



Faculty of Computing, Engineering and Media

**Adaptable Spatial Agent-Based Facility Location for
Healthcare Coverage**

PhD THESIS

**Olukemi Oyefunke Olowofoyeku (MNIS)
BSc. (Hons.), MSc. (Surveying & Geoinformatics)**

**Submitted in partial fulfilment of the requirements
for the degree of:**

**DOCTOR OF PHILOSOPHY IN
COMPUTATIONAL GEOSPATIAL SCIENCE**

September, 2019

Declaration

This is to certify that:

- (i) the thesis comprises only my original work towards the PhD except where indicated,
- (ii) due acknowledgement has been made in the text to all other material used, appendices and footnotes.

Olukemi Oyefunke Olowofoyeku

I dedicate this thesis to my dearest children - ADEOLA, ADEBAYO and ADEYEMI

Acknowledgements

I give praise to GOD ALMIGHTY who has made all things possible.

I would like to express my unreserved gratitude to my supervisor, Dr. Jethro Shell, who has provided excellent dedicated supervision and guidance throughout the period of this PhD. Your 'progress' emails were quite propelling. I am indeed very grateful to other supervisory team members Dr. Lipika Deka and Prof. Francisco Chiclana, including Dr. Eric Goodyer for their motivations, support, encouragement and dedication towards the success of this PhD. I appreciate my independent assessors, Dr. Benjamin Passow and Dr. Daniel Paluszczyszyn for their contribution. I also thank my advisor, Prof. James Olaleye of University of Lagos, Nigeria for his support. I count myself blessed to have greatly benefited from your wealth of knowledge and experience.

My sincere gratitude goes to members of the Faculty of Computing, Engineering and Media - De Montfort University (DMU), especially the Dean of the faculty, Professor David Mba for all assistance rendered to me in the course of this PhD. I also acknowledge all the anonymous on-line software developers and contributors who provided answers to different questions that made this work a success, including Santa Fe Institute for supporting the on-line course on 'Introduction to Agent-Based Modelling', through Complexity Explorer. I also appreciate my office mates, colleagues and friends in DMU for their encouragement, supports and motivations - Sadir Fadhil, Issa Qabajeh, Alejandro Peña, Mohammed Younis and Abdul Ogunji.

I am grateful to Surv. Abodunrin Akinpelu for his professional and financial support. I acknowledge the Surveyor General of Lagos State - Surv. Tomi Sangowawa, Surv. Solomon, Dr. Sunday Alademomi, Dr. Lola Olayinka, Surv. Nurudeen Ishola, Daniel and Gboyega for their assistance regarding data information. I also appreciate Wale Aladeokin and Semiat of Yaba College of Technology, who didn't get fed up with my burdens.

I would like to express my deepest gratitude to my wonderful Bolaji family: my late father Oluwole Oyeniyi and my mother, Aderemi Felicia for leading me through the right path. I am indeed very grateful for your labour and support for me in life. I sincerely appreciate my mother who took up great responsibility of caring for my children to ensure that this PhD is a success. I thank my younger brothers, Olakunle "Kudeti", Olatunji "Ksinger", Temitope "Areatee", and my kid sister Yetunde "Orobo", my sisters-in-law, Abiola, Kemi, Oyinkansola, my brother-in-law, Kayode, including their children who give me 'big mummy' waves and calls. Thank you for your prayers, financial and emotional support. I also appreciate my late eldest brother Bamidele; my senior sisters - Bola Ogunwole, Biodun Abimbola, Remi Olopade. I am also grateful to Dr. Yinka Okuwoga, Bimpe, Dipo, Tosin, Taye, Dayo & Bunmi Bello. Thank you all for the support, prayers, love and encouragement. You are great pillars; I love you.

My profound appreciation to Toyin & Kunle Adesina for being so loving and car-

ing, and ensuring that my stay in Leicester is pleasant. Despite your tight schedule, you found time to always check on me and go through the report. I also thank my first host family in the UK, Mr. & Mrs. Giwa. My sincere appreciation to the Adelekes, Soji Omolafe, my brothers-in-law - Tunde Ikuesan and Sunday Olowofoyeku for checking on my family during my absence for this PhD. Also my 'Big Sis' Anne Akpoyoware who has always been so caring and encouraging. Thank you for the usual 'you can do it'.

I sincerely appreciate Past. Akindele, Past. Ige, Past. Obed, Past. Femi Awokiya, Baba Lasehinde, Baba Adeleke and Evang. Niyi. My appreciation also goes to my wonderful family of God in The Redeemed Christian Church of God (RCCG), Love Assembly, Derby and RCCG Desire of Nations, Leicester. My special thanks to Mummy Ajibola, and Pastors Adeola & Bola Owagbayegun. Thank you all for your care and prayers.

I am grateful to everyone who has given me support in different ways, though their names cannot be mentioned due to space limitation.

Most importantly, none of this would have been possible without the great love, support, patience and endurance of my family: my loving husband, Adeoye whose understanding is unequivocal; and my precious children, Adeola, Adebayo and Adeyemi. I would like to express my heart-felt gratitude to you for giving me all the encouragement and opportunity I need to succeed. Thank you for your perseverance! I love you.

Thank you all for encouraging me to climb this ladder, and not look at 'how far, but how well'. Your supports and motivations gave me hope and energy to move on and overcome the challenges that I faced throughout the PhD programme. Rather than give-up, you all let me see reasons to move-on.

I sincerely appreciate the financial support of the Nigeria Government through the Tertiary Education Trust Fund (TETFund), especially the bailout that came at a crucial period of this programme: with special thanks to the Executive Secretary - Prof. Suleiman Elias Bogoro, the Ag Director (Academic Staff Training & Development) - Arc. Babatunde Olajide and Mr. Abdullahi Baba Imam for their love for Nigeria's development and profound compassion towards the scholars. I also thank Yaba College of Technology, Lagos for granting me study leave.

Most of all, I thank the Almighty GOD, the giver of life who has given ALL to me through Jesus Christ and the Holy Spirit:

Great is Thy faithfulness
O God my Father
There is no shadow of turning with Thee
Thou changest not
Thy compassions they fail not
As Thou hast been
Thou forever will be

Great is thy faithfulness!
Great is thy faithfulness!
Morning by morning new mercies I see:
All I have needed thy hand hath provided—
Great is thy faithfulness, Lord, unto me!

Abstract

Lack of access to healthcare is responsible for the world's poverty, mortality and morbidity. Public healthcare facilities (HCFs) are expected to be located such that they can be reached within reasonable distances of the patients' locations, while at the same time providing complete service coverage. However, complete service coverage is generally hampered by resource availability. Therefore, the Maximal Covering Location Problem (MCLP), seeks to locate HCFs such that as much population as possible is covered within a desired service distance. A consideration to the population not covered introduces a distance constraint that is greater than the desired service distance, beyond which no population should be. Existing approaches to the MCLP exogenously set the number of HCFs and the distance parameters, with further assumption of equal access to HCFs, infinite or equal capacity of HCFs and data availability. These models tackle the real-world system as static and do not address its intrinsic complexity that is characterised by unstable and diverse geographic, demographic and socio-economic factors that influence the spatial distribution of population and HCFs, resource management, the number of HCFs and proximity to HCFs. Static analysis incurs more expenditure in the analytical and decision-making process for every additional complexity and heterogeneity. This thesis is focused on addressing these limitations and simplifying the computationally intensive problems.

A novel adaptable and flexible simulation-based meta-heuristic approach is employed to determine suitable locations for public HCFs by integrating Geographic Information Systems (GIS) with Agent-Based Models (ABM). Intelligent, adaptable and autonomous spatial and non-spatial agents are utilized to interact with each other and the geographic environment, while taking independent decisions governed by spatial rules, such as •containment, •adjacency, •proximity and •connectivity. Three concepts are introduced: assess the coverage of existing HCFs using travel-time along the road network and determine the different average values of the service distance; endogenously determine the number and suitable locations of HCFs by integrating capacity and locational suitability constraints for maximizing coverage within the prevailing service distance; endogenously determine the distance constraint as the maximum distance between the population not covered within the desired service distance and its closest facility.

The models' validations on existing algorithms produce comparable and better results. With confirmed transferability, the thesis is applied to Lagos State, Nigeria in a disaggregated analysis that reflects spatial heterogeneity, to provide improved service coverage for healthcare. The assessment of the existing health service coverage and spatial distribution reveals disparate accessibility and insufficiency of the HCFs whose locations do not factor in the spatial distribution of the population. Through the application of the simulation-based approach, a cost-effective complete health service coverage is achieved with new HCFs. The spatial pattern and autocorrelation analysis reveal the influence of population distribution and geographic phenomenon on HCF

location. The relationship of selected HCFs with other spatial features indicates agents' compliant with spatial association.

This approach proves to be a better alternative in resource constrained systems. The adaptability and flexibility meet the global health coverage agenda, the desires of the decision maker and the population, in the support for public health service coverage. In addition, a general theory of the system for a better-informed decision and analytical knowledge is obtained.

Contents

| | |
|-----------------------|-----------|
| List of Tables | xi |
|-----------------------|-----------|

| | |
|------------------------|------------|
| List of Figures | xiv |
|------------------------|------------|

| | |
|---|-----------|
| 1 Introduction | 1 |
| 1.1 Overview | 2 |
| 1.2 Background and Motivation for the Study | 5 |
| 1.3 Aim and Objectives of the Research | 8 |
| 1.4 Research Questions | 9 |
| 1.5 The Study Area | 9 |
| 1.6 Geospatial Analysis, GIS and ABM | 12 |
| 1.7 Structure of the Thesis | 13 |
| 2 Theoretical Framework | 15 |
| 2.1 Introduction | 15 |
| 2.2 Healthcare Services and Coverage | 15 |
| 2.3 Access to Healthcare | 17 |
| 2.3.1 Spatial Dimension of Access | 18 |
| 2.3.2 Measuring Spatial Accessibility | 19 |
| 2.4 Facility Location Problems | 23 |
| 2.4.1 Median Problems | 23 |
| 2.4.2 Covering Problems | 24 |
| 2.5 Considerations for Locating Healthcare Facilities | 25 |
| 2.6 Covering Problem Solution Models | 28 |
| 2.6.1 Formulation of MCLP | 28 |
| 2.6.2 Formulation of PCLP | 29 |
| 2.6.3 Maximal Covering Problem with Mandatory Closeness Constraints (MCPMCC) | 31 |
| 2.6.4 Capacitated Facility Location | 32 |
| 2.6.5 Suitability Analysis | 33 |

| | | |
|----------|--|-----------|
| 2.6.6 | Facility Type With Establishment Cost Constraints | 34 |
| 2.6.7 | Unified Approach to Covering Model | 34 |
| 2.7 | Large Neighbourhood Search (LNS) Solution Algorithm | 35 |
| 2.8 | Agent-Based Model | 36 |
| 2.9 | Spatial-ABM Adaptability for Facility Location | 39 |
| 2.9.1 | GIS and ABM Coupling | 40 |
| 2.10 | Data Representation | 41 |
| 2.10.1 | Raster Representation | 42 |
| 2.10.2 | Vector Representation | 42 |
| 2.11 | Summary | 44 |
| 3 | Materials and Methods | 45 |
| 3.1 | Introduction | 45 |
| 3.1.1 | Research Design | 46 |
| 3.2 | Spatial Data | 47 |
| 3.2.1 | Coordinate Reference System | 48 |
| 3.2.2 | Spatial Data Acquisition | 51 |
| 3.2.3 | Database Creation and Data Processing | 53 |
| 3.3 | Model Preparation | 53 |
| 3.3.1 | Agents and Environments Representation | 53 |
| 3.3.2 | Agents and Spatial Features Required for a Catchment Definition | 55 |
| 3.3.3 | Agents and Spatial Features Required for Facility Location Model | 56 |
| 3.4 | Travel-Time Catchment Model | 57 |
| 3.4.1 | Catchment Area Estimate | 61 |
| 3.4.2 | No-Coverage Layer Creation | 64 |
| 3.5 | Location-Allocation Model | 65 |
| 3.5.1 | Solution Models | 67 |
| 3.5.2 | Formulation of Capacitated Location-Allocation Model | 68 |
| 3.5.3 | Formulation of Mandatory Coverage Model | 70 |
| 3.6 | Spatial and Agent Characteristics | 73 |
| 3.7 | Location-Allocation Procedure | 74 |
| 3.8 | Large Neighbourhood Search for Location-Allocation | 75 |
| 3.8.1 | Initialization | 77 |
| 3.8.2 | Construction Phase | 79 |
| 3.8.3 | Destruction Phase | 80 |
| 3.8.4 | Repair Phase | 82 |
| 3.8.5 | Sorting Phase | 83 |

