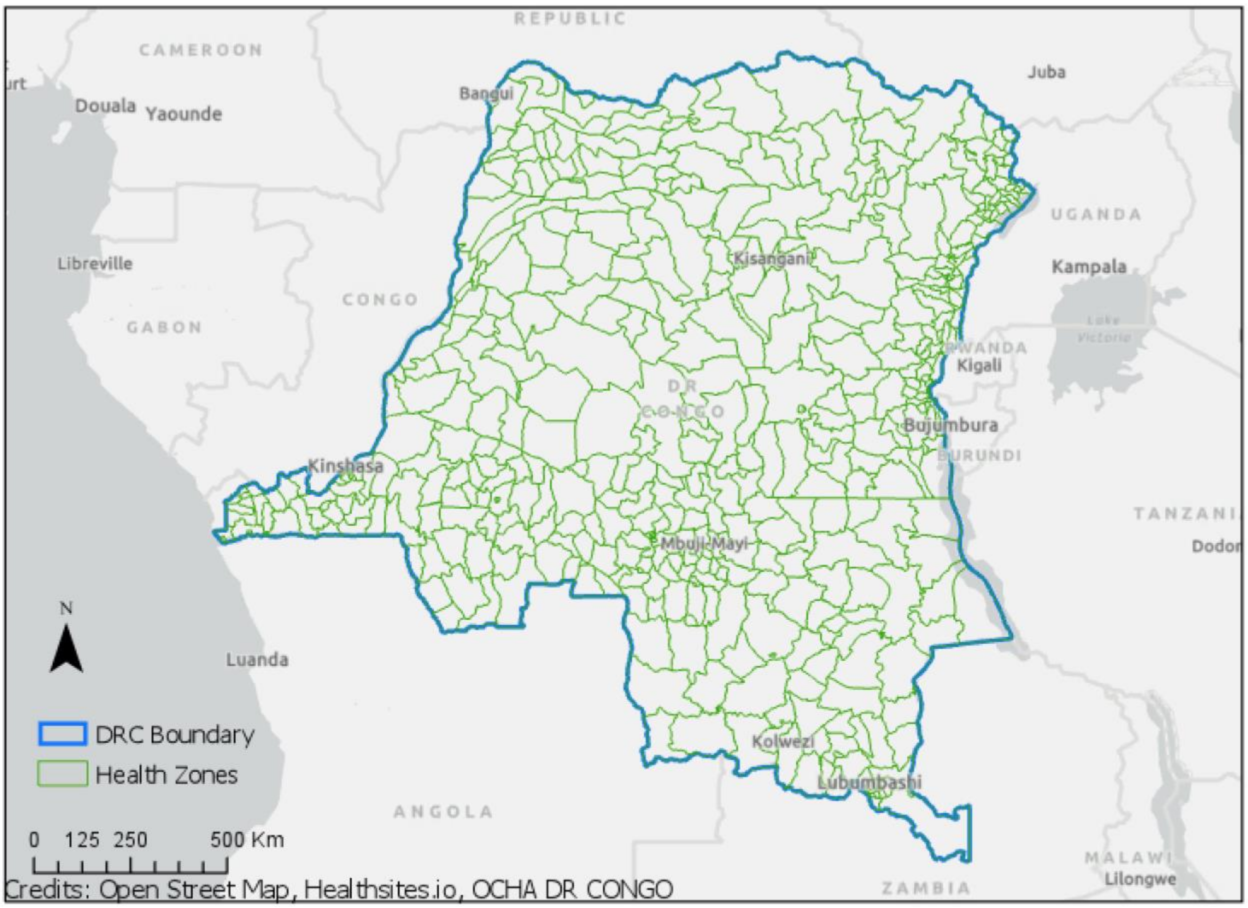

Figure 2 The boundaries of the DRC and the 519 Health Zones. Source: OCHA

3.1.3. Population Table

The population data for the Democratic Republic of the Congo was also obtained from OCHA. The table is updated annually and was last updated 17 September 2018. The table has population data for the 519 health zones. The attributes of the table are health zone codes, health zone names, and total population per zone as well as population totals for male, female, adult male, adult female, boys, girls, and elderly by zone. Since this study is about maternal mortality, only the adult female population attribute was utilized in the analysis. The table was joined to the health zone polygons using the health zone code. Table 1 displays a sample of the adult female population data obtained for the 519 health zones and Figures 3 and 4 show maps of the raw adult female counts and the adult female density in the health zones, respectively.

Table 2 Sample of the adult female population data aggregated by health zone. Source: OCHA

Health Zone	Health Zone Code	Adult Female Pop
Aba	CD5307ZS01	24184
Abuzi	CD4304ZS01	14548
Adi	CD5409ZS01	30686
Adja	CD5409ZS02	23964
Aketi	CD5204ZS01	25811
Alimbongo	CD6105ZS01	41835
Alunguli	CD6301ZS01	15179
Ango	CD5207ZS01	19757
Angumu	CD5407ZS01	31135
Ankoro	CD7406ZS01	48172

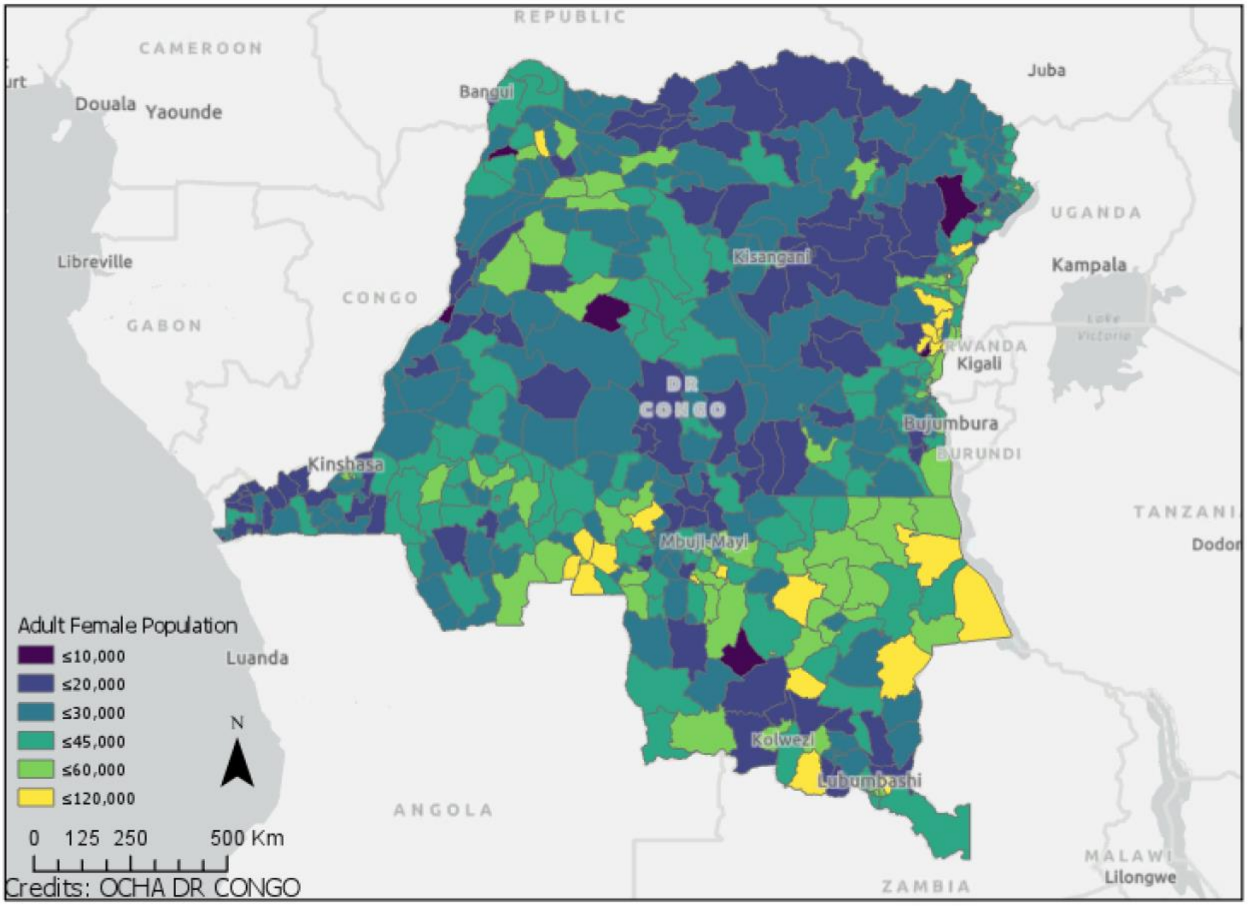

Figure 3 Adult female population count by health zone. Source: OCHA

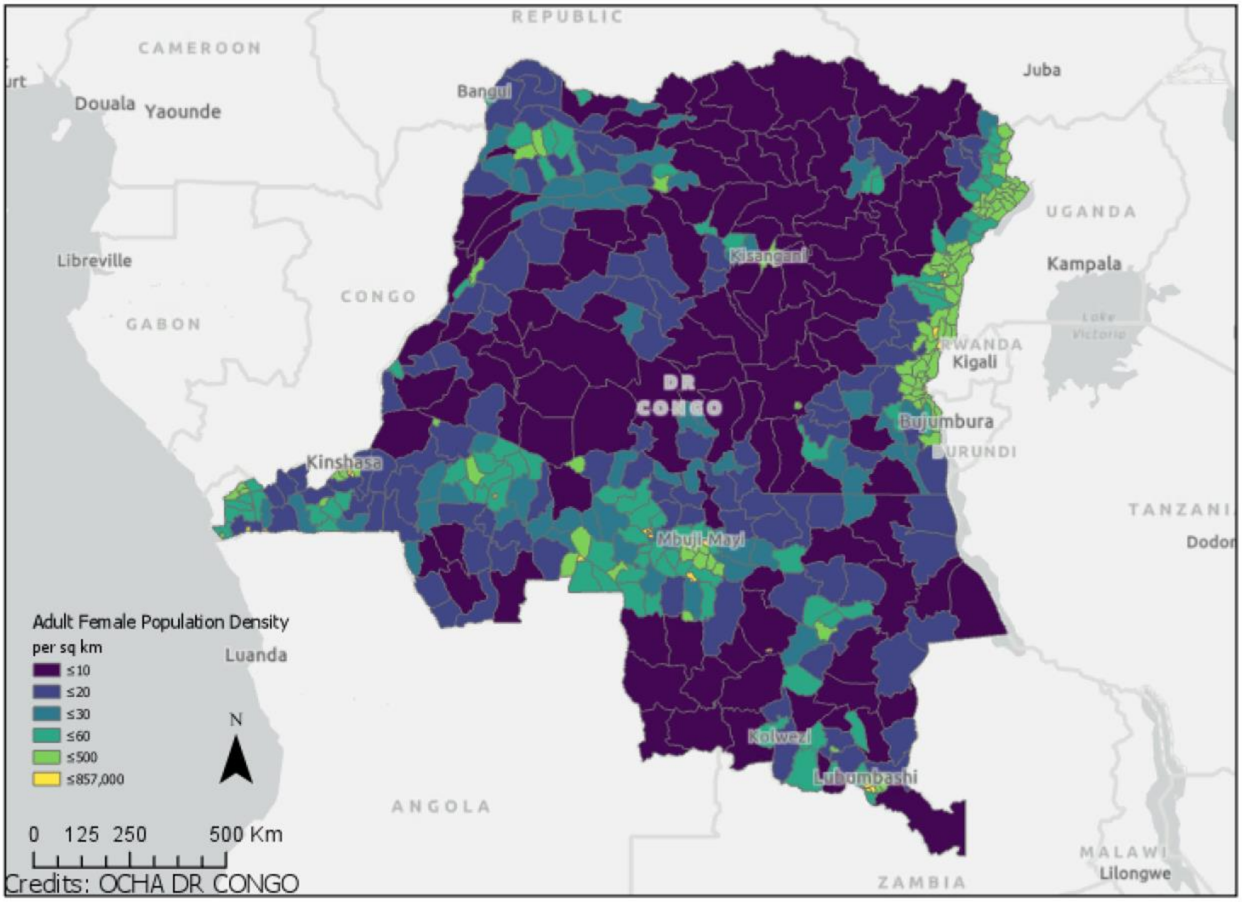

Figure 4 Adult female population density within the health zones. Source: OCHA

3.1.4. Medical Facilities

The existing medical facility location data was available for the Democratic Republic of the Congo from a website called the HealthSites.io that specializes in mapping medical facilities globally. According to their website, HealthSites' mission is to establish accurate health care location data that can be used to support first responders in the event of a natural disaster or disease outbreak. HealthSites produces and manages the Global HealthSites Mapping Project, which is an initiative to create an online map of every health facility in the world and make the details of each location easily accessible. HealthSites noted on their website that their data has proven extremely useful in the past, particularly during the Haiti Earthquake in 2010 and the Ebola epidemic in West Africa from 2013 to 2016.