| | | |
|----------|---|------------|
| 3.8.6 | Improvement and Stopping Phase - PCLP | 84 |
| 3.9 | Visualization and Analysis | 85 |
| 3.10 | Summary | 85 |
| 4 | Model Testing and Application to Existing Case | 86 |
| 4.1 | Description of the 55-Node Dataset | 86 |
| 4.2 | Model Verification and Validation | 87 |
| 4.3 | Travel-Time Model Test | 88 |
| 4.3.1 | Verification on Travel Time-Limit and Geographic Location . . . | 88 |
| 4.3.2 | Estimated Catchment Area Test with Circular Buffer | 89 |
| 4.3.3 | Validation With On-Line Travel-Time Application | 91 |
| 4.4 | Spatial-ABM Optimization Model Verification | 93 |
| 4.4.1 | Initial Number of Facilities Variation on 55-Node Dataset | 94 |
| 4.4.2 | Initial Number of Facilities Variation on Lagos State Dataset . . . | 95 |
| 4.4.3 | Statistical Robustness of Coverage | 97 |
| 4.5 | Location-Allocation Validation on 55-Node Dataset | 99 |
| 4.5.1 | Decision Support for Number of HCFs and Service Distance . . . | 100 |
| 4.5.2 | Comparison With Other Works | 104 |
| 4.5.3 | Point Density With Facility Distribution | 108 |
| 4.5.4 | Significance of the Model With 55-Node Dataset | 109 |
| 4.6 | Summary | 110 |
| 5 | Implementation and Analysis | 111 |
| 5.1 | Introduction | 111 |
| 5.2 | Healthcare in Nigeria | 112 |
| 5.3 | Spatial Data and Agents | 115 |
| 5.3.1 | Healthcare Facility Locations | 115 |
| 5.3.2 | Road Network Data | 115 |
| 5.4 | Access to Existing Healthcare Facilities | 116 |
| 5.4.1 | Catchment Definition With Network Travel-Time Analysis | 118 |
| 5.4.2 | Patient Agents' Travel Characteristics and Time-Limit Compliant | 118 |
| 5.4.3 | One-Sample Equivalent Tests on Travel-Time Threshold | 119 |
| 5.4.4 | Assessment of Road Dataset | 120 |
| 5.4.5 | Service Area of Existing Healthcare Facilities | 120 |
| 5.5 | Location-Allocation Implementation | 128 |
| 5.6 | Optimization Results and Analyses | 129 |
| 5.7 | Decision Support for Facility Locations | 131 |
| 5.7.1 | Apapa_Ajeromi/Island | 132 |

| | | |
|----------|--|------------|
| 5.7.2 | Ikeja | 135 |
| 5.8 | Visualization and Overlay Analysis | 138 |
| 5.9 | Non-Spatial Correlation Analysis | 141 |
| 5.10 | Spatial Pattern Analyses | 143 |
| 5.10.1 | Data Preparation | 144 |
| 5.10.2 | Exploring HCFs Distribution | 144 |
| 5.11 | Spatial Autocorrelation Analysis | 146 |
| 5.11.1 | Spatial Autocorrelation Analyses Results | 148 |
| 5.12 | Overall Pattern Distribution of Optimized HCF Locations With Nearest Neighbour Analysis | 155 |
| 5.13 | Results of Located Healthcare Facilities | 156 |
| 5.14 | Summary | 162 |
| 6 | Conclusions and Future Work | 163 |
| 6.1 | Introduction | 163 |
| 6.2 | Discussion | 163 |
| 6.3 | Conclusion | 166 |
| 6.3.1 | Contributions | 167 |
| 6.4 | Recommendation | 171 |
| 6.5 | Limitations and Future Works | 173 |
| 6.6 | Summary | 174 |
| A | Research Papers and Award | 175 |

List of Tables

| | | |
|-----|--|-----|
| 3.1 | Research design | 47 |
| 4.1 | Travel-Time Distance and Time Experiment Result | 89 |
| 4.2 | Comparison of ABM and GIS catchment area estimates using Euclidean distance equivalent for 30 minutes travel time | 91 |
| 4.3 | Equivalence Test with Paired Data | 93 |
| 4.4 | Range of number of healthcare facilities, p and service distance for 100% coverage | 100 |
| 4.5 | Results for Capacitated Mandatory Closeness-MCLP for 55-node | 102 |
| 4.6 | Percentage coverage within D and corresponding coverage within S . . | 103 |
| 4.7 | Five facilities solutions for MCLP with mandatory closeness constants . | 104 |
| 4.8 | Comparison of total coverage (%) for facilities restricted to the nodes, facilities placed in continuous plane and Spatial-ABM facilities located at discrete plane. | 106 |
| 5.1 | Public Healthcare Facilities in Lagos State by LGA | 116 |
| 5.2 | One Sample Equivalence Test against 30 minutes travel-time (minutes) limit | 121 |
| 5.3 | Paired ABM and Google travel information: travel mode —walking . . . | 122 |
| 5.4 | Summary of solutions for Apapa_Ajeromi/Island | 132 |
| 5.5 | Aapa_Ajeromi/Island selected solutions (required $p = 5$; service distance $D = 12$ m) | 134 |
| 5.6 | Ikeja summary of solutions | 135 |
| 5.7 | Ikeja selected solution (required $p = 4$; service distance $D = 12$ m) . . . | 136 |

List of Figures

| | | |
|------|--|-----|
| 1.1 | Map of Nigeria | 10 |
| 1.2 | Lagos State and Local Government Areas | 11 |
| 1.3 | Cellular Automata Neighbourhood | 13 |
| 2.1 | Network versus Straight-line journey | 21 |
| 2.2 | Raster and Vector data representation | 43 |
| 3.1 | Geoid and Ellipsoid Reference Surface | 49 |
| 3.2 | World Universal Transverse Mercator (UTM) zones | 50 |
| 3.3 | NetLogo environment | 54 |
| 3.4 | Travel distance covered | 61 |
| 3.5 | Travel-Time Flow Chart | 62 |
| 3.6 | Minimum Bounding Circle and Circular Buffer | 64 |
| 3.7 | No-coverage feature extraction from administrative and coverage features | 65 |
| 3.8 | LNS Phases | 78 |
| 3.9 | Patches and Agent location | 81 |
| 4.1 | 55-node map (ReVelle et al., 2007) | 87 |
| 4.2 | Travel time equivalence test with on-line road data | 92 |
| 4.3 | 55-node dataset: selected number of facilities from varying initial candidate facilities | 95 |
| 4.4 | Lagos State dataset: selected number of facilities from varying initial candidate facilities | 97 |
| 4.5 | Coverage verification analysis | 98 |
| 4.6 | 55-node facility coverage | 101 |
| 4.7 | 55-node Optimal solution for mandatory closeness constraints (Church and ReVelle, 1974) | 105 |
| 4.8 | Spatial-ABM solution for MCLP with mandatory closeness constraints for 55-node | 105 |
| 4.9 | Alternative 7-facility solutions for 55-node dataset | 108 |
| 4.10 | capacitated 7-facility solution: population density and facility distribution | 109 |

| | | |
|------|--|-----|
| 5.1 | Existing Lagos State Public Healthcare Facility and Water Bodies | 117 |
| 5.2 | Agent Travel Flow Chart | 118 |
| 5.3 | Histogram of travel-time of simulated agents | 119 |
| 5.4 | Visual comparison of road data and travel paths of agents with Google Maps | 123 |
| 5.5 | Sample road dataset update for Badagri LGA | 124 |
| 5.6 | Variation in catchment area of healthcare facilities | 125 |
| 5.7 | Lagos State Public Healthcare Facility Individual Catchment | 126 |
| 5.8 | Lagos State Public Healthcare Service Coverage | 127 |
| 5.9 | Sample model output view | 130 |
| 5.10 | Apapa_Ajeromi/Island/Ikeja percentage population uncovered and maximal distance to closest HCF | 133 |
| 5.11 | Set of p -facility in Apapa_Ajeromi/Island | 134 |
| 5.12 | p -facility-type in Apapa_Ajeromi/Island | 135 |
| 5.13 | Ikeja percentage population uncovered and maximal distance to closest HCF | 136 |
| 5.14 | p -facility-type in Ikeja | 137 |
| 5.15 | Set of solutions for Ikeja | 137 |
| 5.16 | ABM and GIS visual analysis of healthcare facility locations | 139 |
| 5.17 | Four overlaid feasible solutions to healthcare coverage map | 140 |
| 5.18 | Correlation of uncovered population with number of facilities and maximum service distance | 142 |
| 5.19 | Box graphics of healthcare facilities distribution | 145 |
| 5.20 | Box plot maps for exploratory analysis of aggregated HCFs in grid polygons | 145 |
| 5.21 | Natural breaks of aggregated HCFs in grid polygons | 146 |
| 5.22 | Contiguity illustration | 147 |
| 5.23 | Connectivity histogram | 147 |
| 5.24 | LISA for new HCFs | 149 |
| 5.25 | LISA for new HCFs | 150 |
| 5.26 | LISA for new and existing HCFs | 151 |
| 5.27 | Population map | 152 |
| 5.28 | Correlation of existing and new healthcare facilities vs water bodies and population | 154 |
| 5.29 | Correlation matrix of total healthcare facilities vs population | 155 |
| 5.30 | Average nearest neighbour analysis of new HCFs located within the . . | 156 |
| 5.31 | Region healthcare facilities by type | 157 |
| 5.32 | Total located healthcare facilities by type | 158 |
| 5.33 | Type of healthcare facilities by percentage | 158 |

| | | |
|------|---|-----|
| 5.34 | Lagos State: optimized healthcare facility locations by facility-type . . . | 160 |
| 5.35 | Lagos State: optimized and existing healthcare facility locations | 161 |

Chapter 1

Introduction

Planning and decision-making on the appropriate locations for setting up facilities are vital to any service provider, especially in the provision of healthcare service to the public. This PhD thesis proposes procedure and solution for covering-based facility location problems, with a focus on locating centre healthcare services that will enhance better access to the public, and serve as a cost-saving decision support for universal health coverage. Centre healthcare facilities are expected to be accessible and offer maximum coverage to the entire population (Afshari and Peng, 2014; Ahmadi-Javid et al., 2017). Decisions on locating such services need to consider reasonable proximity of the population to the services in terms of road network traverse, and capacity of the facility in relation to the size of the population, including the suitability of selected facility sites. Such analysis is a complex geographic and computational problem that is simplified in this thesis through a synergy of Agent-Based Models (ABM) and Geographic Information Systems (GIS) analysis, to produce a decision support tool that non-technical healthcare providers and policy makers can relate with. Agents' ability and characteristics are utilised to interact with the spatial environment to simplify the complex problems.

The main objective function of the facility location analysis is to ensure coverage of healthcare service to the whole population. However, the provision of the required number of facilities that will provide total coverage may not be feasible in certain settings due to insufficient resources. Therefore, the Maximal Covering Location Problem (MCLP) seeks to cover as much population as possible with fixed number of facilities within a desired service distance (Church and ReVelle, 1974). The traditional MCLP does not guarantee a total coverage of the population. However, all the populations left uncovered can be assigned to the nearest facility to them. This proffers an additional objective that minimizes the maximum distance that patients outside the coverage of the average service distance will need to travel to receive services. In view of the fact that scarce resources may also be a challenge to the establishment of fixed number of facilities, there may be need to consider different alternative number and sites of facilities, including varying the costs of their establishment (with some being low-cost while others are costly). The problem is therefore complex and multi-criteria.

The Adaptable Spatial Agent-based Facility Location (ASABFL) approach for this research determines the average service distance that patients travel to healthcare facilities (HCFs), which serves as input to endogenously determine the number and optimal or near optimal locations of new HCFs, that will cover as much population as possible. The model assigns the population not covered to HCFs nearest to them and endogenously determines the maximum service distance between the farthest demand and its closest HCF. The resources needed for establishment of the HCFs decide if a HCF is low-cost type or standard type. The location-allocation model considers capacity of HCFs in terms of facility density as a measure of number of HCFs per population size, and utilizes the capacity to determine the type of HCF to be established in relations to resources. This problem is therefore proposed as Capacitated Mandatory Closeness Maximal Covering Location Problem (CMC-MCLP).

The solution for the maximal coverage problem that will locate HCFs and allocate demands to the facilities assumes that the required number of HCFs are not fixed, and the expected travelling distance and the population size are known for the HCF sites to be determined. The second solution to ensure that the maximum travel distance of the uncovered population in the first solution is minimized assumes that the service distance is not known; however, HCF locations are fixed from the previous MCLP solution. The facility location problems are solved simultaneously to give solutions that can aid decision at improving health outcomes. A meta-heuristic optimization approach is considered which may not guarantee an exact solution but can give a set of feasible solutions.

1.1 Overview

Access to healthcare is a basic need and right of everyone (WHO, 2006). However, for a considerable number of people in the world, especially in most developing countries, accessing healthcare is generally poor, despite the importance of good quality healthcare interventions (Baeten et al., 2018). In these countries, healthcare services are sparsely distributed and insufficient to take care of the populations' health needs. Research and health policy therefore needs to be focused on improving healthcare services to enhance the people's well-being and their quality of life, which will consequently reduce their poverty level (Peters et al., 2008). For effective healthcare planning and service intervention, decision makers require evidence-based information on the demand and supply of treatment with the aim of moving healthcare service close to the people. One of the ways to achieve these is through the analysis of existing HCFs, including the measurement of how accessible they are to the population served from different locations, and subsequently address deficiency in healthcare spatial coverage.

Access to healthcare is a complex, multi-faceted concept that involves five prime dimensions which includes affordability (costs of healthcare utilization), acceptability (health service compliance and satisfaction), availability (adequacy of health service provision), spatial or geographical accessibility (travel impedance between patients

and providers), and accommodation (appropriateness and suitability of health services) (Levesque et al., 2013). Of these dimensions, spatial accessibility which is a major contributor to overall population health (Jamtsho et al., 2015), is an essential measure of access to healthcare that can translate to efficient healthcare facility location. This aspect of healthcare service intervention has been previously ignored by health professionals and policy makers in many countries of the world. One major challenge to such analysis is the lack of technical know-how and resources required by the sophisticated and expensive tools that are available. Within the current research, a novel method will be developed for assessing spatial accessibility of the population to healthcare facilities and optimizing locations for new ones to improve health service coverage, using flexible and inexpensive tools that will support flexibility in decision making.

For this study, Lagos State, Nigeria is chosen as a test case. Nigeria is still regarded as being considerably poor despite her recent robust economic growth, because the country has still been unable to achieve remarkable success in the area of healthcare services to the citizens. Morbidity and mortality rates in Nigeria are considerably high. Presently, the country lacks policy on investigating proximity to healthcare services that can inform Universal Health Coverage strategies and locating new HCFs.

Although varying techniques exist for optimizing facility locations, the dearth in technical and economic capability in most developing countries make these methods unrealistic to adopt. A methodology that can be applied to the varying objectives, constraints and features in a generalized form is needed as part of the strategy for universal health coverage and sustainable development. ABM integrated with GIS techniques shall be adopted to achieve the aims of this research. ABM involves using computational methods to investigate and analyse processes and problems that are regarded as intelligent, heterogeneous systems of interacting intelligent and autonomous agents, which can be heterogeneous, fixed or non-fixed (Ali, 2012; Crooks and Heppenstall, 2012). ABM's flexibility enables potential variables and parameters to be specified and explored with different complex spatial scenarios which may be difficult mathematically. Integrating ABM and GIS provides an adaptable model that can create health outcomes of low mortality and morbidity rate, and increased life expectancy for the present and future populations. Therefore, this thesis will develop an interactive adaptable spatial agent-based facility location simulation model that considers complex heterogeneous variables of the population and the environment, to serve as a decision support tool for healthcare provisions for the entire population.

Consequently, the research work includes, spatial and non-spatial data collection, model development, and geospatial analysis. The data requirements include population density data, existing HCFs' geographical locations, a land use/land cover map, large scale satellite imagery and street maps of the study area. All spatial data will form map layers in the GIS environment and be imported to the ABM environment, and be exported back to the GIS environment for further analysis and visualisation. Facility catchment or service area of HCF will be based on travel time and distance of

the population to healthcare services, in compliance with some of the World Health Organization (WHO) (1998) recommendations on catchment and facility locations. WHO (1998) recommends that the catchment for healthcare facilities can be determined based on factors which may be:

- i politico-administrative boundary,
- ii geographical boundaries that are natural physical barriers to population movement, and
- iii time boundaries that assumes the population move towards the most accessible facility in the shortest time

The appropriate and suitable locations for healthcare facilities should consider certain guidelines, including the following:

- i It should be within 15-30 minutes travelling time or about 25km radius in an area with good roads and adequate means of transport.
- ii It should be grouped with other institutional facilities, such as religious (church, mosques), educational (school), tribal (cultural) and commercial (market) centres.
- iii It should be free from dangers of flooding and so should not be sited at the lowest point of the area.
- iv It should be in an area free of pollution of any kind, including air, noise, water and land pollution.
- v It must be serviced by public utilities like water (could be well or bore hole), sewage and storm-water disposal, electricity (could be generators), gas and telephone.

This thesis presents a research that shall consider:

- i 30 minutes travelling time limit for catchment analysis
- ii grouping facility sites with other institutional facilities, such as religious (church), educational (school), and commercial (market) centres
- iii not placing facility on water body or other committed sites, such as airport and stadium

These criteria will be considered heuristically to arrive at a solution. Other recommended factors are beyond the scope of this thesis.

The primary objectives of the facility location spatial simulation model shall be to maximize coverage of services, increase accessibility of population to healthcare

facility locations and reduce the cost of establishing the healthcare facilities. Model validation will compare the developed Spatial-ABM travel time analysis with Google Map Distance Matrix Application Programming Interface. The spatial optimization model will also be tested on the 55-node dataset of Swain (1971) that is widely used in the literature (Church and ReVelle, 1974; Church and Roberts, 1983; Caccetta et al., 2005; ReVelle et al., 2007; Lei et al., 2016). The 55-node network dataset is the locational dataset representing 55 demand locations available in ReVelle et al. (2007).

The simulation model will be a flexible and adaptable decision support tool that will accommodate varying geographical, time, environmental, and social factors to produce results which will aid healthcare policies and decisions on location of HCFs. It will serve as a complement to the existing universal coverage strategies that will facilitate 'health for all' and poverty reduction. It will also serve as an improvement on the traditional GIS method, and be comparable with other costly and highly sophisticated approaches. The research will contribute to knowledge in the application of simulation models in spatial optimization problems.

1.2 Background and Motivation for the Study

This section describes the background of this PhD thesis that was primarily motivated by two factors:

1. **Need for suitable technique:** There is a dearth of research on suitable technique for optimizing healthcare facility locations within a resource and data deficient setting. More importantly, to the best of the author's knowledge, the integration of ABM with GIS has not been employed in multi-criteria covering based facility location problem as applicable to public healthcare facilities and WHO (1998) guidelines. Within the geographical systems, numerous works have used GIS's geospatial analysis approach, but very few have explored the use of a geosimulation method in which simulation modelling is applied to GIS analysis for public healthcare locational research. Although GIS functionalities such as buffering (generating catchment of hospitals), overlay analysis (examining different map data), network analysis (using characteristics of a network) and visualization techniques give understanding and decision-making capabilities for healthcare services; this method only provides static analysis. Geosimulation methods offer a more robust, flexibility and adaptable analysis. Facility location analysis is a geographical phenomenon that is influenced by complex heterogeneous processes, such as environmental, physical (MacKinney et al., 2014), and social variables which should not be overlooked in providing healthcare services to the population. There are five reasons why the traditional GIS approach is not appropriate for this research:

- First, the traditional method assumes stable spatial and temporal conditions (De Smith et al., 2018), which may be appropriate for developed coun-

tries but unsuitable for developing countries; where housing, environmental, and planning policies are unstable.

- Second, spatial accessibility to healthcare services should consider the effect of the transportation structure on travel time - especially when good road surfaces and structured transportation system are deficient, and there are inadequate guidelines on average speed for each road class (Nelson, 2000; Delamater et al., 2012; Rodrigue et al., 2013).
- Third, facility location measure should not assume a uniform socio-economic condition for the population and the country.
- Fourth, a facility location model needs to put into consideration the possibility of locating different types of facilities such as temporary or low-cost facilities, especially when the population to serve is below the optimal. This is an important consideration for low-and-middle-income countries where funds for establishing permanent facilities may not be readily available (Rahman and Smith, 2000).
- Fifth, the traditional method assumes that there is availability of data and technical know-how for analysis and implementation (Luo and Qi, 2009), whereas the use of GIS requires training for understanding and analyzing results.
- Sixth, the traditional method's static modelling assume that the potential facility locations are pre-known and the required facilities are uncapacitated and limited in number, to cover fixed and finite number of demand points within a predefined maximum service distance. Designing a flexible model that can endogenously determine the number of facilities and maximum distance that the farthest population to any facility will have to travel gives the decision maker more flexibility to choose amongst varying alternatives (Berman et al., 2010).

Distance and time are major factors in accessing healthcare. These two factors are geographical phenomena that depict physical access and can be analysed based on travel distance or travel time that can be appropriately measured with basic information such as travel speed on road dataset. In these regards, geospatial analysis with GIS has aided the computation of service area of healthcare services and decisions to determine locations for new HCFs. These have been possible due to vast availability of spatial data; and sophisticated GIS and modelling software, that are predominantly affordable by developed countries. Unfortunately, the under-developed countries are in dearth of such robust resources and technical capability to adopt and implement such techniques. Nonetheless, few researchers have tried to be as simple as they can by employing crude GIS techniques that do not incorporate the required data to carry out analysis, and in turn produce unreliable results and a mis-informed decision. With access to healthcare as a vital need of all populations, the importance of the required spatial

data such as, actual travel pattern, population locations, demographic data, land use and land cover, and location suitability data cannot be overlooked. A novel methodology that will consider these geographical factors is highly needed, with the objectives to maximize access, minimize cost and maximize healthcare coverage.

Although for healthcare facility location problems, GIS has been integrated in different mathematical modelling technique such as linear programming, mixed integer programming and dynamic programming. For example, Dudko et al. (2018) applied such technique for selection of sites for primary healthcare facilities in Australia. Ye and Kim (2016) combined mathematical programming with GIS to develop a covering-based location model and applied it to locating healthcare facilities in Hillsborough County in Florida. Such models have been found to be computationally intensive, both in the amounts of time to find optimal solution (Lei et al., 2016), and the complexity of the structure.

A reasonable alternative is to use a budget-friendly, simple, flexible and adaptable geosimulation method which integrates GIS with ABM in the geospatial analysis. GIS is required to inform the model spatially as agents' spatial location is a significant element in this context. GIS provides the space or environment through which agents will interact. ABM is a simulation technique that supports decision to better understand the operation of a particular system for enhanced informed decision to be taken (Siebers and Aickelin, 2008). For example, decision can be made for a community-level facility that can be upgraded later as the population increases with time. This method allows outcomes over time to be measured due to change in spatial and agents' attributes as applicable to this study. Very few works have explored the use of this unique and more robust, flexibility and adaptability. For example, Chen et al. (2010) applied spatial-ABM to land allocation problem in Panyu, China. Liu et al. (2016) used this synergy to develop a planning support model for the development of creative industries in Jiading, Shanghai.

2. **Need for a decision support tool for facility location:** Coverage and access issues have not been given much attention in many countries of the world, especially at facility level to improve coverage of intervention to all citizens for the expected Universal Health Coverage (UHC) of the Sustainable Development Goals (SDGs) by 2030. The importance of bringing care close to the patient and the role of integrated community case management and community health workers has been identified as one of the ways to reduce health burden (WHO, 2017). There is a dearth of research on suitable methodology to inform and aid decision on HCF location-allocation to match the demands for treatments. For example in Lagos State, there is no policy in place for optimizing the location of HCFs, instead HCFs are randomly placed based on the guideline that each Local Government Area (LGA) should have at least a HCF. Rather than establish new HCFs that are more accessible to the citizens, land spaces within existing facilities are expected to be

utilized through the building of multi-storey structures (Lagos State Ministry of Health, 2015a). This measure is not appropriate for a developing country with no constant supply of electricity for lifts or escalators to convey the sick to higher floors. Understanding the spatio-temporal relationship that exists between the supply of treatment and the demand by the population provides a useful tool for decision makers in planning for the location of additional treatment facilities.

The World Health Organization has recently (WHO, 2018b) called for the return of The Alma-Ata Declaration of 1978 which affirms that Primary Health Care (PHC) is:

“.....essential health care based on practical, scientifically sound, and socially acceptable methods and technology made universally accessible to individuals and families in the community through their full participation and at a cost that the community and country can afford to maintain at every stage of their development in the spirit of self-reliance and self-determination.....” (WHO, 1978)

While countries are allowed to adopt strategies that best suit their circumstances, a robust approach that will produce enhanced health outcomes, without compromising essential details and accuracy should be employed.

Therefore, the priority of this research is to develop a decision support tool with GIS and ABM software to present it as an innovative system for researchers and policy makers in the front-line of healthcare provisions. The potential to transfer the knowledge acquired in this research for further strengthened health system in the global community is apparent, in order to improve the populations' health and reduce death rate due to deficient health service coverage. This methodology can further be applied to locating other centre or public facilities as well.

1.3 Aim and Objectives of the Research

The aim of this research is to develop a novel method of modelling public healthcare coverage and optimizing new public healthcare facility locations. This is to facilitate sustainable healthcare policies and decisions as part of health service coverage strategies, through the use of a methodology that can be applied to varying spatial, economic and social characteristics.

In order to achieve the aims of this research, the following objectives are considered.

1. Develop methodology that can allow the integration of GIS and ABM in geospatial analysis for public healthcare decision-making and management at facility level.

2. Provide understanding on the relationship between travel time or distance and healthcare coverage for future development decisions for healthcare providers and stakeholders.
3. To generate optimized healthcare facility locations for population that is out of health service coverage.
4. Evaluate healthcare facilities that are under-served, over-served or adequate based on healthcare facility to population ratio within the catchment area.
5. To integrate locational objectives into geosimulation models with focus on healthcare facility density in relation to population.
6. To understand the impact of population on costs of healthcare facility establishment.
7. To evaluate suitability of optimized healthcare facility locations based on near amenities, and land use/land cover criteria.

1.4 Research Questions

The questions that this research seeks to ask are therefore:

1. Can Spatial Agent-Based modelling analysis be used to support decision-making in healthcare coverage?
2. Can Geographic Information Systems (GIS) and Agent-Based Models (ABM) be integrated to analyze healthcare coverage and optimize facility locations for health interventions?

This research seeks to answer the above questions by modelling public healthcare coverage and optimizing healthcare facility locations to be applied in Lagos State, Nigeria for improved healthcare, using Geographic Information Systems (GIS) and Agent-Based Models (ABM) to facilitate decision-making in healthcare delivery.

1.5 The Study Area

The study area for the research is Lagos State, situated in southwestern part of Nigeria. Nigeria is the fourteenth largest country in Africa and is located in West African sub-region with a total land area of 923,768 square kilometres. The country lies between latitudes 4°16' and 13°53' north and longitudes 2°40' and 14°41' east. Nigeria shares boundaries with Niger in the north, Chad in the northeast, Cameroon in the east, Benin in the west, and approximately 850 kilometres of the Atlantic Ocean in the south.