HealthSites imports health care location information from OpenStreetMap (OSM) and other trusted partners such as Geomatica, Missing Maps, the International Committee of the Red Cross, the International Hospital Foundation, Radiant Earth Foundation, cartONG, and Medecins Sans Frontières. As Figure 5 shows, like other OSM crowd-sourced data validation workflows, community users of HealthSites update and validate health care center data and then specialist users validate the community user updates. The validated location data is then submitted back to OpenStreetMap where HealthSites access the data. The data for medical facilities is consistently being updated through crowd sourced users, trusted partners, and OSM field workers. Figure 5 shows HealthSites' data lifecycle process, illustrating the care that is being taken to ensure that data in the database is valid and accurate through multiple validations of the crowd sourced data.

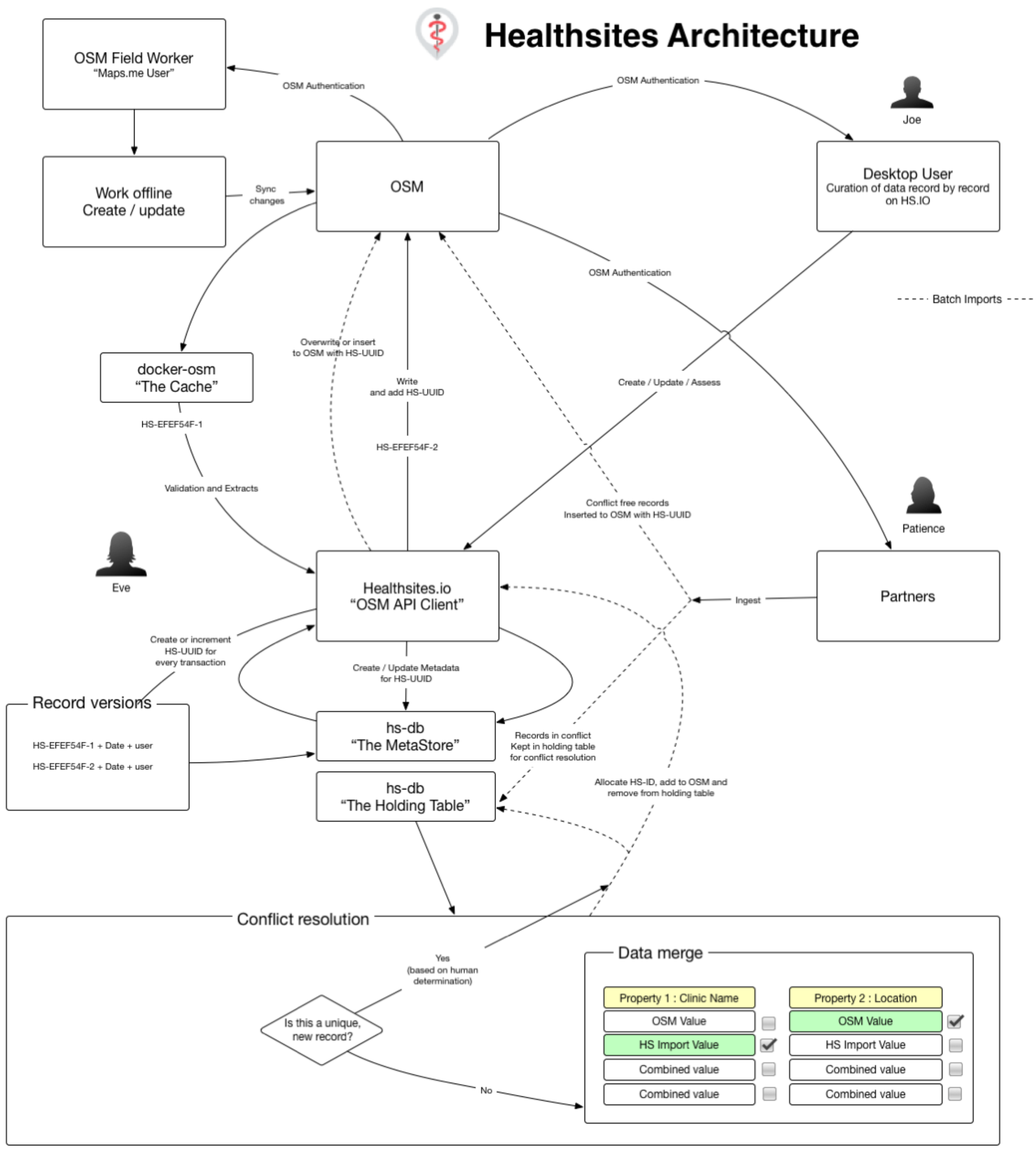

Figure 5 Healthsite.io data lifecycle. Source: healthsites.io

The Global HealthSites Mapping Project provided a point shapefile of 1,086 sites, made specifically for analysis within ArcGIS. The point shapefile attributes include the latitude and longitude in decimal degrees with six decimal places, the type of facility (i.e. hospital or clinic), and the name of the facility. A latitude and longitude in decimal degrees with six decimal places offers a precision of 0.1 meters which is certainly precise enough to tag a building. Figure 6 displays the typical accuracy of a single point location of a medical facility.

Figure 6 Sample of the Medical Facility Data showing the accuracy of the reported point location of the General Hospital of Banila

Figure 7 displays the location data for the current medical facilities. From first glance, it is apparent that the 1,086 medical facilities are tightly clustered together and not evenly distributed among the health zones or the DRC as a whole. Figure 8 shows a heat map of these facilities that illustrates this clustering.

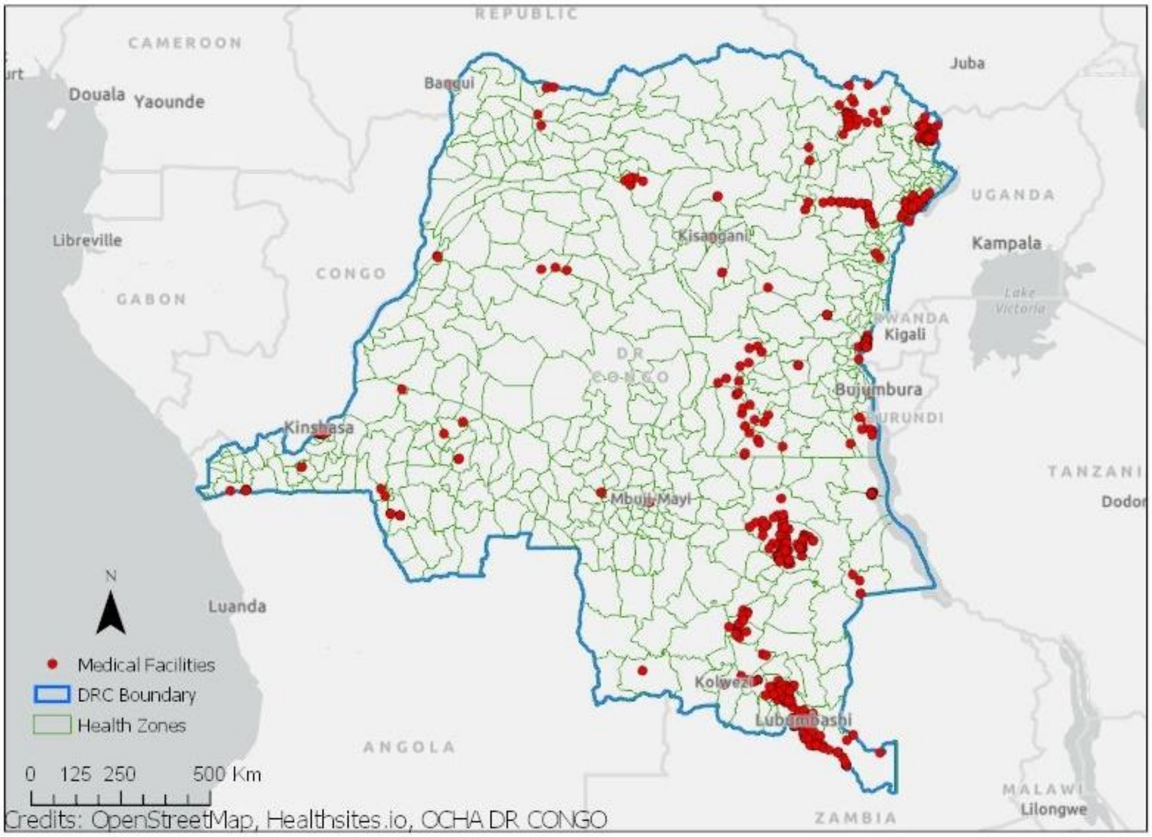

Figure 7 The locations of the 1,086 existing medical facilities within the health zones.

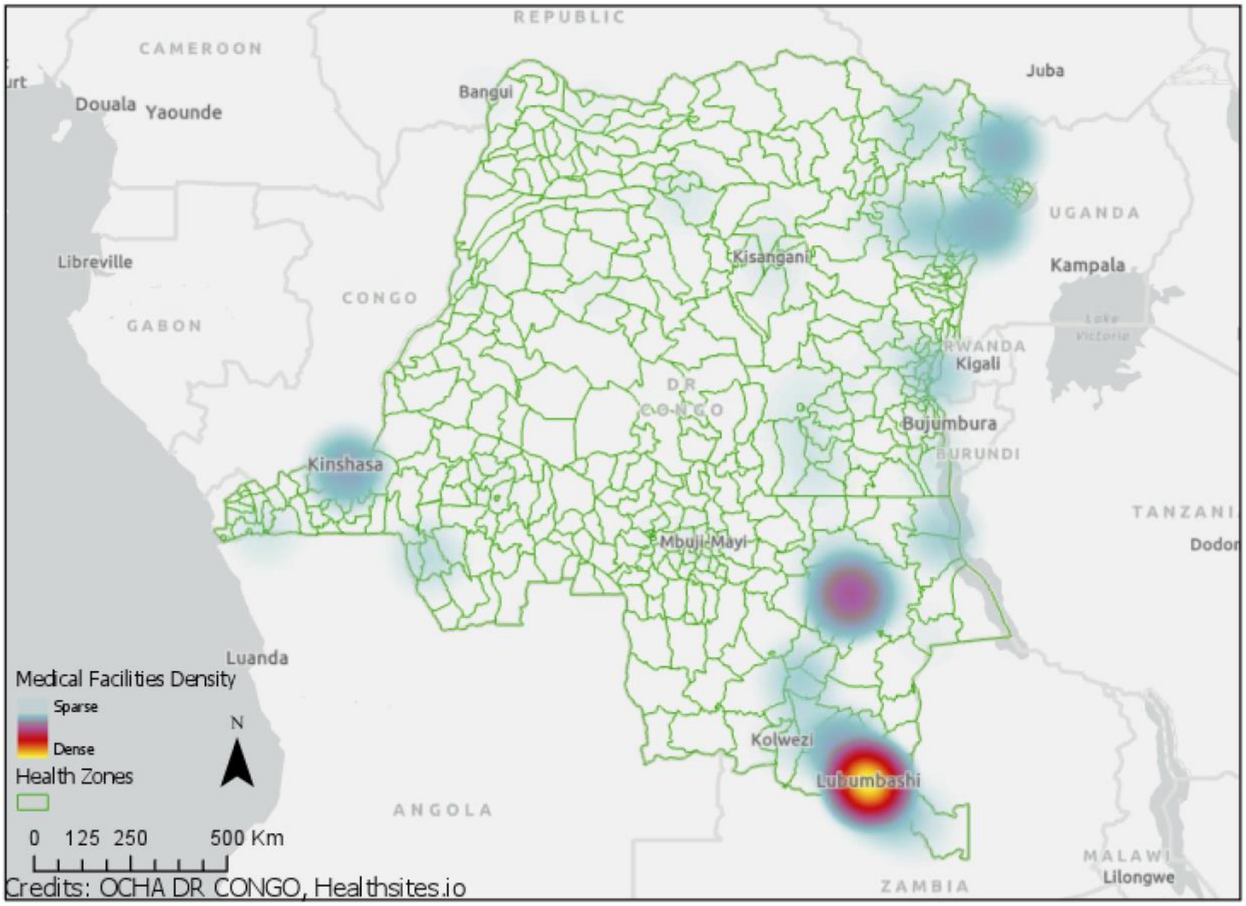

Figure 8 The 1,086 existing medical facilities displayed as a heat map to show the density of the facilities. Source: OCHA, HealthSites

3.1.5. Elevation

The digital elevation data called the Global Multi-resolution Terrain Elevation Data (GMTED2010) was obtained from the United States Geological Survey (USGS). This data set was created in collaboration with the National Geospatial-Intelligence Agency (NGA) and was obtained at a 30-arcsecond spatial resolution. The cell size is 0.0083 degrees (or approximately 928 meters at this latitude) and the datum is WGS 1984. The USGS and NGA's sources for the elevation data were the Digital Terrain Elevation Data (DTED) from the Shuttle Radar Topography Mission (SRTM), Canadian elevation data, Spot 5 Reference3D data, and data from the Ice, Cloud, and land Elevation Satellite (ICESat). Figure 9 shows the elevation distribution of the DRC. The DRC is part of the Congo River Basin and a majority of the central portion of the country is the low lying basin. This basin area increases in elevation as land moves east to the inland part of the country. The basin area primarily consists of rainforests, rivers, terraces, savannahs, and graslands.