Nigeria has 36 states plus the Federal Capital Territory. Lagos state covers an area of approximately 356,861 hectares (3,569 square kilometres) of which 75,755 hectares

are wetlands (lagoons and creeks), and lies approximately on latitudes $6^{\circ}22'$ and $6^{\circ}41'$ north and longitudes $2^{\circ}42'$ and $4^{\circ}21'$ east. The state is bounded on the North and East by Ogun State. It shares boundaries with the Republic of Benin in the West, and the Atlantic Ocean in the south. Figure 1.1 shows the map of Nigeria, and the map of Lagos State, the study area, is presented in Figure 1.2. Although Lagos state is the smallest state in Nigeria, it is the most populated with an estimated population of 17.5 million in 2006, and an estimated growth rate of 3.2% (600,000) per annum. The population was projected to be approximately 24 million in 2016 (Lagos State Bureau of Statistics, 2017). The United Nations estimates that Lagos will be third largest mega city in the world after Tokyo in Japan and Bombay in India as a result of its population growth rate. There is heavy migration to Lagos from other parts of Nigeria and surrounding countries. Migration takes the biggest part of the population growth and the main motivation for migration is economic as Lagos is known to be a unique national centre for trade and commerce in Nigeria with lots of job opportunities.



Figure 1.1: Map of Nigeria

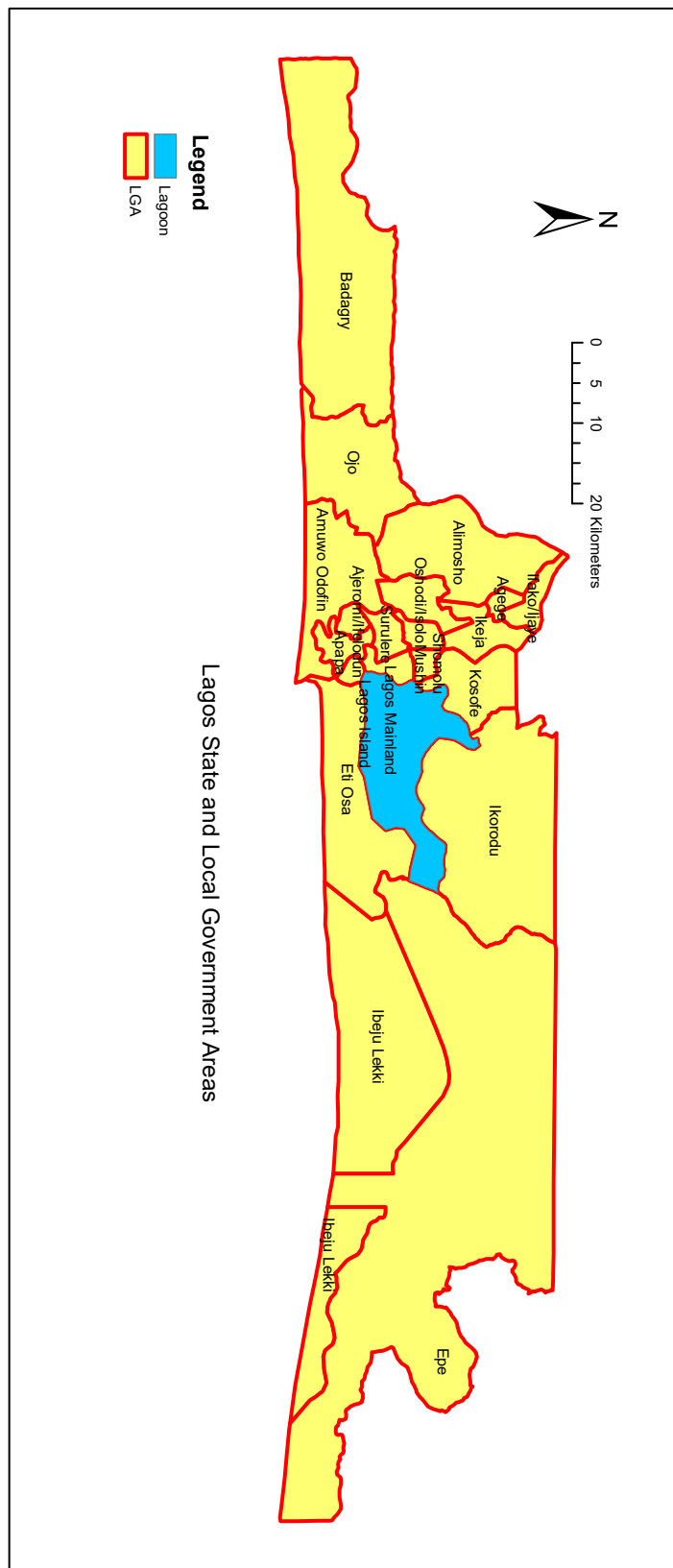


Figure 1.2: Lagos State and Local Government Areas

The inadequacy of public facilities and access to treatment, amongst other reasons have made private hospitals to be major healthcare providers to over 60% of the population within Lagos State (Lagos State Ministry of Health, 2015b), and has consequently increased the poverty level of the people.

1.6 Geospatial Analysis, GIS and ABM

This section provides a brief explanation on geospatial analysis, GIS, and ABM. Further details are provided in Chapter 2 and Chapter 3.

Geospatial or spatial analysis uses techniques and tools such as GIS, remote sensing, Global Positioning System (GPS) and computational methods to examine pattern or trend, or processes of georeferenced or location-based data and phenomena. While it is not easy to give a definition of GIS, Maguire (1991) has suggested that the way to define the system is to summarize the diverse ideas from different perspectives which are mainly, the map, database and spatial analysis. The author described GIS as an integrated collection of hardware, software, data, and liveware which operates in an institutional context. Using the definition of Cowen et al. (1990) in the context of this thesis, GIS is "a system of hardware, software and procedures designed to support the capture, management, manipulation, analysis, modelling and display of spatially referenced data for solving complex planning, management and research problems".

GIS can manage spatial data and can be used for harmonizing and continuous investigation of data (Chanda et al., 2012). GIS spatial analysis and functions combine different data layers such as satellite imageries, digital maps, orthophotos, and statistical data that can be spatially referenced for different analysis to be performed to support healthcare decision making. Such analysis include - assisting public health practitioners to identify gaps or inequities in healthcare delivery; recognising and developing solutions to address shortfalls in healthcare services and deliveries; giving enhanced geographic visualization techniques, to enhance faster, better, and more robust understanding and decision-making proficiencies for healthcare services and delivery. McLafferty (2003) affirmed that adoption of GIS techniques for healthcare researches and policy-making is a function of access to related data on healthcare providers, customers and spatial data.

The complex spatial and non-spatial variables and parameters needed for this research would require the acquisition of vast amount of data which will be time consuming, tedious and costly. The processing of such data also requires high computing powers, special skills and software. The reasonable approach is to create a model of the world and simulate the real world phenomena on such model to address the complex and heterogeneous issues. The simulation model enhances the ability to control, forecast or understand the behaviour of the system being modelled. Such models help to explore what is known and what is not known about the system by changing the variables or parameters. Borshchev and Filippov (2004) describe a simulation model

as a set of rules that define how the system being modelled will change in the future from its present state. Computer simulations are now being used as a complement or substitution for the traditional mathematical modelling.

Two distinct classes of simulation models are Cellular Automata (CA) and ABMs. CA uses cells or grids to represent objects or agents in space. Agent in a cell can change state and only interacts with the nearest neighbour. Cells of neighbourhood can increase in order to increase agent interaction (Figure 1.3). The immobility of cells in CA gave rise to Agent-Based Models or Multi-Agent Systems which allow mobile objects such as, vehicles and pedestrians to be modelled. This capability makes ABMs more suitable for this research. Agent-based modelling is a computational simulation method that enables the creation, analysis, and exploration of models that are composed of agents that interact within an environment (Gilbert, 2008). The simulation where agents are linked to the geographic space, and the environment where they interact represent real-world spatial data is referred to as **Spatial Agent-Based Modelling** (Filatova et al., 2013) or **Agent-Based Geosimulation** (Moulin et al., 2003). Geosimulation is the process of studying the behaviour and interaction of objects or agents in space using geographic or spatial models in order to understand and analyse complex and heterogeneous geographic systems.

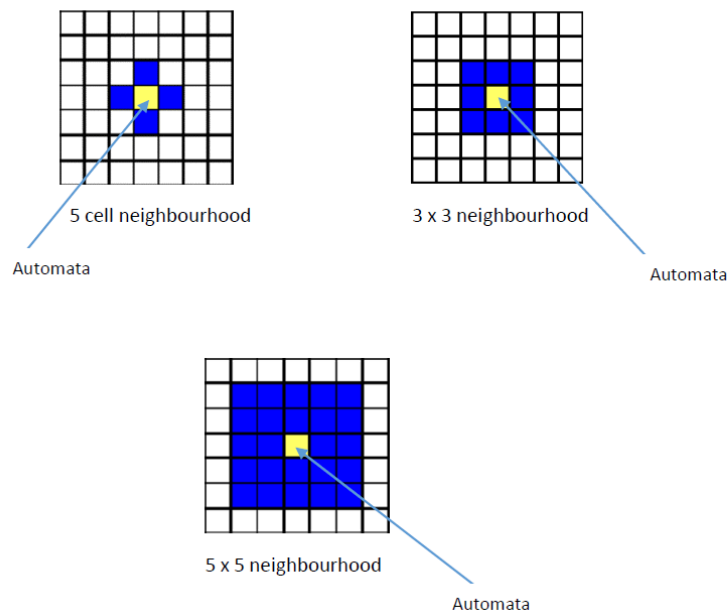


Figure 1.3: Cellular Automata Neighbourhood

1.7 Structure of the Thesis

The thesis is organised into six chapters as follows:

Chapter 1 covers the introduction, which includes the research overview, the motivation for the research, description of the study area, and the primary aim and objectives of the research. It also includes a brief introduction of ABM and GIS and concludes

with the structure of the thesis.

Chapter 2 provides a general concept of healthcare services and access to healthcare, with relevant background on spatial accessibility and its measures. It also includes a background and review on facility location problems, solution and models for location problems with emphasis on covering-based models and the applicability of ABM and GIS in facility location problems.

Chapter 3 introduces the GIS data and their representation, including basic coordinate reference systems. The proposed ABM and GIS travel-time and facility location models which integrate spatial data and ABM simulations technique into health service coverage modelling and location-allocation process are described. It includes detailed methodological process of the proposed research approach, which are data collection, database creation, agents and their characteristics, ABM simulations, and geospatial visualisation and analysis.

Chapter 4 describes the experimentation of the models and validations on the Google Maps web mapping application and the 55-node dataset. The analysis and the impact of the test results on healthcare access and coverage are also described.

Chapter 5 describes the implementation of the models on a case study of public HCFs coverage in Lagos State, Nigeria. Detailed explanation about the data collection, database creation, travel-time model, location-allocation model, and the results of the models are described. The spatial pattern analysis of the existing HCFs and the newly located HCFs, including their relationships with spatial features are compared. The chapter concludes with the presentation of results.

Lastly, Chapter 6 concludes the thesis with explanation on the impacts, and limitations of the research.

Chapter 2

Theoretical Framework

2.1 Introduction

This chapter describes the theoretical framework upon which the presented research is based. A detailed review of the literature is provided to understand different classification of healthcare services and the methods for carrying out geospatial analysis with a focus on geographic accessibility and spatial optimization techniques for selecting sites for new healthcare facilities (HCFs). Concept of accessibility and the measurement of geographic accessibility; and the different classification of facility location problems and their solutions are described. There are also further discussions on Agent-Based Model (ABM) and Geographic Information Systems (GIS) techniques and their integration as employed in this thesis, in proffering solution to the covering-based facility location problem required in locating public health centre facilities. The chapter also describes how spatial data is represented.

2.2 Healthcare Services and Coverage

Aside from the consideration of the health needs of the people, decision on the type and number of facilities to be located in a healthcare facility location problem will also depend on what type of service such facilities are providing. Following the classifications of Afshari and Peng (2014) and Ahmadi-Javid et al. (2017), healthcare services can be categorized into three broad classes: preventive healthcare services, emergency healthcare services and health centre services.

Preventive healthcare services: These are healthcare services that are provided for preventive programs such as immunization of children, anti-natal services for pregnant women, blood tests for disease detection and control, other services of control and early detection of diseases. Patients may not necessarily visit the closest HCF for these services as choice of visitation may be based on quality or cost of service among other factors.

Emergency healthcare services: Emergency services include ambulance services and disaster operations management, such as earthquake, flood mishaps and disease outbreak, emergency service stations vehicles, relief distribution and ambulance stations. The location of such services is dictated by the place or scene and the time of occurrence, which cannot be predicted. For such services, maximizing coverage and response time from service to demand location are key factors to determine health outcomes.

Health centre services: These services provide varying cares which are not regarded as emergency or preventive; however, some preventive services may be rendered. Visitation to these services are generally based on ease of access. Examples of such services include, primary healthcare facilities, specified service facilities, home healthcare centres, rehabilitation centres and preventive centres. Facilities that provide health centre services are public facilities that must provide service to the entire community and should be easily accessible to all. They are usually funded and managed by the government. Such health centres are the initial point of visits by patients for advice or treatment. This category of healthcare facilities are also referred to as central facilities (Rahman and Smith, 2000) and are the focus of this thesis. Since the services are required to cover the entire population, maximum coverage and accessibility are important factors for health service providers and decision makers.

This research is primarily focused on healthcare coverage to the entire community by public health centre or central facilities, to which people have to travel to receive services. Unfortunately, the decisions to find the appropriate service locations to enhance maximal coverage and ease of access are hindered by a number of factors. These challenges include:

- how to minimize the travel distance or travel time between demand (population) and supply (service)
- how to determine the capacity of the HCF
- how to measure the time or distance between demand (population) and supply (service)
- how not to over-allocate or under-allocate demand (population) to supply (service)
- how to ensure full coverage of services
- how to know the number of HCF that the population require
- how to know where to place the facilities
- how to know how suitable is the proposed location for placing new facilities
- how to minimize the cost of establishing the facilities

Various modelling techniques have been used to address few of these problems as a multi-objective (Zhang et al., 2016) or multi-criteria optimization (Basu et al., 2018) strategy. Although there is a considerable large number of literature on location models for central facilities, there is no approach that combines these objectives to provide a decision support for improving the public healthcare systems with integration of GIS and ABM, as provided by this thesis.

2.3 Access to Healthcare

Access to healthcare has long been considered a vital notion in healthcare management (Penchansky and Thomas, 1981), even though access is still generally regarded as a complex concept with no universally accepted definition and dimensions of measurement (Gulliford et al., 2002; Levesque et al., 2013; Penchansky and Thomas, 1981). However, there is a general consensus on the characteristics of access. Access is an act and a right of approaching; and it possesses attraction capacity and feature of being easy to reach. These various characteristics of access have resulted in diverse opinions by many authors on how access can be conceptualized and measured to fit into health service delivery.

In a healthcare context, Aday and Andersen (1974) in their work suggested that access to healthcare needs to be considered on whether individuals who need care get into the healthcare system - viewing access in this case as involving input and output. In a different context, Gulliford et al. (2002) also identified a need for healthcare services in access, but argued that access to healthcare involves opportunity to receive healthcare when it is required. Consequently, there is an interaction between supply and demand for healthcare. This is also supported by Levesque et al. (2013) who see access to healthcare as opportunity to have healthcare needs fulfilled - considering supply-side and demand-side features and processes that describe how access is realized.

While access may be viewed as an attribute of health services, access cannot be regarded as service utilization (Oliver and Mossialos, 2005; Penchansky and Thomas, 1981). Access may not actually translate into service utilization if the demands in need do not make use of the opportunity, possibly due to contact time interval (Aday and Andersen, 1974). For example, if the population that is served by a HCF is large compared to its carrying capacity, it may increase waiting time. In other words, health policy is considered the starting point of the concept of access, which runs through the characteristics of the healthcare system in terms of resources and organization, and of the susceptible population (Mosadeghrad, 2012; Arah et al., 2003; Aday and Andersen, 1974). There is therefore a complex interaction between healthcare services and population in terms of demand and supply.

2.3.1 Spatial Dimension of Access

Some authors (Gulliford et al., 2002; Peters et al., 2008) have suggested four dimensions of access with demand and supply elements to be: geographic accessibility, availability, financial accessibility or affordability and acceptability. In their submission, geographic accessibility relates to the physical distance or travel time from service delivery location to the user; availability is getting the right type of care available to those in need of the care, such as the operation and waiting times that meet the needs of those using the care, which also includes having the suitable type of service providers and resources; financial accessibility is the relationship between the cost of services and the user's willingness and capability to pay, including user's protection from the economic effects of health costs. Lastly, they defined acceptability as the linkage between health service providers' responsiveness to the socio-cultural expectations of the users and the communities.

Levesque et al. (2013), in their work proposed five dimensions of access as approachability, acceptability, availability and accommodation, affordability and appropriateness. Approachability is the ability to perceive a need and have knowledge of health services which is made possible through transparency and outreach activities of service providers. However, approachability may not result to utilization if services are not conveniently located for the population to reach as other physical, environmental, financial and social barriers may hinder approachability. The authors described availability and accommodation as the ability to reach health services physically and timely, including physical existence of health resources with sufficient capacity to produce services. Affordability is also related with the economic capacity for people to spend resources and time to use appropriate services, while acceptability relates to socio-cultural factors that determine the possibility of people accepting the services being provided. Appropriateness relates to adequacy and quality of health services, and its integrated and continuous nature.

It becomes apparent, from the above literature, that there is a spatial dimension to access which is a determinant of utilization, affordability, availability and outcomes. Spatial accessibility is an important consideration in health centre facility location (Ingram, 1971; Jacobs et al., 2012; Peters et al., 2008) as it specifically has influence on utilization (McColl, 2005; Kroll et al., 2006), cost, and health outcome. The longer the distance, the costlier the service in time and price, and the less the utilization. A study (Awoyemi et al., 2011) in Nigeria revealed that approximately 62.1% of the households who live within 0-4km of a public health centre utilize government facilities while others who live farther off seek healthcare from traditional health centres or resolve into self-medication. Reduced travel time or distance between patient location and healthcare facility location implies that healthcare is geographically reachable as an indication of good access.

2.3.2 Measuring Spatial Accessibility

There are several techniques of measuring spatial accessibility, encompassing both simple and sophisticated methods. Spatial accessibility measure is based on two major concepts (Levesque et al., 2013):

- i location where the activity takes place
- ii distance which is the separation between locations

Examples of geographic or spatial accessibility measures widely used in the literature are:

- Supply/Demand ratio
- Distance or Time measure
- Gravity Model
- Floating Catchment Area

2.3.2.1 Supply/Demand ratio

One common approach to planning HCFs distribution is a simple supply–demand ratio within a predefined administrative boundary such as county or Local Government Area (LGA), based on the ratio of a particular service such as number of physicians, clinics, or hospital beds to population served. This measure has been applied in some research. For example, Abbas et al. (2012) considered a HCF to population ratio of 1 to 500 in HCF distribution in Chikun LGA of Kaduna State, Nigeria. Ali and Onokala (2008) used population to medical doctor, nurse and bed space ratios in assessing adequacy of healthcare facilities in Enugu State, Nigeria. Unfortunately, this indicator for assessing healthcare provision and coverage has some drawbacks. For example, they do not consider the spatial interaction between demand and supply, and they do not incorporate travel impedance.

2.3.2.2 Distance or Time Measure

Distance can be a straight-line distance (Love and Lindquist, 1995), also referred to as Euclidian distance or "as the-crow-flies". Distance can also be network distance (Brabyn and Skelly, 2002). The distance factor can also be evaluated in terms of travel time (Schuurman et al., 2006; Ray and Ebener, 2008) and/or travel cost (Burns and Golob, 1976; Breheny, 1978). Travelling time has been suggested by Ray and Ebener (2008) to be more preferable to distance when measuring proximity to healthcare services because transportation mode is taken into account. Liu and Zhu (2004) consider travel mode an important aspect of accessibility. Drive times is a preferable measure in

a structured setting (Jordan et al., 2004) with extensive availability of vehicular transportation and good roads, but may not be suitable where there are no good roads and most people cannot afford vehicular transportation; and a large number of people either walk (Noor et al., 2006) or travel by unstructured public transport. The choice of method depends on the prevailing situation being analysed and availability of dataset.

- Straight-line distance/time measure:** The straight-line or Euclidean distance measure is a simple distance or time measure that has an advantage of minimal data requirements (Love and Lindquist, 1995). Spatial separation between population location and HCF location is measured as the straight-line distance between the two locations (Douglas and Peucker, 1973). While previous research (LaMondia et al., 2010) has also shown that this measure can give more accurate value than network distances in areas with grid-like street patterns, it is not appropriate for countries where streets are not well structured due to planning deficiencies. Other disadvantages are that distance in straight-line does not accurately reflect travel patterns, service use patterns, topography and barriers; it assumes uniform utilization rates of facilities within the catchments, and utilization of the nearest facility to patient's location (Alegana et al., 2012; Cromley and McLafferty, 2011; Gething et al., 2004). Also, different direction of travel is not considered. This may lead to incorrect travel time or distance estimation which is a fundamental factor in accessibility measure.
- Network distance/time measure:** Network measure may be more difficult to define and analyze, it offers a more realistic measure of spatial accessibility because the actual road network travelled from one location to another is considered. It therefore provides a better representation of travel times and distances, and offers a better method for accessibility application in healthcare service utilization, availability and locational-allocation planning. Network analysis has been applied in various studies using road characteristics such as travel speeds (Brabyn and Skelly, 2002) or public transport availability (Djurhuus et al., 2015) to gauge time taken to travel between two locations.

Figure 2.1 illustrates the network and straight line journey from service location to demand location.

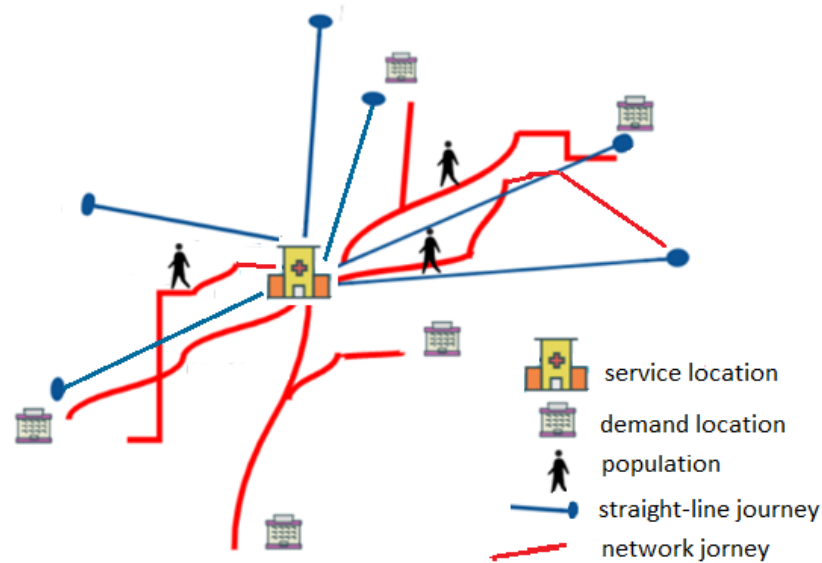


Figure 2.1: Network versus Straight-line journey

2.3.2.3 Gravity Model

The gravity model is a spatial interaction model that is similar to Newton's law of gravity that says the force of attraction between two objects is directly proportional to the product of their masses and inversely proportional to the distance between them (Alegana et al., 2012; Schuurman et al., 2006). In its application to healthcare accessibility measure, patient's spatial interaction with healthcare is determined by representing force of attraction with the flow from the patient origin to the healthcare service; and the body masses are denoted by different utilization factors such as cost of service, type of service and size of the HCF. Based on this concept, things that are close are regarded as having a greater attraction than things at a distance, and attraction is viewed as decreasing with increase in separation. Although this method integrates distance, demand location and supply location in its analysis, it does not account for the availability of service to demands within a defined boundary.

2.3.2.4 Floating Catchment Area (FCA)

The Floating Catchment Area (FCA) method is an improvement over the gravity model, and hence over the demand to supply ratio method. The method is based on the gravity model, and therefore has most of its advantages. Spatial accessibility is measured as a ratio of healthcare provider to population within a defined catchment or service area. For example, the healthcare workers and resources can be allocated in areas where needs are revealed and optimal locations for new facilities identified based on service area or catchment zones (Boulos et al., 2001). The definition of catchment may be based on travel time or distance measured on network or straight-line. There have been several modifications to this method in the literature (Luo and Qi, 2009) which

are collectively referred to as the FCA methodologies family (Vo et al., 2015).

Spatial accessibility measure requires complex and large spatial datasets, and is usually analysed with GIS techniques (Vadrevu and Kanjilal, 2016; Higgs et al., 2019; Gilliland et al., 2019; Olowofoyeku et al., 2019). However, little attention has been given to the fact that:

- decision on the method to adopt for accessibility measure cannot be isolated from technical and resource availability.
- the GIS techniques are static.

The Spatial-ABM technique for accessibility measure for this research shall be adaptable and flexible (Li et al., 2011) by considering both straight-line and network travel time to determine the catchments of HCFs for the FCA analysis. The model's flexibility shall allow analysis on varying travel time and average travel speed threshold. The supply-demand ratio within the catchment of the HCF will reveal if a HCF is under-served or over-served.

Catchment Area Estimate: The healthcare service area is the zone around a healthcare facility location that enhances the analysis and understanding of health outcomes (Zinszer et al., 2014) and improvement of healthcare service delivery. Such analysis can reveal under-served communities. Catchment area of healthcare services can be determined using constructed polygons based on straight-line or road network (Carlson et al., 2011). The catchment can be determined centred on either the facility or the population location. One example of straight-line methods is the drawing of circular buffer around a HCF using arbitrary distance radius as travel distance. Many works have used this approach (Ayoade, 2016; Fadahunsi et al., 2017; Ismaila and Usul, 2013; Olowofoyeku et al., 2019). Such works have overlooked the significance of road network datasets while trying to be as simple as they can, hence have produced results that do not provide optimal access. A small number of other works have used the GIS network-based analysis to determine the catchment area (Isma'il and Musa, 2014), which is a more realistic method.

The method proposed in this research for estimating catchment area proffers an improvement over the traditional GIS circular buffering method, and a better alternative to the sophisticated network method that requires quality and costly input data, and expert knowledge in GIS. A non-static, flexible and interactive travel-time model is developed to demonstrate the applicability of GIS and ABM in patient travel analysis to determine the catchment area of HCFs.

2.4 Facility Location Problems

Facility location problem involves the search and identification of a specific position or place for a particular activity or function from which demands that are spatially distributed can be serviced (Church, 1999). It can be applied to aid decision in both private and public sectors within a region, a community or a country. Facility location problem is believed to have its origin in the theories of Johann Heinrich von Thünen, Alfred Weber and Walter Christaller. Heinrich's theory in 1826 was based on agricultural land location where transportation costs were said to be related to distance travelled and the volume of good shipped. Weber in his location theory in 1909 claims that the location of manufacturing industry depends on availability and location of raw materials and is influenced by the transportation costs (Friedrich and Weber, 1929). He established that the weight of the raw materials and the final commodity are important factors in the transport costs and the location of production. It costs less to transport commodities that lose mass during production from the production site to the market than transporting from the raw material site to the production site. Walter Christaller also formulated the central place theory in 1964. He related settlements' economic activities with the population by considering the growth and development of towns, human behaviour and fundamentals of economics. Using the settlement pattern and size of southern Germany, he used geometric shapes to model pattern of settlement locations (Christaller, 1966).