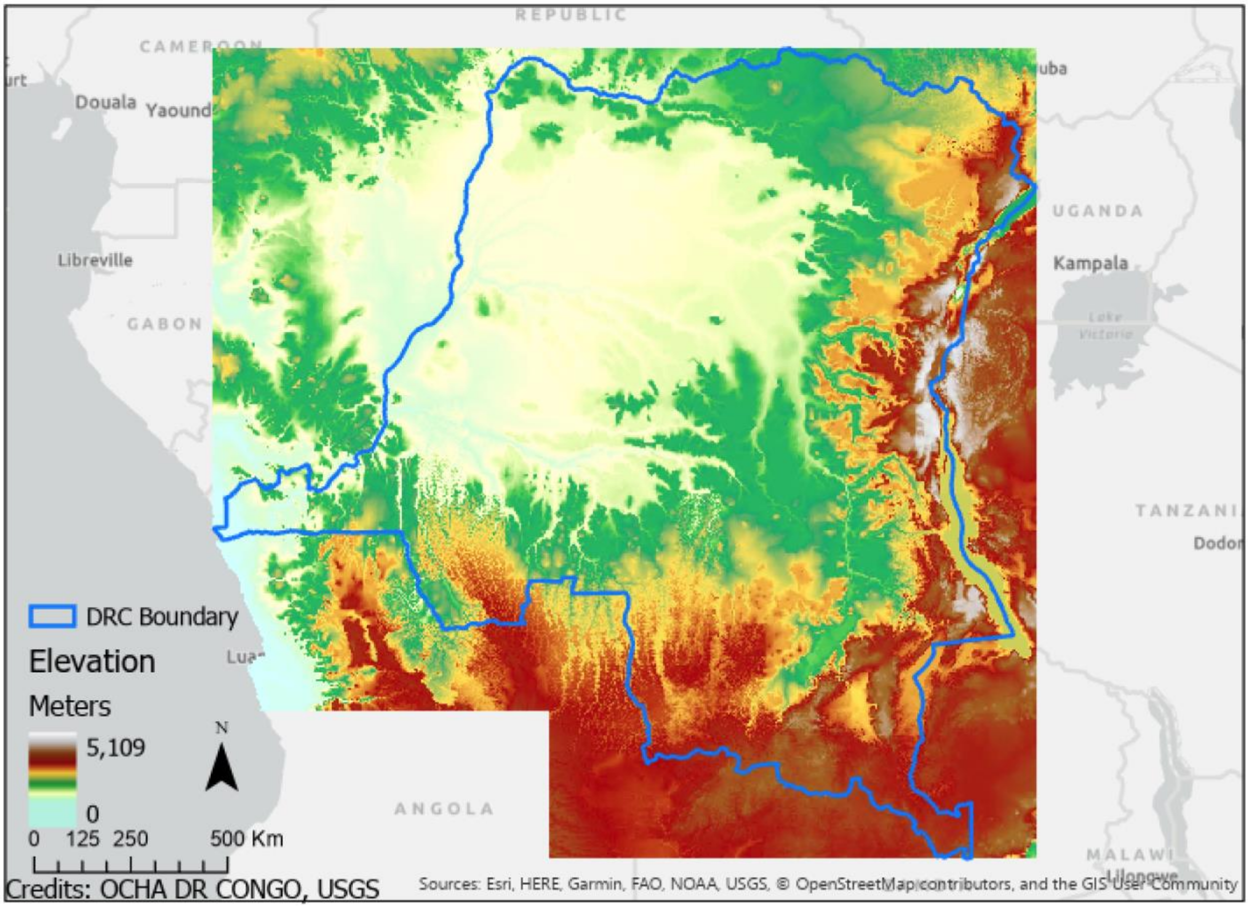

Figure 9 Elevation data of the DRC. Source: USGS

3.1.6. Land Cover

Land cover data were obtained from the Global Land Cover Classification (GLCC) created by the USGS in 1993. The land cover from the GLCC was used because the major land types were already classified and just needed to be reclassified to suit this study. While this data set is relatively old, it was easily acquired and deemed sufficient for this project given that the focus of the study was on the methodology used to find optimal locations to implement medical facilities.

The land cover raster is a representation of land cover based on the unsupervised classification of 1-km AVHRR (Advanced Very High-Resolution Radiometer) 10-day NDVI (Normalized Difference Vegetation Index) composites. These data were available both globally and by continent, both with a spatial resolution of 1 kilometer. Figure 10 shows the land cover trimmed to the polygon boundary of the DRC. To simplify the display, the values associated with the 23 categories of land cover within the clipped area are displayed in Table 3. The original dataset in its full fidelity can be found online at the source indicated in Table 1.

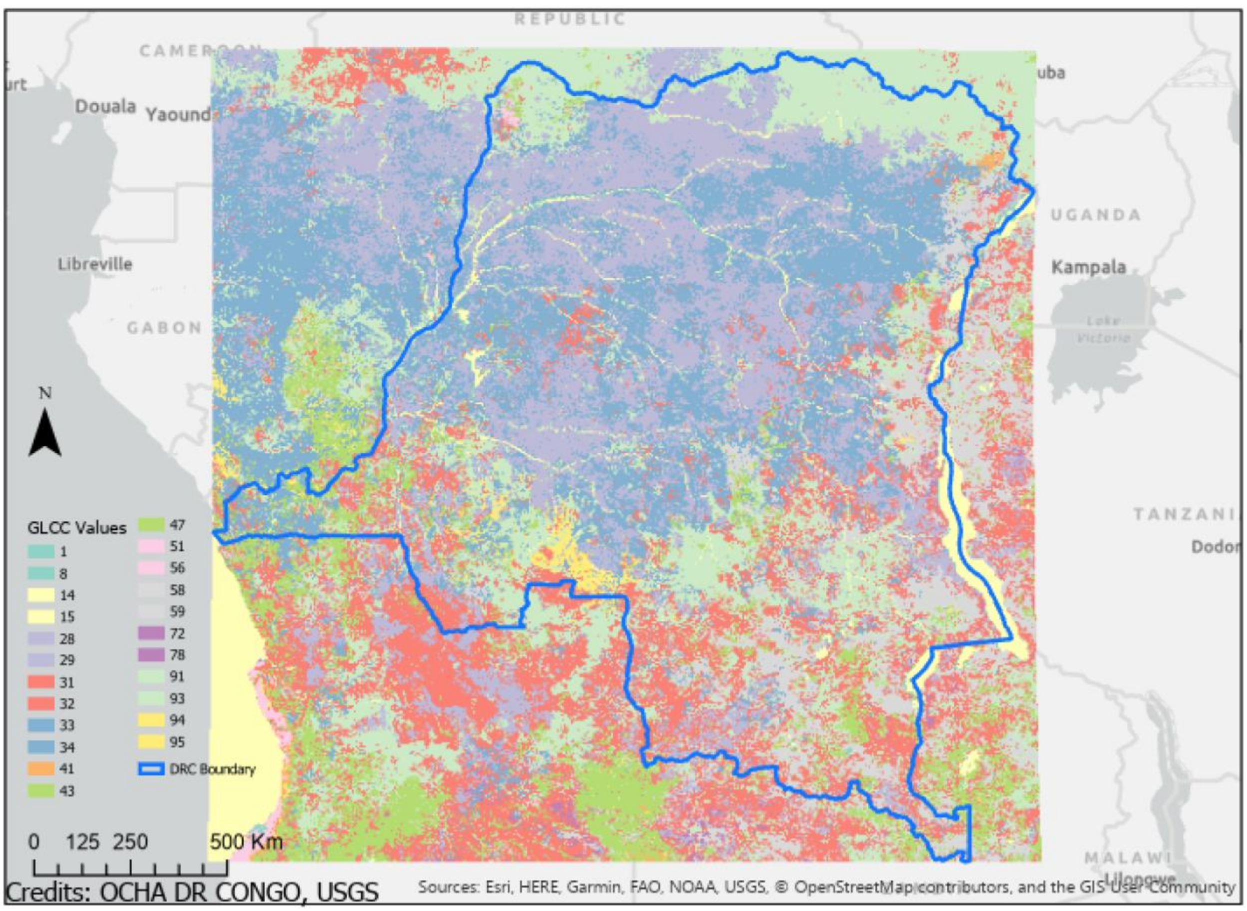

Figure 10 The GLCC raster trimmed to the boundary of the DRC. Source: USGS

Table 3 The classification of land use in the GLCC. Source: USGS

Value	Description	Value	Description
1	Urban	47	Dry Woody Scrub
8	Bare Desert	51	Semi Desert Shrubs
14	Inland Water	56	Forest and Field
15	Sea Water	58	Fields and Woody Savanna
28	Montane Tropical Forests	59	Succulent and Thorn Scrub
29	Seasonal Tropical Forest	72	Mangrove
31	Crops and Town	78	Southern Hemisphere Mixed Forest
32	Dry Tropical Woods	91	Woody Savanna
33	Tropical Rainforest	93	Grass Crops
34	Tropical Degraded Forest	94	Crops, Grass, Shrubs
41	Hot and Mild Grasses, Shrubs	95	Evergreen Tree Crop
43	Savanna (Woods)		

To use these data in the cost path analysis, it was necessary to reclassify the land cover categories to reflect the difficulty of traversing across each type. As explained below in the Methods section, the land cover data was manually reclassified based on Blanford et al.'s (2011) work to a scale where the more difficult land cover to traverse was assigned a higher value. The pixel size is 1000 meters and the datum is WGS 1984.

3.2. Methods

The following section describes in detail the methodology used to determine ideal locations to implement medical facilities based on need. Figure 11 displays the generalized workflow of the methodology used in this study. A complete model of the methodology can be seen in Appendix A.

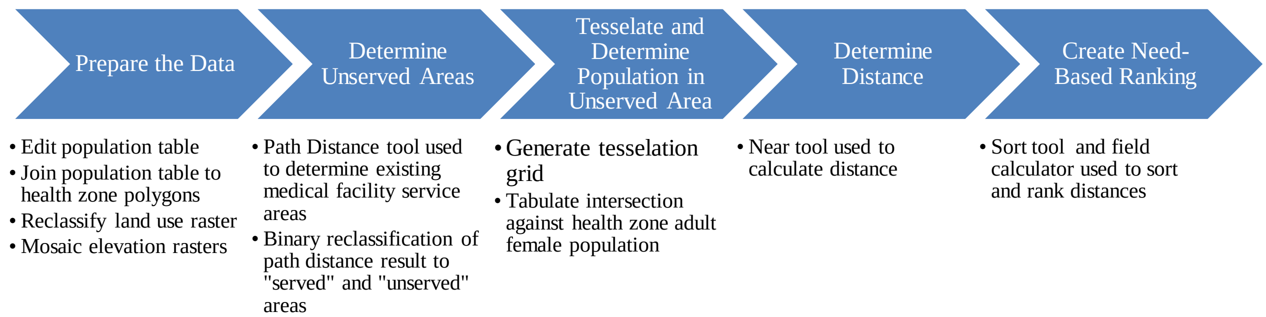

Figure 11 The workflow of the methodology used.

3.2.1. Prepare the Data

The population data table had eight columns of population data and only the adult female population data was needed for this study. The health zone code, health zone name, and the adult female population data columns were extracted. The columns were saved on to a new table and imported into ArcPro. The population table was joined to the health zones polygon boundaries by the health zone codes.

The layers were all obtained with a geographic coordinate system of WGS 84 and no projected coordinate system. In order to perform the analysis, each layer was projected to WGS 1984 UTM Zone 33S which covers the DRC well.

The land cover data were reclassified using the Reclassify tool based on published travel speeds through varying land types derived from Blanford et al. (2011). The travel speeds and associated cell crossing times through various terrains used in this study can be seen in Table 4. Since Blanford et al. did not include river travel in their model, an average speed on rivers was assumed to be the typical speed a canoe could travel on a slow-moving river, 6 km/hr. The average speed of 6km/hr was derived from an article by Munn (2006) in which he describes the speed for an advanced paddler as 6km/hr, an intermediate paddler at 4-5km/hr, and a novice paddler at 2-3km/hr (2006). The fastest speed was taken into account to determine the largest three-hour service area a person would be capable of traveling regardless of ability. More difficult terrain to traverse has a higher cost while easy terrain has a lower cost value. Using the travel times in Table 4, the value of the cell crossing time was determined by the cell size of 1000 meters and the speed to cross the land use type in the cell. Equation 1 shows the formula used to identify the cell crossing time, *t*, and Equation 2 displays the equation if the 6 km/hr terrain speed is applied for a single cell.

$$t = (1 \div Speed) \times distance \tag{1}$$

$$t = (1 \div 6) \times 1 \tag{2}$$

Table 4 Land use reclassification based on travel speed. Source: Blanford et al. 2011

Land Cover Type	Average Speed (km/hr)	Cell Crossing Time (hr)
Rivers	6 km	0.16667
Open or sparse grasslands, croplands, urban areas	3 km	0.33333
Shrubland, woodland, desert	1.5 km	0.66667
Lowland forest, swamp	1 km	1

The reclassified land use layer was then clipped to the extent of the DRC utilizing the Clip Raster tool and the polygon layer of the DRC boundary.

The elevation data came as three separate rasters, so it was necessary to use the Mosaic to New Raster function to merge the three rasters together to be used as one. The raster was then clipped to the DRC boundary using the Clip Raster tool. The elevation raster then had to have the cell sizes altered to match the land use raster. The Resample tool was used with the output cell size set to the land use raster at 1000 meters. The bilinear resampling technique was used due to the fact that bilinear interpolation calculates the value of each pixel by averaging (weighted for distance) the values of the surrounding four pixels and it is suitable for continuous data.