These theories formed the foundation of several other location models for different applications and purposes, and have led to different classifications of the location problems: depending on the objective to be achieved and problems aimed at solving. In measuring the effectiveness of facility location, a facility is regarded to be more accessible if the total weighted distance or time is comparably smaller. Facility location models can be discrete or continuous space site location. The discrete site selection assumes that the location of potential facilities are predefined and finite, while the continuous site selection also referred to as planar location (Church, 1984) assumes infinite number of potential locations for facilities anywhere in the geographic space. The classifications of facility location problems are mainly variants of two main classifications which are described in the following sections: median-based problems and covering-based problems (Ahmadi-Javid et al., 2017; Persson et al., 2013; Church, 1999).

2.4.1 Median Problems

The median problem introduced by Hakimi (1964) seeks to minimize the total (or average) distance or time between service and demand. The median-based facility location model is a single-objective model based on locating a pre-specified number of facilities (usually denoted as p) at locations that ensures that the average distance between demand locations and their closest facility is minimized. The p -median model has been widely applied to facility location problems. For example, ReVelle and Swain (1970) formulated a p -median model for locating central facilities using linear programming.

The p -median model is one of the most popularly applied models in healthcare facility locations and planning. For example, Caccetta et al. (2005) developed p -median heuristics for locating emergency facilities. Mehretu (1985) formulated a p -median model to locate health clinics in Bourkina Faso to ensure that no population travelled more than a maximum distance of 5 km and that every village is assigned a health clinic. Beheshtifar and Alimoahmadi (2015) developed a spatial optimization model with the p -median problem to locate clinics in Tehran, Iran.

2.4.2 Covering Problems

Covering problems (also called maximum service distance) involves locating facilities to cover all or most demand locations within a specified service distance or time to ensure service coverage. This comes from the concept that the demands can be better served if they are closer to a service. Coverage or closeness to a service can be measured in terms of the service distance or service time, which is the maximum distance or time between a facility and a demand location serviced by that facility. A demand location is therefore regarded as being covered by a facility if the travel distance or travel time between the demand location and the facility location is less than or equal to a specified service distance or time. Any demand location that is farther than the service distance or time to a facility is regarded as not being covered by the service. Many covering location models have been developed to address shortage of HCFs. These include location of emergency service facilities such as ambulance deployment (Yin and Mu, 2012) and response to disaster onset (Balcik and Beamon, 2008; Jia et al., 2007); location of healthcare waste disposal facilities (Chauhan and Singh, 2016); and location of public healthcare clinics (Ye and Kim, 2016; Shariff et al., 2012).

Covering-based problems are further categorised as Location Set Covering Problems (LSCP), Maximal Covering Location Problems (MCLP) and p -Centre Location Problems (PCLP):

Location Set Covering Problems (LSCP): While the p -median model tries to ensure that the demand gets to the nearest facility within a short time, it does not account for users that are far from the closest service and may not likely utilize the facility. This setback leads to the formulation of the LSCP that considers maximum distance or time to ensure that every demand location is covered by at least one facility (Toregas et al., 1971).

LSCP model seeks to minimize the number of facilities or facility locations required to ensure the total demand coverage with a specified service distance or time. The LSCP tries to minimize the cost of locating facilities and at the same time ensures a specific coverage level within the maximum distance or time between the facility and the demand location. The facilities are located such that no demand travels farther than the maximal service distance or time to receive service. In the LSCP, the only cost factor that is considered in the decision making is the number of facilities (Church and ReVelle, 1974), with no distinction between large and small demand location (Daskin

and Owen, 1999).

Maximal Covering Location Problems (MCLP) was developed by Church and ReVelle (1974) to increase the cost factor of the decision making process of the LSCP. They observed that out of many solutions of a particular LSCP, it is possible that the minimum service distance that can provide total coverage for a fixed cost (number of facilities) is quite larger than the desired service distance, D . This may lead the decision maker to shift focus from total coverage with the larger service distance, to total coverage with the desired service distance, D . This will require more facilities and expenditure. With limited resources, the resort may be to cover as many demand as possible within D and the available resources. This led to the formulation of the MCLP by Church and ReVelle (1974) that seeks to locate p facilities to maximize the demand or population covered within a specified service distance or time. In this case, if the available resources cannot meet the desired total coverage level, the objective is relaxed to provide the coverage to as many demands as possible. In other words, the total coverage requirement within the exogenously specified distance is relaxed. The MCLP has been applied in many health service facility location problems. For example, Meskarian et al. (2017) used the MCLP for planning clinic locations for sexual health services in Hampshire, United Kingdom.

p -Centre Location Problems (PCLP) seeks to minimize the maximum travel distance or time between demand location and facility location. In this case, service distance or time is not known. It is a location-allocation problem that simultaneously finds optimal locations of facilities and allocates demand to those facilities. In this circumstance, all the demand locations must be covered within an endogenously determined maximum distance. Lu (2013) developed a p -centre model to locate relief distribution centres for emergency natural disaster on-set that was applied to earthquake case in Taiwan.

The facility location problem for this research falls in the category of covering problems because it seeks to provide service coverage to the population, with the objective function of maximizing coverage within a service distance. It also seeks to minimize the maximum travel distance between the farthest demand and its closest facility.

2.5 Considerations for Locating Healthcare Facilities

Despite the considerable uncertainties and imbalance underlying real-life decision-making process of location selection for HCFs, most location models are focused on the intrinsic assumptions of proportionate basic input parameters, such as the population size, HCF capacity, the number and the location of HCFs, and the service distance. These models take these parameters as static and constant inputs, and tend to ignore the reality varying population sizes and access to health services that affect the distribution of scarce service providers, HCFs and resources. These location problems are solved as static (Ahmadi-Javid et al., 2017) or deterministic (Boonmee et al., 2017)

location problems. For example, Dekle et al. (2005) used the deterministic approach to locate disaster recovery centres in a Florida County. The objective was to minimize the total number of recovery centres that will cover the residence within fixed distance radius. Also, Ye et al. (2015) developed emergency warehouse location problem (EWLP) as an extension of the P-centre problem and used Variable Neighbourhood Search (VNS) heuristic to locate fixed number of emergency warehouse within given coverage radius in China. These models assumed equal capacity for the facilities. However, Wang and Ma (2018) in their deterministic model considered unfixed capacity in location-allocation for locating nursing homes for the elderly people in Kongjiang Road area, Shanghai. The objectives were to minimize the total construction costs and minimize the total weighted distances from the nursing home to the community. Kim and Kim (2013) also solved a public healthcare facility location problem using Lagrangian heuristic algorithm applied to North and South Chungcheong Provinces, Korea. The aim was to maximize patient coverage with fixed budget constraint. The model considered allocating low-income patients to the public HCFs and high-income patients to both public and private HCFs.

In real-world situation however, healthcare facility locations are characterised with uncertainties and are therefore not often predictable (Afshari and Peng, 2014). Interest in recent facility location problems has shifted to such uncertainties which in turn have influence in the input parameters. The integration of the varying uncertainty factors has been identified as a major challenge in real-world location problems (Daskin and Owen, 1999; Snyder, 2006). Different approaches have been applied to address this in HCF location decision (Snyder, 2006). Murawski and Church (2009) identified the lack of all-weather roads in the under-developed countries where changes in weather conditions such as rains affect the states of the roads, and therefore has impact in accessibility. The authors developed the Maximal Covering Network Improvement Problem (MC-NIP) model to predict the accessibility to health services based on road improvement and applied it to a rural area in Ghana. Mestre et al. (2015) developed two stochastic location-allocation models for planning hospital networks re-organization under uncertainty associated with the demand and supply of hospital services, when the decision maker considers improving geographical access, while minimizing costs. The models were applied to the Portuguese National Health Service. Harper et al. (2005) also developed a discrete-event geographical location-allocation simulation model using stochastic approach for location of health service centers that considered variable patient flows, travelling times, and transport preferences, including different services. Also, in their work, Wang et al. (2018) addressed uncertainties in medical demands and costs with a bi-level multi-objective particle swarm optimization (BLMOPSO) using fuzzy and stochastic factors that was applied to a case study in China.

Another limitation in the traditional covering HCF location models is the modelling of the real-world as homogeneous system where the environment is assumed to have similar characteristics. Spatial heterogeneity of the real-world in terms of varia-

tions of environmental characteristics or conditions (roads, vegetation, wetlands, population distribution) influences the HCF locational choice, resources, demand allocation, travel time, service area, service distance, and number and location of HCFs. Spatial heterogeneity has been identified to influence imbalance in accessibility to healthcare services (Yin et al., 2018) and ambulance service time (Leknes et al., 2017). The simplification of the complex real-world heterogeneous problem are location models with the same geographic characteristics. However, each HCF has a unique combination of population and spatial features within its catchment that results in distinguishable capacity, size of healthcare workers, and resource allocation. Therefore, HCFs characteristics are location-dependent based on the social, economic, demography and geographical environment (Zhang et al., 2016).

Although the number of facilities, for example, may be suggested based on a facility density ratio as a function of the aggregate population size, it is evident that the spatial distribution of population is heterogeneous (Su et al., 2010). Such distribution has a direct impact on the population size within a setting. A group of people may dwell within certain regions based on ethnicity or religious belief, while a settlement may be influenced by economic activities and geographical features. For example, fishing activities may be a factor for dwelling close to rivers. Other settlement pattern may be due to commercial activities and proximity to roads, schools or workplaces. Few works have considered spatial heterogeneity and combined GIS with other locational techniques (Malczewski, 1999). In this regard GIS plays a valuable means of informing the decision. For example, Zhang et al. (2016) considered the heterogeneous spatial distribution of population and economic development in Hong Kong, that is characterised with mountainous topography. The authors used Genetic Algorithm-based multi-objective optimization in locating HCFs and allocating demand to the HCFs.

Decisions for establishing HCFs is typically a long-term commitment of substantial resources - technology, equipment, human and infrastructure. Such decisions may be very difficult to reverse or change after it has been implemented. Static facility location may result in waste of scarce resources and redundant expenditure due to inequity in spatial circumstances. HCF planning should possess adaptation to changing socio-economic and geographic conditions. Due to the complexity of such decision making in site selection for HCFs, integrating GIS and ABM as proposed in this thesis provides alternative adaptable solutions to aid decision making that combines such factors. The heterogeneous spatial characteristics and population distribution are considered to flexibly suggest the number and location of HCFs, and the service distance or travel time. Agents are capable of sensing and adapting to changes in the environment. For example, changes in the population size, changes from land to water, changes from committed to uncommitted zones. As opposed to static modelling, HCFs represented by mobile agents can change their locations. The emergent behaviour from these interaction with environment changes results in different service area, service distance, HCF network, including location and number of HCFs. This adaptability also makes it possible for the agents to change their states, reflecting the current HCF capacity that

may induce additional HCF or a reduced resource allocation.

Adaptability models are robust and consider different scenarios and conditions (Daskin and Dean, 2005). In this case, a perception of the real-world system, which is sometimes difficult or impossible to predict, is considered in suggesting HCF types. HCFs that do not meet the requirements for standard HCF in the current circumstances can serve as temporary medical centres or community health clinics. Upgrade or addition of resources or HCF therefore may suffice with time. This is especially useful in planning, considering the political, budget, economic and policy contexts in the provision of complete healthcare coverage for all populations.

2.6 Covering Problem Solution Models

Facility location models are developed upon the problems which they intend to solve and the objectives to accomplish. In the literature, different facility location models have been proposed for different applications (Rahman and Smith, 2000). Covering location models for locating healthcare facilities can be of varying objective functions. For example, the objectives can be to minimize travel time or distance between facilities and demand locations, minimize costs of facility establishments, and/or maximize coverage of population to be served by the facilities. Ye and Kim (2016) developed covering model to locate HCFs in Hillsborough County, Florida. The model used network-based distance to determine coverage. Zarandi et al. (2011) and Galvão and ReVelle (1996) also developed a covering model with MCLP.

As solutions for addressing the challenges of locating HCFs to provide universal health coverage, the spatial optimization model for this research is two-stage, and falls mainly in the category of covering-based problem, specifically the MCLP and PCLP.

The mathematical formulation of the MCLP (Church and ReVelle, 1974) and the PCLP are considered for this research and provided as follows:

2.6.1 Formulation of MCLP

The MCLP can be formulated using the notations below:

I = set of demand locations or nodes i

J = set of potential facility sites j

h_i = total sum of population to be served at demand location i

p = total number of facilities to be located

d_{ij} = the shortest distance between demand location i and facility location j

D = the maximum acceptable travel distance between demand i and facility j

$$N_i = \{j \in J | d_{ij} \leq D_{ij}\}$$

X_j = binary decision variable indicating if the facility is located at site j , 0 otherwise

Z_i = decision variable is 1 if location i is covered; 0 otherwise

N_i is the set of facility sites that are eligible to provide cover to demand location i . When the closest facility to a demand location is less than or equal to D , such location is regarded as been "covered". When the closest facility to a demand location is at a distance greater than D , then the location is regarded as "uncovered".

The objective function of the model is:

$$\text{Maximize } \sum_i h_i Z_i \quad (2.1)$$

Subject to:

$$\sum_j X_j \geq Z_i; \forall i \in I \quad (2.2)$$

$$\sum_j X_j = p \quad (2.3)$$

$$X_j = \{0, 1\} \quad \forall j \in J \quad (2.4)$$

$$Z_j = \{0, 1\} \quad \forall i \in I \quad (2.5)$$

The objective (2.1) maximizes the total covered demands within the desired service distance. Constraint (2.2) is the constraint that describes the relationship between the coverage and location variables, and states that demand location is covered if it is covered by at least one facility. Constraints (2.3) limits the number of facilities to p . Constraints (2.4) and (2.5) are integrity constraints.

The solution to the problem specifies both the largest amount of population that can be covered, and the p facilities and their locations that provide the maximum coverage.

2.6.2 Formulation of PCLP

Using the following notation:

I = set of demand locations or nodes i

J = set of facility sites j

p = total number of facilities to be allocated

d_{ij} = the maximum distance or time between demand location i and facility location j

S_{ij} = the maximum acceptable travel distance or time between demand i and facility j $d_{ij} \leq S_{ij}$

N_i = set of all candidate facilities which can cover demand location i

X_j = binary decision variable indicating if the facility is located at location j or not

Z_i = decision variable is 1 if location i is covered; 0 otherwise

The objective function of the model is:

$$\text{Minimize } W \quad (2.6)$$

Subject to:

$$\sum_{j \in N_i} y_{ij} = 1 \quad \forall i \in I \quad (2.7)$$

$$\sum_{j \in J} x_j = P \quad (2.8)$$

$$y_{ij} \leq x_j \quad \forall i \in I, \quad \forall j \in N_i \quad (2.9)$$

$$W \geq \sum_{j \in N_i} d_{ij} y_{ij} \quad \forall i \in I \quad (2.10)$$

$$x_j \in \{0, 1\} \quad \forall j \in J \quad (2.11)$$

$$y_{ij} \in \{0, 1\} \quad \forall i \in I, \quad \forall j \in N_i \quad (2.12)$$

$$W \geq 0 \quad (2.13)$$

The objective (2.6) minimizes the maximum distance or time between a demand location and the nearest facility allocated to it. Constraints (2.7) ensure that each demand location is covered by only one facility. Constraints (2.8) specify the total number of facilities that demand will be allocated to. Constraints (2.9) indicate that demand locations are only covered by established facilities. Constraints (2.10) enforce that W is the maximum distance or time from demand location to facility (W is the auxiliary variable used to determine the maximum distance). Constraints (2.11), (2.12) and (2.13) are domain constraints.

2.6.3 Maximal Covering Problem with Mandatory Closeness Constraints (MCPMCC)

While the MCLP seeks to maximize coverage within a specified service distance, D such that as few people as possible are outside coverage, the probability of distance between facility and certain uncovered demand being greater than a desirable value cannot be overlooked. As a fair consideration to the few population not covered, the solution to the MCLP should also ensure that no demand is farther from a facility beyond certain distance, S (where $S > D$), where D is the desired service distance for the maximal covering location problem. Church and ReVelle (1974) called this problem Maximal Covering Problem with Mandatory Closeness Constraints (MCPMCC) and stated the problem as:

Locate a fixed number of facilities in order to maximize the population covered within a service distance D , while maintaining mandatory coverage within a coverage distance S (where $S > D$).

Two types of covering constraints are now included in the MCLP, which are:

$$\sum X_j \geq Z_i; \quad \forall i \in I \quad (2.14)$$

$$\sum X_j \geq 1; \quad \forall i \in I \quad (2.15)$$

where:

$$N_i = \{j \in J | d_{ij} \leq D_{ij}\}$$

$$M_i = \{j \in J | d_{ij} \leq S_{ij}\} \quad N_i \subseteq M_i$$

Constraints 2.15 ensure the mandatory closeness constraints for each demand location i : there is at least one facility within distance S .

The MCPMCC can provide an improvement to the MCLP by increasing the coverage of demand from the feasible solution of D service distance with additional coverage with S mandatory closeness distance. If small number of facilities, p' provides total coverage of demand with S distance units, then a solution to the LSCP has been obtained. Example of such solution was demonstrated by Church and ReVelle (1974) using linear programming to obtain optimal solution. In their solution, the value of S is pre-defined and used to obtain the required number of facilities by solving the LSCP to obtain the smallest number of facilities necessary to cover all demand within distance S . A maximal covering solution within D is then obtained with the number of facilities while specifying a mandatory closeness of S for demand locations beyond D coverage. The problem with this solution is that the number of facilities cannot be increased and a larger maximal coverage can only be achieved if the mandatory closeness constraint of S distance is relaxed. This may not result in total coverage, but improved percentage of coverage to demand locations.

In the centre HCF location, a total coverage is desired. It is possible to maximize coverage with $D(< S)$ distance, and still maintain total coverage with mandatory value of S distance with different feasible solutions. In situations where p^* number of facilities exists, the uncovered demand with D distance can be allocated to p^* facilities in addition to the set of p' facilities that have been selected, such that the maximum distance between each demand location and its closest facility is not greater than a specified value of S , ($S > D$) distance and ensure that all demand locations are covered. Although the exact value of S is not known, but a set of solutions that makes S as close as possible to D can serve as feasible solutions to help decision makers in planning facility locations that will ensure maximum coverage within a service distance coverage D , and still achieve total coverage of all demands within S as proposed in this thesis. If S is not predefined, a desired total coverage can be achieved through a generalised model that endogenously provides a range of S distance values from feasible solutions that will give the decision maker the flexibility of choosing a combination of S distance units and number of facilities that will ensure total coverage. With the existing facilities p^* , the number of facilities can be increased to $p^* + p'$, and demand locations assigned to any of the facilities that is closest to their locations.

2.6.4 Capacitated Facility Location

A key limitation in the work of Church and ReVelle (1974) is that their model did not consider the capacity of the facilities in the solution of the MCLP. While this is an essential consideration in facility planning, many works in the healthcare facility location literature have assumed that the facilities have unlimited capacity to serve all demands - a situation referred to as "uncapacitated" facility location problem (Farahani et al., 2012; Rahman and Smith, 2000; Galvão, 2004). In reality, healthcare facilities do not have infinite capacity to serve all demands. The size or capacity of a facility may be measured in terms of number of admissions, number of population to serve, number of healthcare workers, or number of beds. In this case the facility location analysis

is "capacitated" (Bolouri et al., 2018; Haghani, 1996). Although Rahman and Smith (2000) are of the opinion that it is a common practice in developing countries that rural healthcare facilities have similar medical equipment and number of healthcare personnel, there is a limit to the number of patients that can be attended to in reality.

The most common way of integrating facility capacity in healthcare facility location models is to include a capacity constraint in the model and represent it with population data that addresses the required capacity. Stummer et al. (2004) for example, consider a maximum number of beds in the hospital capacity constraint in their multi-objective optimization model to locate medical departments in a hospital network. Similarly, Yin and Mu (2012) developed a model which they refer to as Modular Capacitated Maximal Covering Location Problem (MCMCLP) to ensure that the size of population of the State of Georgia is considered in location-allocation of emergency vehicles for healthcare services. This is based on the assumption that the quality of healthcare service reduces when a facility exceeds its capacity limit. Hence, to improve the decision-making reality of facility placement, it is important to include the size of population to be served by that facility in the process of allocation.

Haghani (1996) in addressing the mandatory closeness of uncovered demand to facilities proposed a capacitated maximum covering location model that maximizes the coverage of demand, and at the same time minimizes the average distance from uncovered demand to the located facilities with excess capacity. Unfortunately the model does not consider the spatial distribution of the uncovered demand and the facilities with excess capacity. If there exists any other facility that is closer to demand than the facility with excess capacity, such facility can likely be utilised. Another issue with this model is that there is no feasible solution if the total maximum capacity of the facilities is less than the total demand. Although the author recommends the location of a dummy facility at an imaginary site to accommodate all of the excess demand. This solution is also not feasible where demands are located at considerable distances apart, and separated with access barriers such as rivers or mountains.

2.6.5 Suitability Analysis

An important factor that needs to be considered in site selections for HCFs is the suitability of the suggested location in terms of environmental and geographical features such as topography; water body, such as river or pond; and proximity to other features and amenities. It is unrealistic to locate healthcare facilities inside water body, such as river or pond. In like manner HCFs are expected to be close to some areas where vulnerable groups can easily receive treatment. A survey of healthcare provisions in the schools in Ogun State, Nigeria revealed a shortage of healthcare personnel in schools (Kuponiyi et al., 2016). Keeton et al. (2012) observed that having healthcare close to schools has revealed improved access to healthcare, health and education outcomes, and satisfaction level. World Health Organization (WHO) (1998) has suggested that primary HCFs should be close to schools and religious facilities. These

considerations have not been widely considered in locating HCFs. It is therefore important that location-allocation models integrate these constraints to assess suitability of the site choice. Suitability analysis is usually incorporated in GIS tools to create a suitability map (Carlson et al., 2011; Church, 2002) by combining different spatial data layers of the suitable and unsuitable features. Such map can be integrated in facility location selection model. For example, Beheshtifar and Alimoahmadi (2015) developed a multi-objective optimization model that integrates GIS site suitability and land-use compatibility analysis for healthcare facility location as one of the model objective functions.

2.6.6 Facility Type With Establishment Cost Constraints

Another important criterion in healthcare facility location is the minimization of cost. Cost may be defined by the cost of establishing the facility in terms of cost of land or resources (Beheshtifar and Alimoahmadi, 2015). The decision maker may consider locating more number of HCFs if the total cost of establishing those facilities is less than establishing fewer facilities. This is possible if weight is assigned to the carrying capacity of each HCF. HCF serving population less than a threshold value is assigned a lower cost of establishment and resource allocation.

2.6.7 Unified Approach to Covering Model

Many facility location models satisfy the objectives of the different location problems, either as variants or extensions, or they are related to one or more of the problems (Lei et al., 2016). The findings of Church and ReVelle (1976) show that the MCLP is a special case of the p -median problem. Daskin and Owen (1999) formulated the partial p -centre and partial set covering model that involves partially solving a series of maximal covering problems. For a combined solution of the LSCP and maximizing the number of demand locations with multiple coverage, Daskin and Stern (1981) developed the Hierarchical Objective Set Covering (HOSC) model for emergency medical service vehicles in Austin, Texas. Jia et al. (2007) used a general model to solve the MCLP, p -median problem and p -center problem for locating medical service facilities for large-scale emergencies.

The problems identified in this research are proposed as special cases of the MCLP and PCLP with objectives to:

1. maximize demand coverage within a desired service distance by locating healthcare services to ensure maximum percentage of coverage
2. minimize total number of HCFs needed to cover all demands
3. minimize costs of establishing HCF if the allocated demand is less than a specified threshold number of population

4. minimize the number of uncovered demand outside the desired minimum travel distance or time coverage
5. minimize the maximum travel distance or time between demand location and HCF location
6. maximize covered demand while minimizing the maximum distance from uncovered demand to the closest HCF to ensure total coverage of healthcare services

This complex and computationally intensive multi-criteria facility location problem in this thesis is proposed as Capacitated Mandatory Closeness Maximal Covering Location Problem (CMC-MCLP). An adaptable spatial agent-based optimization solution model is proposed.

2.7 Large Neighbourhood Search (LNS) Solution Algorithm

Solutions to facility location problems can be provided with different optional means using exact, heuristic or meta-heuristic. In exact method, a set of locations are found as solutions and can be confidently regarded as an optimal solution. Exact solutions are generally used in small-size problems. With an increased size of problem and complexity, this solution ceases to be efficient, however review shows that a large number of facility location problems employ the exact solution (Afshari et al., 2014; Ahmadi-Javid et al., 2017). The different methods that have been used include linear programming, branch and bound, cutting planes, decomposition, Lagrangian relaxation, and dynamic programming.

Heuristic approach is employed as an alternative to exact solution. Heuristic is an approximate method of achieving optimal or near optimal solution (Gu et al., 2010; Kokash, 2005). It is a way of attaining a fast solution at the expense of quality solution. Many applications of heuristics use local search or swapping (interchange) procedure (De Smith et al., 2018). Meta-heuristic technique is a heuristic development outside the local search measures to find a global optimum. One of such meta-heuristic techniques is the Large Neighbourhood Search (LNS). LNS which was first proposed by Shaw (1998) uses iterative process in solving optimization problems. Each possible solution and generation of output serves as input for the next iteration until a solution is found. The LNS procedure starts with an initial feasible solution and iteratively obtains an improved solution to the problem. An approximate solution is first found with a greedy algorithm and further improvements are made with local search in the neighbourhood through a destroy and repair mechanism (Ahuja et al., 2002; Pisinger and Ropke, 2010). The destroy method is stochastic whereby a certain part of the solution is simply destroyed, while the repair method uses greedy constructive algorithm. The greedy algorithm makes a locally optimum choice that looks best at a particular instant, which may or may not yield optimal solutions.

LNS has been largely employed in combinatorial optimization problem, mostly in vehicle routing problem. Demir et al. (2012) observe that unlike other heuristics whose accuracy depends on the quality of the initial solution with increased number of constraints, the LNS recovers quickly and easily from poor initial solution. This was corroborated by Bruglieri et al. (2018) who in their work for relocating vehicles in electric carsharing services find the LNS to be very effective and efficient compared to other heuristics such as Tabu Search. Their works use the Adaptive Large Neighbourhood Search (ALNS) which is an extension of the LNS heuristic that allows multiple destroy and repair methods to be used within the same search (Pisinger and Ropke, 2010).

In this research, the meta-heuristic algorithm for selecting facility location is therefore built upon the LNS algorithm due to its simplicity, accuracy, flexibility and capability to search the neighbourhood of a subset of an initial solution to provide an efficient search of the neighbourhood, and give an improved solution. To the best of the author's knowledge, this approach has not been applied in spatial agent based simulation model for optimizing public healthcare facility locations. The LNS algorithm is presented in Algorithm 2.1.