3.2.2. Determine the Unserved Areas of the DRC

In order to determine the unserved areas, it was necessary first to identify the areas that are served by medical facilities. The conventional approach, such as that used by Becker (2018), for determining the service area of medical facilities using the ArcGIS Network Analyst Generate Service Areas tool was not feasible due to a lack of a transportation network in the country since its civil wars. The most common form of travel in the DRC is walking. Thus, a service area around existing facilities of no more than three-hour walking distance was determined to be the most feasible for a pregnant woman. At the fastest walking speed of 3 km per hour (Blanford et al. 2011), this would be a maximum distance of 9 km. Using this logic, a medical facility located at a distance greater than a three-hour walking distance during a woman's pregnancy or labor is deemed inaccessible.

The Path Distance tool was used to determine the area of the country that is beyond a three-hour walk from any medical facility. The Path Distance tool determines the minimum cumulative travel cost from a source, in this case a medical facility, to each cell location on a raster while accounting for impedance by land use and slope. The reclassified land use raster defines the impedance or cost to move planimetrically through each cell. The value at each cell location represents the cost-per-unit distance for moving through the cell. Each cell location value is multiplied by the cell resolution while also compensating for distance added by crossing cells diagonally (where direction of travel is indicated by the tool's pre-calculated backlink raster) and moving across a slope (as determined by the tool's pre-calculated slope, calculated from the user supplied elevation raster), to obtain the total cost (i.e. time) of passing through the cell. By incorporating the elevation raster, the Path Distance tool also takes into account the actual surface distance that must be traveled as well as the difficulty of walking up or down a slope by utilizing values from the Tobler's Hiking Function which determines the difficulty of walking on slopes. By setting river impedance at the highest speed (hence lowest time to cross a river cell), this tool optimizes for river travel when available, which is a primary mode of transportation within the DRC, secondary to walking.

In the absence of local knowledge, a number of assumptions were made regarding how people move across this landscape. Boat travel on rivers was determined to be the fastest means of travel as most rivers are navigable. Bridges were not taken into account in determining the cost distance given the badly degraded condition of the transportation network. Any functional bridges are most likely to be small foot bridges whose locations quickly change. It was also assumed that crossing larger rivers by canoe would be relatively easy to accomplish within the 1 km resolution of a grid cell. Finally, based on the low impedance value given to rivers and the prevalence of river travel in the DRC, the tool will choose river routes to minimize travel time by balancing time on rivers against time spent walking overland.

The output path distance raster was Reclassified to either 1 (served) or 0 (unserved) to indicate areas within a three-hour walking distance from the medical facility locations and areas beyond a three-hour walking distance. The reclassified raster output of the path distance tool was then converted to polygons using the Raster to Polygon tool. Next, the polygons with attribute 1 (served) were selected using Select By Attributes and then the selection was deleted. This new layer was then restricted to the boundary of the DRC using the Clip tool. This result is the 'Unserved Areas of the DRC' and displays the extent of the country that is beyond a three hour walk from any medical facility.

3.2.3. Generate Tessellation and Calculate Population

Next, the Generate Tessellation tool was used to create a grid of hexagons to cover the extent of the Unserved Areas of the DRC. The tessellation grid size chosen was 210.44 sq kilometers, which produces hexagons with a circumcircle radius of 9km. At a 3km average walking speed, if any new medical facility were placed at the center of the hexagon it would be a 3 hour walk to the extent of the hexagon. Each hexagon serves as an individual potential service area for any new medical facility. When the grid was generated, each hexagon was assigned a GRID_ID with letters starting at AA being the columns and the rows were numbers.

The Clip tool was then used on the hexagons using the Unserved Areas of the DRC polygon to limit the hexagon coverage to only the unserved areas. In doing this, the hexagons that extend beyond the DRC boundary or cover the existing service areas would be omitted or partially clipped resulting in 9,911 full or partial hexagons in the Unserved Areas.

Next, to determine an estimate of the adult female population within each hexagon the Tabulate Intersection tool was used. The 9,911 hexagons that cover the Unserved Area of the DRC were used as the Input Zone Features. The health zones that were joined with the adult female population data were used as the Input Class features. The Adult Female Population was designated as a sum field. The result is a generated table of the hexagons that list the estimated adult female population count based on the percentage of area the hexagon occupies in a health zone.

The 9,911 hexagons that cover the Unserved Area were joined with the generated table based on the hexagon GRID_ID. Table 3 displays a sample of the adult female population data by hexagon. The result is the original 9,911 hexagons but it now includes an attribute that is the adult female population for the percent of the health zone the hexagon occupies.

Table 5 A sample of the calculated adult female population counts within each hexagon by GRID_ID

GRID_ID	Adult Female Pop
AA-70	36,224
AA-71	3,844
AA-72	1,409
AA-73	3,238
AA-74	2,893
AA-75	1,973
AA-76	1,315
AA-77	1,165
AA-78	1,717
AA-79	1,812

3.2.4. Determine Distance

To determine the ranking of hexagon locations that are not accessible medical facilities, the Near tool was used to calculate the distance between the edge of each hexagon and the closest medical facility point location. The Near tool added the nearest medical facility's ObjectID and the distance to the nearest medical facility to the hexagon grid attribute table.

3.2.5. Create Need-Based Ranking

Next, to assign ranks to individual hexagons, the Sort tool was used on the hexagon grid layer attribute table to sort by distance from the nearest medical facility in ascending order. A new field called Distance Rank was added to the hexagon grid attribute table, and the Field Calculator was used to generate a sequential number based on the sorted near distance column. The closest hexagon has a value of 1 and the furthest hexagon from a medical facility has the highest number.

The Sort tool was used again but now to sort the population attribute in ascending order. A field was again added and called Population Rank. The Field Calculator was used based on the ascending sort order of the population to generate sequential numbers in the population rank field. The lowest population count within a grid hexagon has the rank of 1 and the highest population count has the highest rank. Thus, all hexagons are ranked by both distance and population.

To create a final needs-based rank, it was necessary to decide what metric should be used to combine these two rankings. There are many ways this could be done, many of them very complex computationally, but for the demonstration purpose of this study, it was decided that a mean of the two ranks would provide a value that balances the criteria evenly. Thus, to create the final needs-based rank of the hexagons, a field was added to the layer's attribute table called Mean. The Field Calculator was used to calculate the mean of the Population Rank and Distance Rank for each hexagon. The Sort tool was used on the Mean field, this time in descending order. Since the previous sorts gave the highest population and the furthest distances the highest values, these high values produce a high mean and these should be ranked highest (closer to 1) because it means the hexagon would likely be far from medical facilities and have a higher population. A final field was added to the attribute table called Needs-Based Rank and a sequential number was populated in the Needs-Based Rank field based on the descending order of the Areas in Needs field.

From the results, the top (lowest rank) location will be the most ideal for a new medical facility. If ten, twenty, or fifty new medical facilities were to be added, the service area rank and coverage should ideally be recalculated after the addition of each new medical facility to optimize coverage. Within the selected hexagons, a facility can be constructed at the center point as the center point is 9 kilometers, or a 3 hour walk, from the edge. Each hexagon is its own service area. However, once constructed in one location, all adjacent hexagons would have a much closer distance to an existing facility and in the combined ranking their ranks would drop (i.e. grow larger). All the ranks throughout the unserved areas would adjust.

Chapter 4 Results

The intent of this study was to develop a methodology for identifying the optimal locations to create new medical facilities or birth centers in order to reduce maternal mortality. The ideal location to implement new medical facilities is based on population that could be served within a three-hour walking distance service area and distance from existing medical facilities. This chapter presents and discusses the results of each step in the methodology.

4.1. Intermediate Results

There were several stages in the methodology, and each produced an intermediate result. These are shown in the following sub-sections.

4.1.1. Determine the Unserved Areas of the DRC

In order to identify the Unserved Areas of the DRC, identifying and eliminating the service areas of the existing medical facilities needed to occur first. To identify the three-hour walking distance service areas of the medical facilities the Path Distance tool was used. A cost distance raster was prepared and reclassified based on walking speeds through various terrains for use in the Path Distance tool. Figure 12 displays the results of the reclassification of the land use based on travel speed. From the results, it is observed that most of the country is shown to have either a 3km or 1km travel speed. The lack of variation in land use travel speed likely did not alter the results in determining the 3-hour walking distance service area much more than a simple Euclidean Distance calculation might have provided. However, this study is on the development of a methodology using the DRC as the study area. The same methodology could be used in an alternate country that has more variation in their land use travel speed.

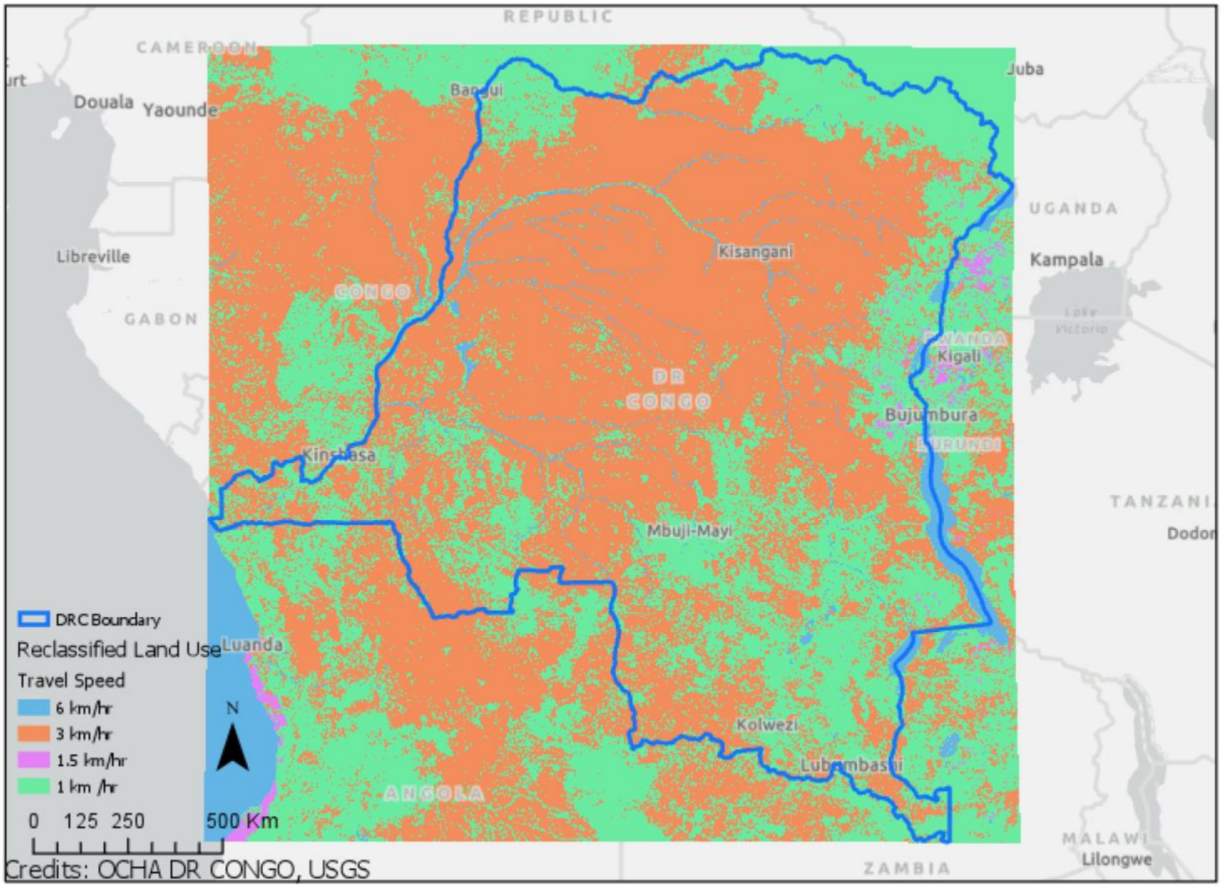

Figure 12 The land cover raster reclassified based on travel speed.

The result of the Path Distance tool was a raster that was then reclassified to two categories: within a three-hour walking distance and beyond a three-hour walking distance and converted to polygons. This result is the ‘Unserved Areas of the DRC’ (Figure 13). This makes it very clear the vast extent of the country that is beyond a three-hour walk from any medical facility.

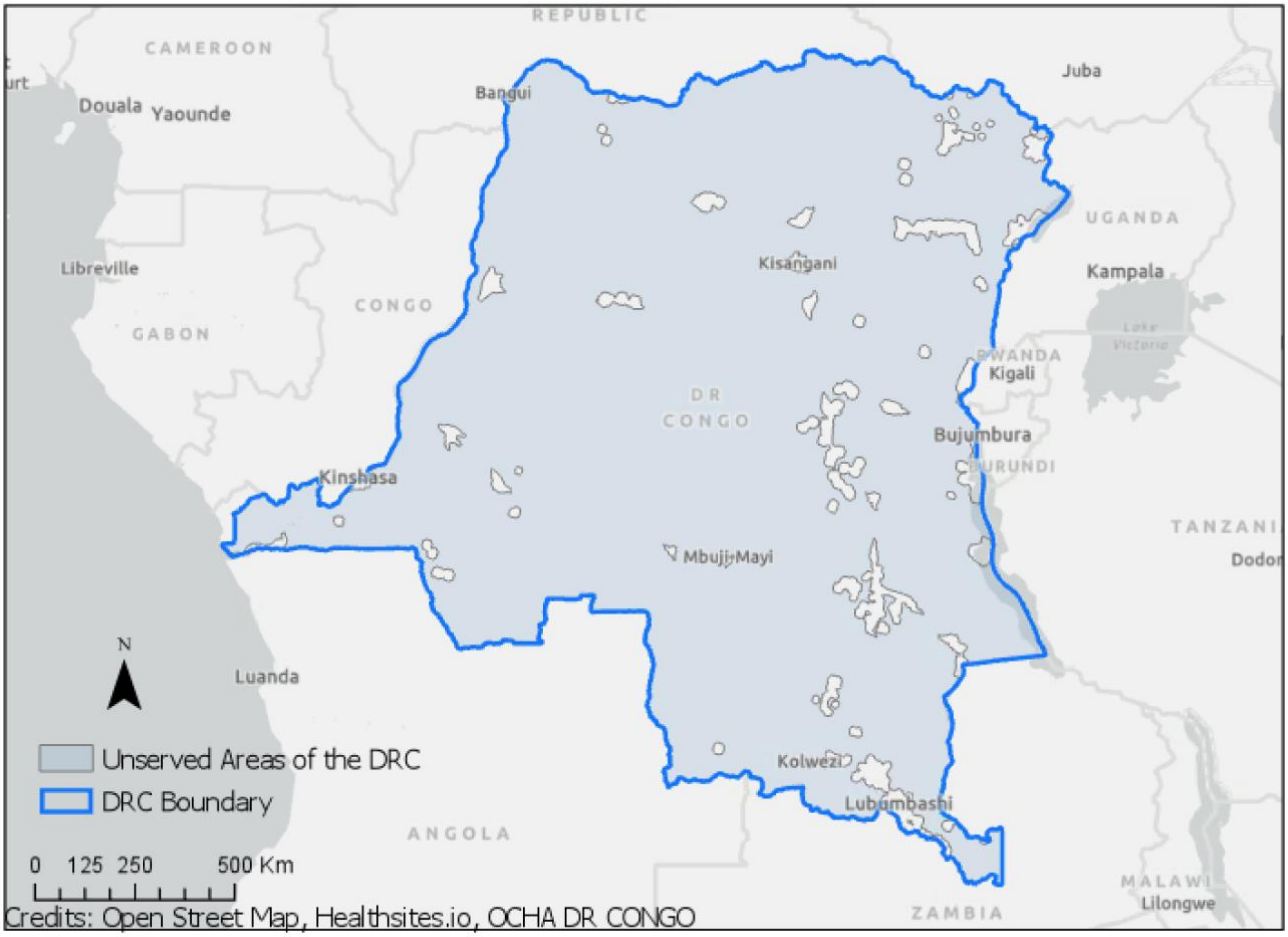

Figure 13 The results of the path distance tool display the areas of the DRC that are beyond a 3-hour walk from a medical facility and are considered Unserved Areas in the DRC

4.1.2. Generate Tessellation and Calculate Population

The Generate Tessellation tool created a grid of 9,911 hexagons to cover the extent of the Unserved Areas of the DRC (Figure 14).

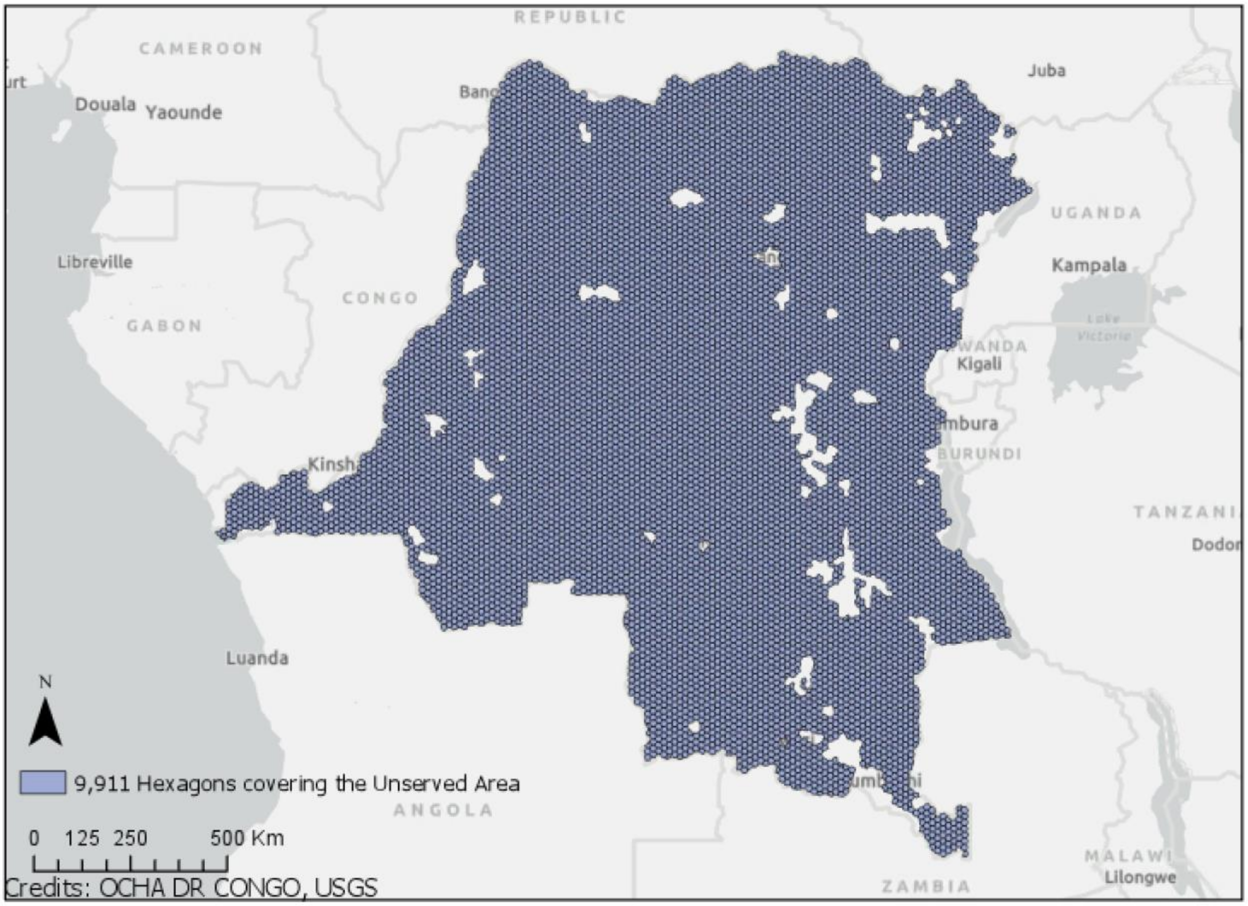

Figure 14 The 9,911 generated hexagons covering the Unserved Area of the DRC.

After the hexagon grid was established for the Unserved Area of the DRC, the population that could potentially be served within each hexagon was determined using the Tabulate Intersection tool. After joining the intersection output table to the hexagon grid based on GRID_ID, the results of the Adult Female population per hexagon can be seen in Figure 15.

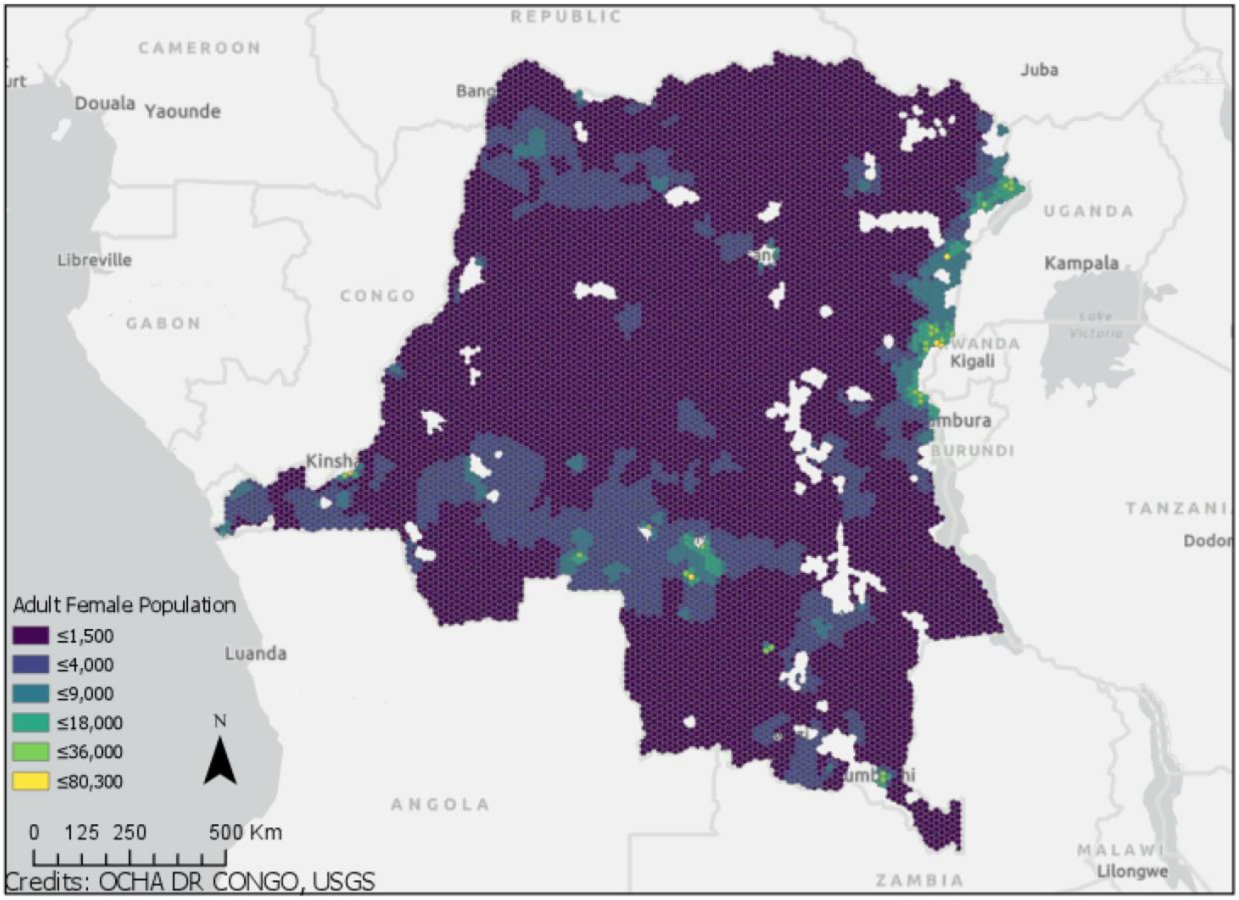

Figure 15 Distribution of the estimated adult female population by hexagon.

The sum of the adult female population in the DRC is approximately 17.6 million while the sum of the adult female population within the grids of the unserved area is 12.5 million. These sums show that only about 5.1 million adult females in the DRC are within a three hour walking distance of a medical facility, which is less than 1/3 of the country being served. Table 5 displays a sample of the adult female population data by grid. While a portion of the population is located near medical facilities on the east side of the country, a large number of the adult female population are located in remote rural areas of the south central.

4.1.3. Create Need-Based Ranking

Figure 16 is a visualization of the Adult Female Population Rank. A high rank in this map indicates high population numbers. The hexagons with the higher population counts are in yellow.

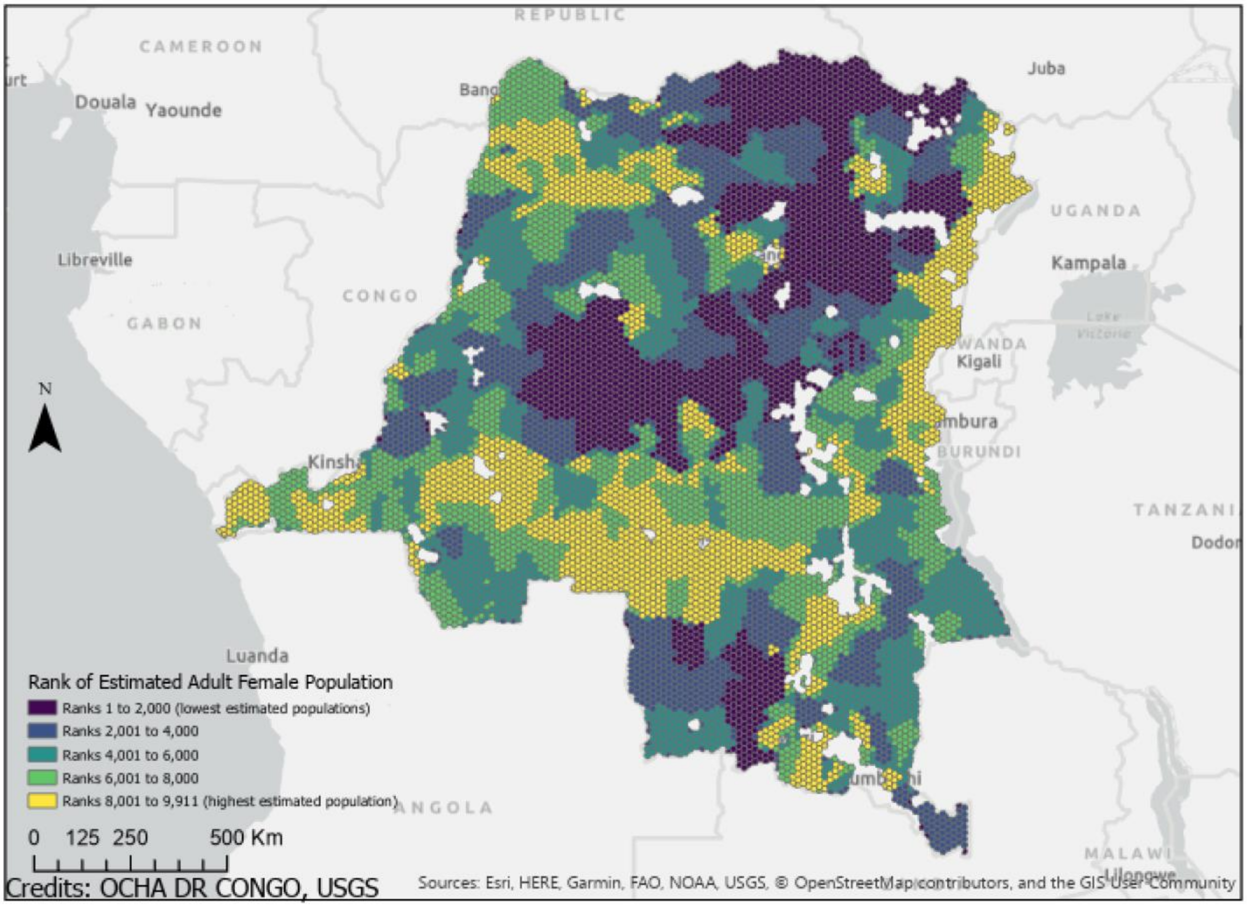

Figure 16 Visualization of the Adult Female Population Rank. Recall that rank closer to 1 in this map indicates hexagons with the lowest population counts. The hexagons with the highest population counts are in yellow.

4.1.4. Determine Distance

To determine the ranking of hexagon locations that are not accessible to medical facilities, the Near tool was used to calculate the distance between the edge of each hexagon and the closest medical facility point location. Figure 17 displays a visualization of the ranked distance of hexagons from the existing medical facilities. The closest hexagon to a medical facility has a rank of 1 and the furthest hexagon from an existing medical facility has a rank of 9,911. The hexagons that are the furthest distance from medical facilities are in yellow.

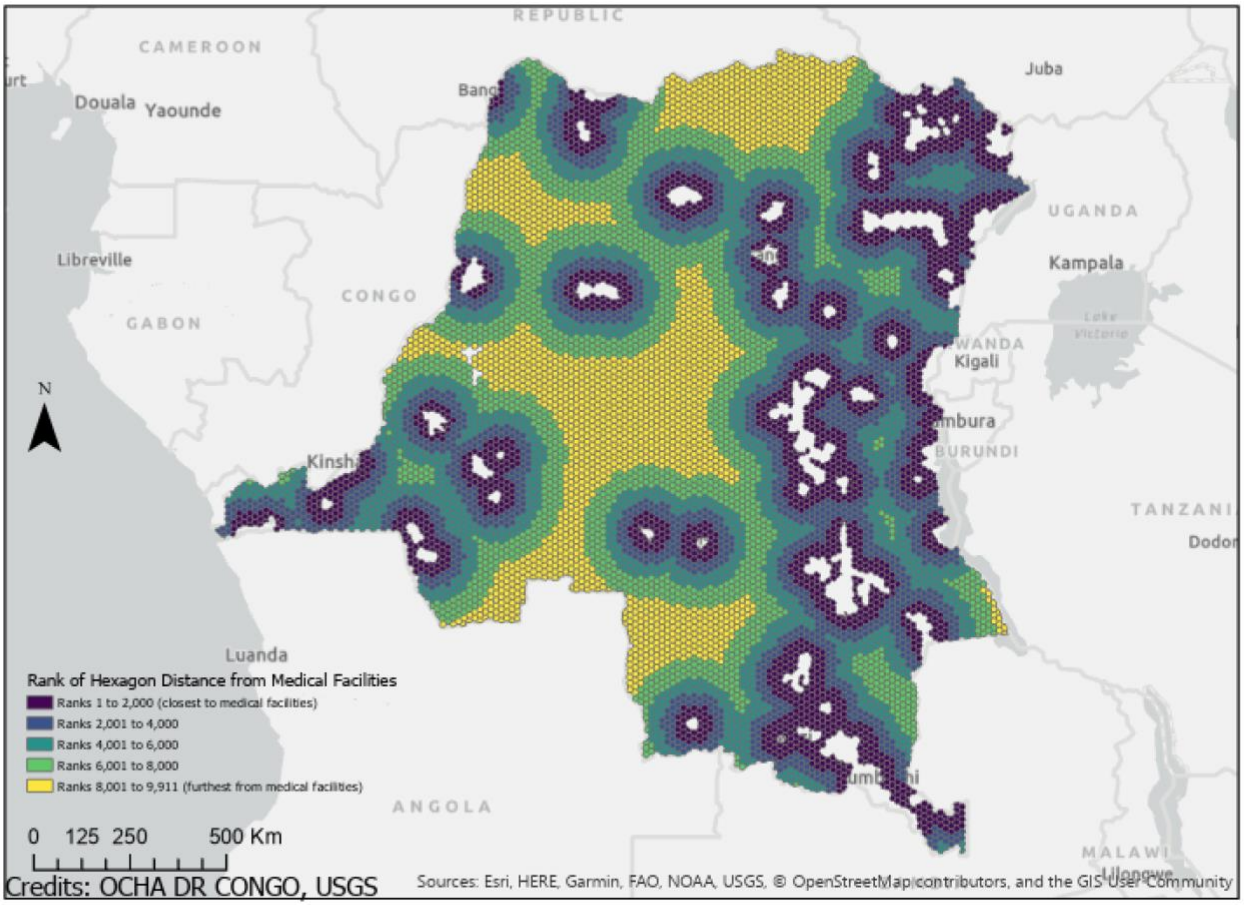

Figure 17 Visualization of the ranked distance of the 9,911 hexagons from the medical facilities. A further distance is a higher rank. The hexagons with the furthest distance from medical facilities are in yellow.

4.2. Final Needs-Based Ranking

The results of the overall ranking of the hexagons based on estimated adult female population and the distance from existing medical facilities can be seen in Figure 18. Low ranks mean a combination of high distance and high population. The results show that the areas displayed in yellow, primarily the west and south-central portion of the country, are in the most need for medical facilities.

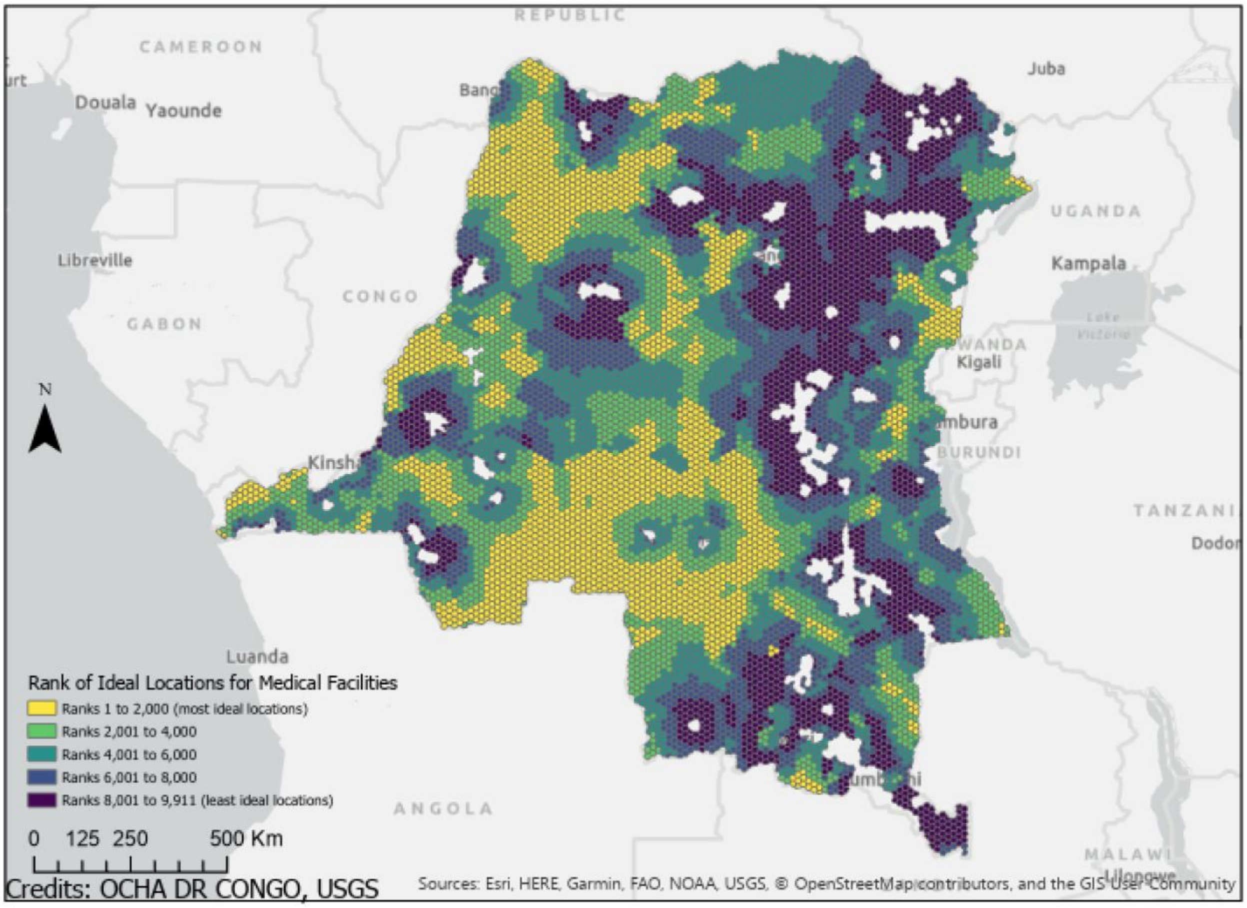

Figure 18 Visualization of the hexagons based on need for a medical facility. The result is the mean of the population rank and the distance rank.

Recall from Figure 15 that a large adult female population is present in the south-central portion of the DRC. This area also appears as high ranking (low value rank) for need of a facility in the resulting map in Figure 18 and would, thus, seem to be the optimal first place to implement a new medical facility. Additional facilities might be evenly distributed within the yellow areas of the need-based map in Figure 18.

As a way to evaluate the validity of the mean rank index used, histograms of the hexagon distance and population data as well as a scatter plot of the relationship between the two were evaluated. As would be expected, that the data was not evenly or normally distributed due outliers in the data at the high end of both distributions. Figure 19 displays the distribution of the distances of the hexagons from existing medical facilities. The majority of the hexagons are between 20 and 125 km away and only a few are at much greater distances.

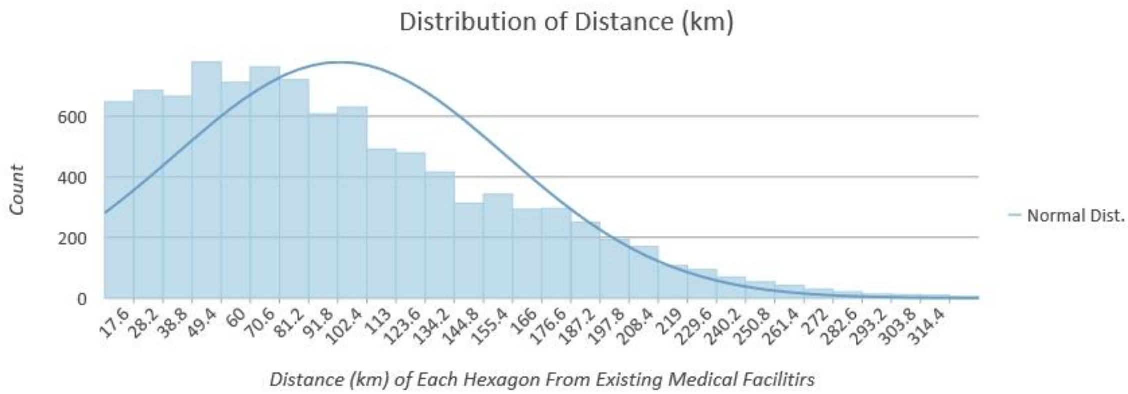

Figure 19 Histogram displaying the distance in kilometers of each hexagon from existing medical facility.

The histogram of the distribution of the adult female population can be seen in Figure 20. The data is not normally distributed due to outliers at the extreme high end where some hexagons are located in very urban areas with high populations and a large number of hexagons with a population estimate of under 1,000 due to large rural areas of the country.

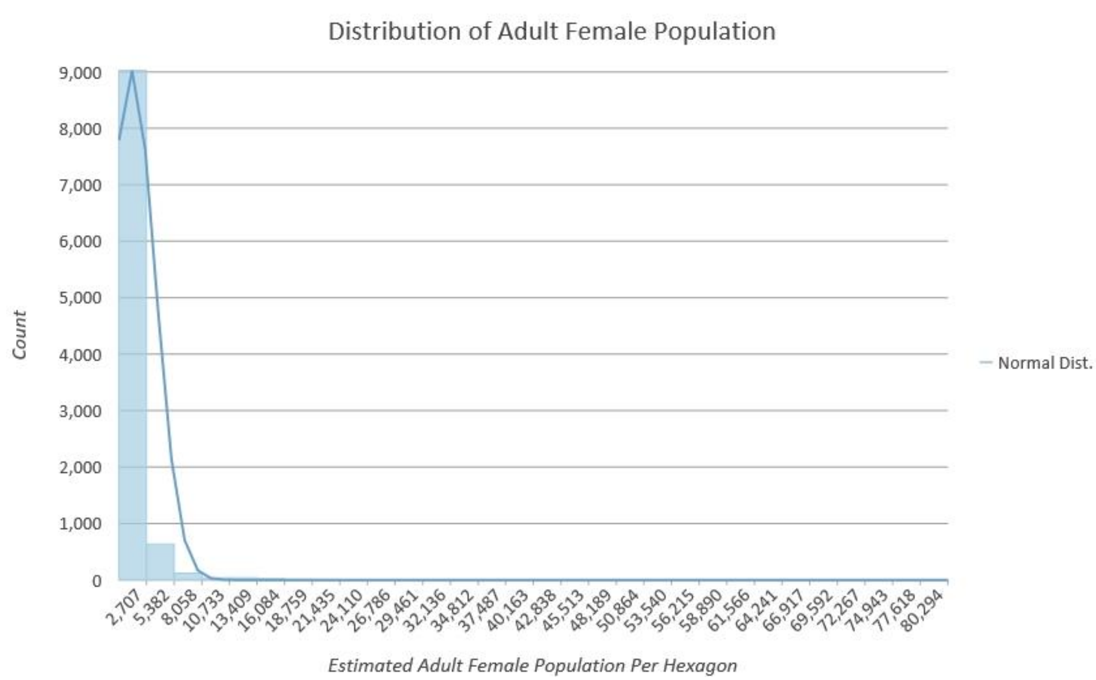

Figure 20 Histogram displaying the estimated adult female population per hexagon

Figure 21 is a revaluation of the adult female population histogram with the adult female population data normalized with a logarithmic transformation to account for the outliers at the high end. The logarithmic transformation squeezes together the larger values in the data set and stretches out the smaller values. Logarithmic transformations are used to make highly skewed distributions less skewed. This can be valuable both for making patterns in the data more interpretable and for helping to meet the assumptions of inferential statistics.

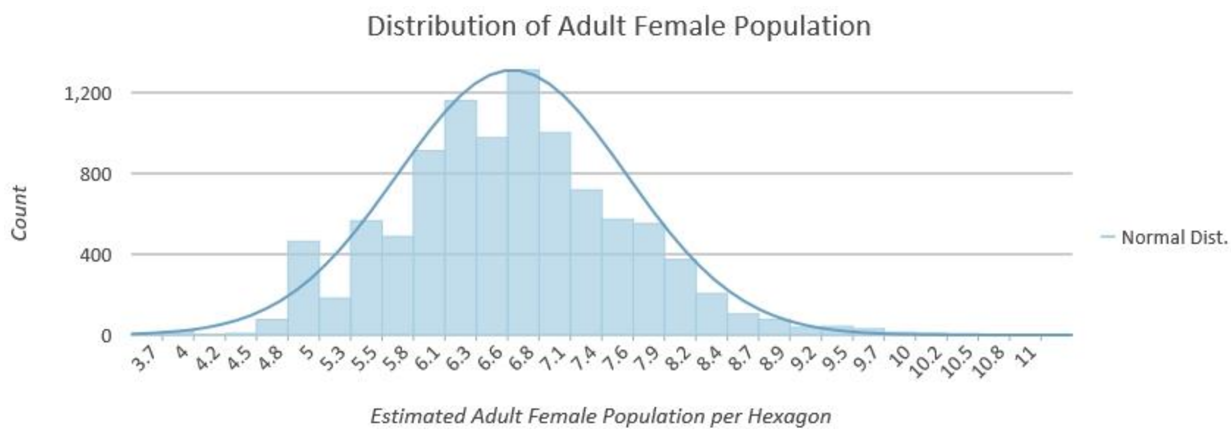

Figure 21 Histogram displaying the estimated adult female population per hexagon with the adult female population data normalized

Scatter plots display the relationship between two variables. Figure 22 displays the relationship between estimated adult female population and distance from existing medical facilities. Due to the outliers at the high end of the data, the scatter plot is condensed and practically unreadable. Figure 23 displayed the scatter plot relationship with one axis, the adult female population, normalized using the logarithmic transformation so that the distribution could be more easily viewed and understood.

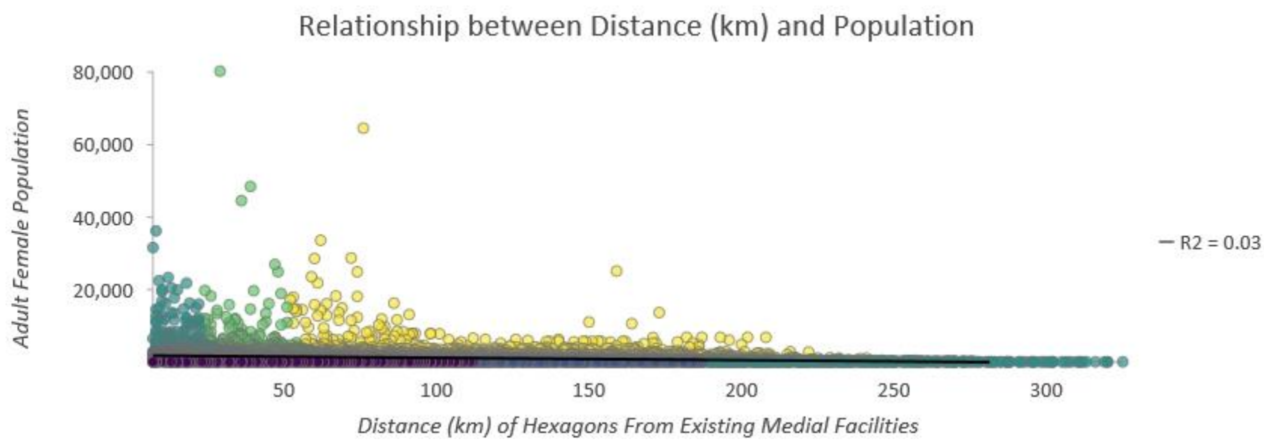

Figure 22 Scatter plot displaying the relationship between the estimated adult female population count within a hexagon and the distance a hexagon is from existing medical facilities. Colors indicate the hexagons' final ranks as displayed in the maps above.

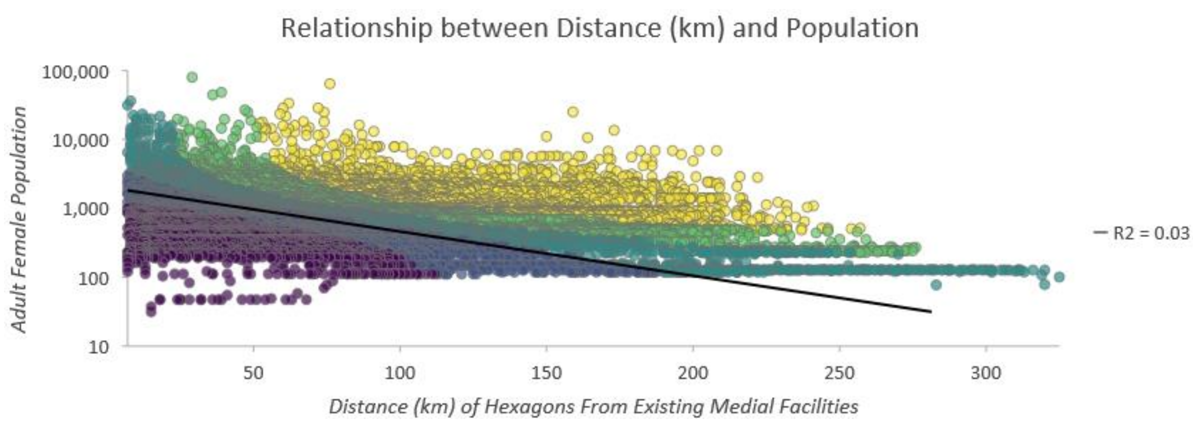

Figure 23 Scatter plot displaying the relationship between the estimated adult female population count within a hexagon and the distance a hexagon is from existing medical facilities. The Y axis (adult female population) has been log transformed. Colors indicate the hexagons' final ranks as displayed in the maps above.

The last scatter plot in Figure 24 displays the relationship between the two variables with both axes having a logarithmic transformation. The logarithmic transformation makes it evident that the mean of the two sets of data produced the intended result. As both the distance and the population increase (moving to the upper right), hexagons are correctly categorized in the highest rank, being most in need of medical facilities. Areas that are closer with a smaller population are in the bottom left in dark blue. The colors of the scatter plots reflect the final ranks of hexagons displayed in Figure 18 where yellow is the most desirable location to build medical facilities and the dark blue is the least desirable.

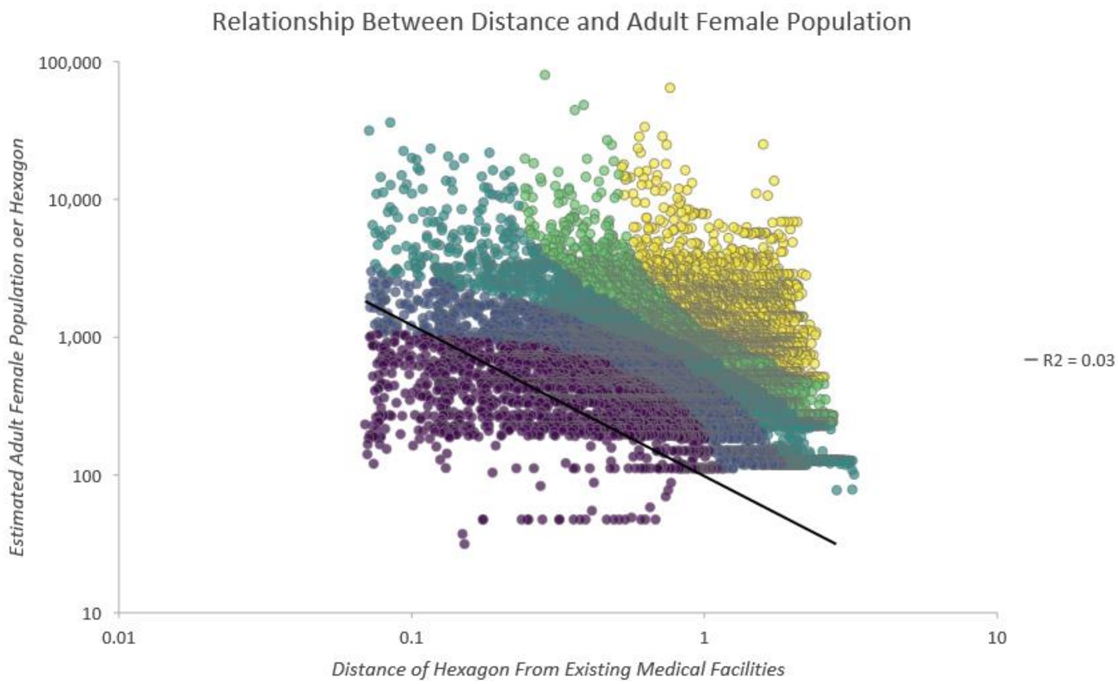

Figure 24 Scatter plot displaying the relationship between the estimated adult female population count within a hexagon and the distance a hexagon is from existing medical facilities. Both axes have been log transformed.

Chapter 5 Discussion and Conclusion

The objective of this study was to develop a simple methodology that could be used to identify ideal locations for new birth facilities where large populations are the furthest distance from existing medical facilities. This methodology was implemented in the study area of the Democratic Republic of the Congo but could be used in any similar rural, largely undeveloped country needing to identify areas in need of medical facilities.

The methodology developed was successful in identifying the areas of the DRC that are currently unserved by existing medical facilities. There were a few reasons the path distance tool was used to first identify the service areas of the existing facilities instead of just omitting the hexagons that had an existing facility within it. First, the existing facilities would not fall perfectly in the center of a hexagon and its service area may encompass parts of two or three hexagons. Second, this was a methodology developed that could be used in other countries with different parameters. If another country had a 5-hour walking distance service area for already present medical facilities, it could be easily identified without altering the hexagon grid size used here. Further, the study was successful in identifying areas that have the highest populations and areas that are the furthest distance from existing medical facilities. Most importantly, the study was successful in combining both distance from medical facilities and estimated population to determine areas most in need of medical facilities.

The locations were identified through the generation of a tessellation grid over the areas of the DRC with low walking accessibility to medical facilities. For each tessellation grid cell, the distance to the nearest medical facility and the population within the cell were calculated. Based on this combination of population size and distance from a medical facility, a rank of locations for new facilities was created. Facilities built in these highest ranked locations would have the maximum impact by supporting the highest population that is the furthest distance from medical facilities. The resulting increase in medical accessibility could greatly decrease birth complications and preventable death.

5.1. Lessons Learned

When I began this study, I initially chose the Democratic Republic of the Congo as my study area due to the fact that the maternal mortality rate was so high. In hindsight, if I had done more discovery into the country prior to committing so much time and effort into the research, I would have discovered earlier that the entire country lacked a usable road network which drastically altered my original plan of utilizing the Service Areas tool. Had I been able to use this tool to identify the service network of a medical facility in say a state in the United States where personally owned vehicles are a primary means of travel then the service areas would have been more accurate and a lot easier to find. I made my research a lot more difficult by committing to the country that chose as a study area.

Through the data discovery phase, it was revealed that the existing medical facilities do not adequately service the population of the DRC nor do they provide any significant amount of service area coverage. In a country that has been unable to reconstruct their road infrastructure in the course of three decades since their civil wars and where mothers are routinely dying from preventable causes, it would be expected that their current medical support is extremely and critically limited. Such a detailed analysis was perhaps not needed for the DRC given the enormity of need.

5.2. Improvements for Future Applications of the Methodology

A different approach could have been taken with the population data. The readily available source population data was aggregated by the health zones and the population of each hexagon was estimated based on the proportion of a health zone the hexagon occupied. There is also a dasymetric-based population estimate available from Oak Ridge National Laboratory, the LandScan population grid, which comes in 1 km spatial resolution and represents an ambient population (average over 24 hours) distribution. This scale would have been smaller for a more precise estimate, and since the dasymetric estimation accounts for areas that are unlikely to have residents (such as lakes or preserved areas) it would have likely produced a more accurate hexagon population estimate than the DRC health care areas. In using the population data at the health zone level and estimating the population within the hexagons based on the percentage of the health zone the hexagon occupies, no hexagons had zero population. This likely is not the case as there are large areas of rivers and open greenspace that are not occupied. However, for the proof of concept that was the goal of this study, the health zone population data was sufficient and the more detailed categories in the health zones data allowed the analysis to focus only on adult females, the target of maternal mortality reduction efforts.

Additionally, the data utilized for the medical facility location data was obtained primarily from crowd sourcing. While this data goes through multiple verification processes, the data at times is incomplete, could potentially be inaccurate, and the update rate of the data is not regulated so a facility could close or open and it could be unreported. Another source of validated medical facility data in Sub-Saharan Africa is the spatial database developed by Maina et al. (2019). This comprehensive spatial database was assembled from government sources such as Ministries of Health or government-developed Master Facility Lists. However, only public, government-owned facilities are included in the Maina et al. database, unlike the healthsite.io data that include private facilities. What type of data is more important to a particular study should dictate which data is best to use. I chose to use the healthsites.io data, and while it might not be the most accurate, it did provide the largest list of potential facilities. Since my goal was to identify unserved areas of any medical facility, including private facilities and small clinics was more important than only referencing government-owned public services.

The bridges and boat crossings for rivers were not factored in at any different travel time because it was assumed that most bridges and boat crossings are not permanent locations in the DRC. Most bridges are temporary structures made from natural materials. If this study were conducted in a country with permanent bridge locations and dedicated ferry docking points, then this could be factored into the cost surface when identifying service areas.

All of the major rivers were included in the classified land use raster. These were assumed to be navigable at the same speed either when crossing or going up or down stream. To improve the accuracy of the served areas, a horizontal factor that could account for the variations in travel costs between the three modes of travel on the river might be added.

While the rivers were accounted for in the classified land use and river pixels were assigned a fast travel speed, they are represented as disconnected pixels. To further improve the analysis, a polygon layer of rivers could be used that has the accurate width of the rivers accounted for. This could be merged with the land use raster so the pixels that intersect the river polygon would obtain the river cost. This was not done in this methodology as only lines of the DRC rivers were available, so depending on the land use raster to categorize the pixels on wide rivers like the Congo was considered more reliable.

This methodology could be used for more than just identifying new locations for medical centers. This could be used to identify new locations for any new facility such as a school, a gym, or a retail store. The service area of existing locations would need to be identified and removed from the study area to find the unserved areas. Then the targeted demographic could be identified. The target demographic that has the largest population the furthest from any new type of facility would be the new ideal location.

5.3. Value of this Research

From the results, the top location we be the most ideal for a new medical facility. If ten, twenty, or fifty new medical facilities were to be added the service area rank and coverage should ideally be recalculated after the addition of each new medical facility to optimize coverage. Within the selected hexagons, a facility can be constructed at the center point as the center point is 9 kilometers, or a 3 hour walk, from the edge. Each hexagon is its own service area. However, once constructed in one location, all adjacent hexagons would have a much closer distance to an existing facility and in the combined ranking their ranks would drop (i.e. grow larger). All the ranks throughout the unserved areas would adjust.

High maternal mortality rates are caused in large part by preventable complications that could be circumvented by increasing geographic access to medical facilities for pregnant mothers. Walking long distances during pregnancy or while in labor and especially at night is a significant barrier for women seeking medical care. Improving the conditions of roads alone would not increase access to care as most of the population does not own a personal vehicle. Establishing more healthcare facilities in the rural areas to increase access to care would be the greatest benefit to women seeking care during pregnancy and labor. Establishing optimized medical facility coverage through implementing new facilities based on need in a country like the DRC requires a through consideration of all factors ranging from infrastructure, to the location of the target population, and the current facilities.

According to historian Dan Snow, the DRC has the potential to be one of the richest countries on earth, but colonialism, slavery and corruption have turned it into one of the poorest (Snow 2013). As recounted in the Related Work chapter, the women in DRC face many perils related to the high maternal mortality rate, with the level of poverty in the country being the immediate cause of medical problems endangering maternal lives (Ellison 2017). Expanding medical services to the interior of the country to cater to the rural population would aid the adult female population that lacks accessibility to the current medical facilities due to the poor infrastructure.

References

Becker, Charles. 2016. “A Spatial Analysis of Veteran Healthcare Accessibility.” Thesis, Spatial Sciences Institute. University of Southern California. https://spatial.usc.edu/wp-content/uploads/2016/07/Becker-Charles.pdf.

Blanford, Justine I., Supriya Kumar, Wei Luo, and Alan M. Maceachren. 2012. “It’s a Long, Long Walk: Accessibility to Hospitals, Maternity and Integrated Health Centers in Niger.” *International Journal of Health Geographics*, 24, 11 (1): 1–15. doi:10.1186/1476-072x-11-24.

Chong, Shanley, Michael Nelson, Roy Byun, Liz Harris, John Eastwood, and Bin Jalaludin. 2013. “Geospatial Analyses to Identify Clusters of Adverse Antenatal Factors for Targeted Interventions.” *International Journal of Health Geographics*, 46, 12 (1): 1–10. doi:10.1186/1476-072x-12-46.

Ebener, Steeve, Maria Guerra-Arias, James Campbell, Andrew J. Tatem, Allisyn C. Moran, Fiifi Amoako Johnson, Helga Fogstad, Karin Stenberg, Sarah Neal, Particia Bailey, Reid Porter, and Zoe Matthews. 2015. “The Geography of Maternal and Newborn Health: the State of the Art.” *International Journal of Health Geographics*, 19, 14 (1): 1–10. doi:10.1186/s12942-015-0012-x.

Ellison, Katherine, and Nina Martin. 2017. “Severe Complications for Women During Childbirth Are Skyrocketing - and Could Often Be Prevented.” ProPublica. December 22. https://www.propublica.org/article/severe-complications-for-women-during-childbirth-are-skyrocketing-and-could-often-be-prevented.

Gabrysch, Sabine, Simon Cousens, Jonathan Cox, and Oona M. R. Campbell. 2011. “The Influence of Distance and Level of Care on Delivery Place in Rural Zambia: A Study of Linked National Data in a Geographic Information System.” *PLoS Medicine* 8 (1). doi:10.1371/journal.pmed.1000394.

HealthSites. 2019. “About healthsites.io.” https://www.healthsites.io/#about.

Howard, Jacqueline. 2017. “Childbirth Is Killing Black Women, and Here's Why.” *CNN*. Cable News Network. November 15. https://www.cnn.com/2017/11/15/health/black-women-maternal-mortality/index.html.

Khan, Khalid S., Daniel Wojdyla, Lale Say, A. Metin Gülmezoglu, and Paul Fa Van Look. 2006. “WHO Analysis of Causes of Maternal Death: a Systematic Review.” *The Lancet* 367 (9516): 1066–74. doi:10.1016/s0140-6736(06)68397-9.

Maina, Joseph, Paul O. Ouma, Peter M. Macharia, Victor A. Alegana, Benard Mitto, Ibrahim Soce Fall, Abdisalan M. Noor, Robert W. Snow and Emelda A. Okiro. 2019. “A Spatial Database of Health Facilities Managed by the Public Health Sector in sub Saharan Africa.” *Scientific Data* 6 (1): 134.

Massey, Philip. 2011. "Reducing Maternal Mortality in Senegal: Using GIS to Identify Priority Regions for the Expansion of Human Resources for Health." *World Health & Population* 13 (2). doi:10.12927/whp.2011.22633.

Molla, Yordanos B., Barbara Rawlins, Prestige Tatenda Makanga, Marc Cunningham, Juan Eugenio Hernández Ávila, Corrine Warren Ruktanonchai, Kavita Singh, Sylvia Alford, Mira Thompson, Vika Dwivedi, Allisyn C. Moran, and Zoe Matthews. 2017. "Geographic Information System for Improving Maternal and Newborn Health: Recommendations for Policy and Programs." *BMC Pregnancy and Childbirth*, 26, 17: 1–7. doi:10.1186/s12884-016-1199-y.

Munn, Richard. 2006. "How Far Can We Go?" *Canadian Canoe Routes - Wilderness Canoe Association*, July. https://www.myccr.com/articles/how-far-can-we-go.

Nour, Nawal M. 2008. "An Introduction to Maternal Mortality." *Reviews in Obstetrics and Gynecology* 1 (2): 77–81. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2505173/#.

Ntembwa, Hyppolite Kalambay, and Wim Van Lerberghe. 2015. "Democratic Republic of the Congo: Improving Aid Coordination in the Health Sector." World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/186673/WHO_HIS_HGF_CaseStudy_15.4_eng.pdf;jsessionid=A85ED662A12F68FD5D5B8842F0D0B4D1?sequence=1.

Nwakeze, Ngozi M., and Ngianga-Bakwin Kandala. 2011. "The Spatial Distribution of Health Establishments in Nigeria." *African Population Studies* 25 (2): 680–96. http://www.bioline.org.br/pdf?ep11047.

OCHA. n.d. "Quality Assurance Framework" Humanitarian Data Exchange. https://centre.humdata.org/wp-content/uploads/HDX_Quality_Assurance_Framework_Draft.pdf.

Okolocha, C., J. Chiwuzie, S. Braimoh, J. Unuigbe, and P. Olumeko. 1998. "Socio-Cultural Factors in Maternal Morbidity and Mortality: a Study of a Semi-Urban Community in Southern Nigeria." *Journal of Epidemiology & Community Health* 52 (5): 293–97.

Ronsmans, C., J. F. Etard, G. Walraven, L. Hoj, A. Dumont, L. Bernis, and B. Kodio. 2003. "Maternal Mortality and Access to Obstetric Services in West Africa." *Tropical Medicine and International Health* 8 (10): 940–48.

Ronsmans, Carine, and Wendy Graham. 2006. "Maternal Mortality: Who, When, Where, and Why." *The Lancet* 368 (9542): 1189–1200. doi:https://doi.org/10.1016/S0140-6736(06)69380-X.

Ruiz, Ivan Zahinos. 2010. "Effectiveness Assessment of Maternity Waiting Homes in Increasing Coverage of Institutional Deliveries Using Geographical Information Systems in Six Districts of Cabo Delgado Province (Mozambique)." Masters of Public Health. University of South Africa. http://uir.unisa.ac.za/bitstream/handle/10500/4843/dissertation_ruiz_iz.pdf?sequence=1

Say, Lale, Doris Chou, Alison Gemmill, Özge Tunçalp, Ann-Beth Moller, Jane Daniels, A. Metin Gülmezoglu, Marleen Temmerman, and Leontine Alkema. 2014. "Global Causes of Maternal Death: a WHO Systematic Analysis." *The Lancet Global Health* 2 (6): e323–e333. doi:10.1016/s2214-109x(14)70227-x.

Snow, Dan. 2013. "DR Congo: Cursed by Its Natural Wealth," October 9. https://www.bbc.com/news/magazine-24396390.

The Guardian. 2012. "Congo's Maternal Mortality Rate Exacerbated by Poverty and Ignorance." https://www.theguardian.com/global-development/2012/may/11/congo-maternal-mortality-rate-poverty.

United Nations. 2019. "World Population Prospects." *Population Division*. https://population.un.org/wpp/.

United Nations Population Fund. 2016. "Maternal Health." https://www.unfpa.org/maternal-health.

The World Bank. 2015. "Maternal Mortality Ratio (Modeled Estimate, per 100,000 Live Births)." https://data.worldbank.org/indicator/SH.STA.MMRT?year_high_desc=true.

Appendix - Model Builder Diagram of the Complete Methodology