Algorithm 2.1 Large Neighbourhood Search Algorithm

```

1: input: a feasible solution  $S$ 
2:  $S_b = S$ ;
3: repeat
4:    $S_t = r(d(S))$ ;
5:   if accept( $S_t, S$ ) then
6:      $S = S_t$ ;
7:   end if
8:   if  $c(S_t) < c(S_b)$  then
9:      $S_b = S_t$ ;
10:  end if
11: until stop criterion is met
12: return  $S_b$ 

```

The variable S_b is the initial best solution observed during the search; S is the current solution; S_t is a temporary solution that can either be removed or be promoted to be a current solution.

$d(S)$ is the destroy method that returns a set of S that has been partially destroyed. $r(d(S))$ is the repair method that is applied to the partially destroyed S to give a feasible rebuilt solution. $c(S)$ is the objective value of solution S . A stop criterion will return the best accepted solution.

2.8 Agent-Based Model

Agent-Based Model (ABM) is defined as "a collection of multiple, interacting agents, situated within a model or simulation environment such as represented by the artificial world" (Crooks and Heppenstall, 2012). Agents are defined based on their characteris-

tics which include autonomy, heterogeneity, pro-active or reactive behavior, bounded rationality, communication capabilities, mobility capabilities, and learning capabilities (De Smith et al., 2018). They can be used to represent moving/non-fixed objects such as the population travelling in space, or fixed objects such as a healthcare facility whose location is fixed, but can change state from open to closed. Agents have several characteristics in common which include:

- *Autonomy*: Agents are capable of carrying out activities independently without any influence or control. They can process information on their own and exchange the information with other agents. They also interact with each other and take independent decisions while still maintaining their autonomy.
- *Heterogeneity*: Agents allow other autonomous agents to be developed with varied attributes. Similar agents can also merge together to form a group of agents.
- *Reactivity*: Agents can sense or be aware of their environment and perceive the presence of obstacles and other agents.
- *Pro-active*: Agents can display goal-directed behaviour by taking independent decision to achieve a goal. While solving a problem or making decision or resolving a conflict, an agent is capable of taking independent deliberation (Abar et al., 2017).
- *Interactive*: Agents can interact or communicate with other agents and their environment.
- *Mobility*: Agents can move around the space within the environment. Although, they can remain fixed in the environment. Their autonomous ability to take decisions can make them change locations.

ABMs proffer some advantages over traditional modelling such as:

Representation. Representation is the key to understanding a phenomenon. In ABM, new representation can help solve problems that cannot be solved before, while changing the representation helps to ask new questions. With ABM representation, results can be well communicated. For example, writing a series of equations that describe the particular location of individual agents and how they affect each other and their environment can be very complex. But with ABM, connected interacting parts in a complex system can be easily represented.

Third way of doing science. Two traditional way of doing science are - induction (inferring from particular data to a general theory), and - deduction (reasoning from first principles to a general theory). ABM on the other hand is generative. That is using first principles to generate a particular set of data that create a general theory.

Modelling very complex local interaction. Very complex, but local interaction between individual agents can be modelled in ABM, especially if the interaction can happen over time. For example in a history dependent interaction, two agents in a system can have knowledge of the past history of how they interacted, and this can be used to determine how they might interact in a particular situation.

Physical location. Agents can actually be physically located in the model, for example in an actual real geographical environment such as road, schools, elevator.

Adaptation/Learning. It is possible to design agents to be adaptive, and to adjust or modify their current state subject to their previous states. Agents' adaptation can be at individual level or at group level.

By contrast to GIS static analysis, ABM systems are designed to allow adaptable and flexible analysis. Agents can move and change locations in time steps, not necessarily in geographic space. Although they cannot perform spatial analysis. Agents' movement is of importance in non-static and heterogeneous spatial phenomenon, and with ABM, movement is possible from agent a to agent b , movement can take place from location i to location j and from time t to $(t + 1)$. This can be via links across space and time.

Agents' ability to interact and their intelligence permit a vast range of potential uses in optimization analysis. ABM's flexibility enables potential variables and parameters to be specified and explored with different complex spatial scenarios which may be difficult mathematically. Aringhieri et al. (2007) developed a model to locate ambulances over the urban area of Milano, Italy. Agents represented the call, the operator at the operation center, and the ambulance. The operation center collects a wide amount of data describing the services from the instant in which a call is received by the operator to the time an ambulance leaves the hospital after the service. The categorization of patients is based on the severity of their injury. Hartmann and Zerjav (2014) also used ABM to support decision making in the placement of out-care centers based on proximity and economic factors. Agents represented out-care centers that move to economically superior locations, and patients that choose to visit the out-care centres based on certain parameters. Outside the healthcare service application, Ciari et al. (2011) developed a model using ABM for location decisions of retailers and implemented it into the agent-based trac simulator (MATSim-T). Retailers were modelled as agents with a goal to optimize the location of their shops in order to maximize the number of customers. The simulation provided the agents the opportunity to experiment with different changing factors to provide the opportunity to relocate their shops.

2.9 Spatial-ABM Adaptability for Facility Location

Traditionally as a standalone tool, GIS has been employed to solve healthcare facility location problems either in its simple form or sophisticated form. Works in this area range from emergency healthcare, specialised healthcare, and public healthcare. GIS spatial analyses are performed in a static environment where all factors and parameters are regarded as fixed and constant within the planning period. However, in real-life situation, lots of factors and parameters are likely to change within planning and intervention period. These include resources, economy, demography, and environmental factors. GIS is discovered not to deliver adequate and sufficient performance, especially when large datasets and heterogeneous factors that will entail varying parameters and several numbers of iterations are required. For such situations, more flexibility is more appropriate.

Spatial optimization modelling involves complex processes of movement and change of location which GIS operations cannot perform. One of the ways that have been employed at overcoming these challenges is to set a balance between the benefits of GIS and facility location or optimization models by integrating mathematical simulation modelling or computational algorithms and GIS. For example, Beheshtifar and Alimoahmadi (2015) combined GIS and genetic algorithm in solving optimization problems for locating clinics. Lei et al. (2016) also combined GIS with tabu search for location-allocation analysis.

While the traditional approach to modelling geographical systems treat geographical components as static and homogeneous entities where for example, static aggregation of population is regarded as having the same characteristics, ABM allows the simulation of diverse agents with discrete characteristics to interact with space. At times, agents may need to evaluate the spatial distribution of resources or other agents with respect to their present location by considering different alternative positions (O'Sullivan, 2008), as will be required in the site selection for HCFs. It is therefore pertinent that this research utilizes these agents' characteristics and ABM ability, with other spatial aspects to model the access to healthcare services and determine where new HCFs can be located. With the complexity of the human, social and physical systems, more detailed analysis are required, and attempts are being made to move from static to non-static, aggregate to disaggregate modelling techniques.

ABM-GIS integration is an innovative method that can enhance better healthcare outcomes in terms of distribution and planning to effectively meet the healthcare needs of the population. For example, in answering the question on how to reduce the cost of establishing facilities to accommodate the budget at hand, instead of the GIS static analysis, a realistic solution to such problems is to develop a flexible model that can give planners the choice of either to obtain a total coverage or certain percentage of coverage. Additional flexibility that the Spatial-ABM can proffer is to create HCF that can be established at lower cost if the population within the catchment of such facility is below certain capacity value. WHO country target for HCF density is given as one

facility per 5,000 population (WHO, 2013). This and other WHO (1998) indicators described in Section 1.1 will serve as a guide to determine the population coverage capacity of each HCF.

Agents' autonomy makes them to be distinguishable from their environment by a spatial, temporal, or programmed limit. Recent developments in technology have benefited from GIS and ABM synergy and its application in optimization techniques. For example, Afshari et al. (2014) used ABM in combination with GIS maps to optimize locations of regional warehouses within logistic and services network. Agents represented warehouses in interaction with city agents. GIS-based maps provided information on location and interaction of agents. Barbati et al. (2011) also built agent-based model to solve a single facility location problem with equity objectives.

The application of ABM–GIS integration in facility locations in healthcare delivery is few compared to other areas such as marketing and production planning (Barbati et al., 2012), and to the best knowledge of the author, the method has not been applied in taking decisions to locate HCFs in Nigeria. Spatial models can involve thousands of agents which can be of varying types and can make complex decisions. For example, the model of a whole city can have pedestrian, vehicle, decision makers, hospital, and school agents interacting with one another and the environment.

The combination of GIS and ABM can give an effective outcome on decisions regarding the appropriate and suitable locations for healthcare facilities by varying the different parameters and attributes of agents, depending on the prevailing situation. Decisions can be taken whether to place a temporary facility now, and plan for permanent structure at a later period based on availability of resources and the existing population.

2.9.1 GIS and ABM Coupling

GIS and ABM stand-alone tools can complement each other while taking advantage of their capabilities and overcoming their limitations. They can be combined through coupling. Coupling as defined by Castle and Crooks (2006) is the linkage of two stand-alone systems by data transfer. There are three different ways by which coupling can be done, loose coupling, moderate coupling, and tight coupling:

Loose coupling

Loose coupling entails the sharing of files or data between the two systems, either by generating data in GIS format for inputs in the ABM system or the results of modelling system for inputs into GIS for visualisation and spatial analysis. This type of coupling is widely used as it has the advantages of fast program linking time and requires low level of software development or programming knowledge. Another benefit is that execution error can be easily detected. One major disadvantage is that the speed of execution is slow, with low level of simultaneous execution. An additional software may be needed to integrate the systems. Although Brown et al. (2005) be-

lieve that the computational efficiency of loose coupling is low and existing database query and spatial analysis cannot be used within the model, it poses fewer technical issues due to its ability to easily discover and locate programming and execution errors. Errors inherent in each system can be addressed independently. While the loose coupling process execution speed is slow, the availability of computers with high processors has brought improved computational speed.

Moderate coupling

Moderate coupling lies between loose and tight coupling. The characteristics are medium integration and execution speed, high programming knowledge and low level of simultaneous execution.

Tight or close coupling

Tight or close coupling is the simultaneous execution and communication within the two systems. It requires medium level of programming knowledge and execution speed is fast. The integration speed is slow and error detection is hard. Tight coupling has posed a difficult task to achieve and an alternative approach is to integrate the functionality of one into the other, depending on which is the dominant software. Tightly coupled models will incur more costs for updates because as each stand-alone system is improved and updated, a simultaneous adjustment is expected of the other to keep-up with the pace. This may not be possible, especially when the systems are maintained by different developers.

In comparison with other coupling methods, loose coupling to integrate GIS and ABM is discovered to be more appropriate coupling method for this research taking advantage of its simplicity, cost effectiveness and low level of required programming skill. There are numerous ABM tools and software but not all have spatial file integration functionalities. Of those with such functionalities, (Repast, MASON, NetLogo), the NetLogo tool kit is used for having similar beneficial efficiency as the reasons for using loose coupling. The NetLogo–GIS integration can input and read GIS database in shapefile or raster format, and apply it to the agents' properties.

2.10 Data Representation

Facility location problem is basically geographic, and entails the use of spatially referenced data, such as the location of service, demand or population and road network. A common concept of GIS data is that it consists of both spatial data and attribute data. The spatial data is the geographical location and spatial dimension of things in geographic space, while the attribute data, also called non-spatial or aspatial data is a descriptive information or properties of the spatial data. For example, the decision variables on the location to site a facility correspond to the coordinate pair (x, y) , that define the facility location in geographic space. The solutions of the site of the facilities in the optimization problem would produce a coordinate pair that indicate the geographic location. Geographic data are basically in raster or vector format, illustrated

in Figure 2.2.

2.10.1 Raster Representation

In the raster data representation, the world is represented as a space divided into regular grids of cells or pixels (picture elements). Each cell is associated with a value that describes a feature which may represent elevation, population, rainfall or other conditions. The location of each cell or geographic feature or condition is defined by the row and column number or position of the cell that they are situated in. The spatial resolution of the raster is the area that a cell covers. For example, if a cell covers an area of 2 m x 2 m, the raster image resolution is 2 m, and this defines the accuracy to which the position of a feature can be defined. Cells of similar values represent the same type of feature or characteristics. Raster data are widely useful for storing data that varies continuously in the geographic space such as rainfall, temperature, vegetation, chemical desparation or elevation; and may be in form of aerial photograph or satellite imagery.

2.10.2 Vector Representation

In vector data representation, objects or phenomenon in geographic space are represented by points, lines and or polygons that define their locations or boundaries. The position of each object is defined by its positional value (x, y) coordinates referred to as coordinate reference system.

It is possible to convert one data representation to another. How information is represented in GIS depends on the purpose of analysis and the map scale. Features that are too large to be represented as area can be represented as point features, depending on the scale and type of analysis. For example, a region may be represented as a point feature such as population centroid, if defining the boundaries is not possible due to the map scale or availability of data.

Spatial modelling of travel time is typically done by vector or raster spatial analytical techniques. The Raster technique is based on path distance and cost distance analysis to obtain the minimum accumulative travel cost from a source to each raster cell location. Although raster-based analysis has been found to be faster, scalable and replicable (Mulrooney et al., 2017), its accuracy depends on the resolution of the raster image (Schuurman et al., 2006), and the size of the dataset can be very large because of the stored values for each raster cell. Vector travel time analysis on the other hand, uses Euclidean (straight line) or network distance (represented by linear features) calculation from an origin to a destination, which are represented by point locations. In a related research, Fisher and Lassa (2017) developed a flexible and interactive travel time scenario model using GIS and ABM. Despite its focus on health service access in developing countries, this methodology is raster-based and requires more input data on road classification and travel speed assignment to different road classes and land

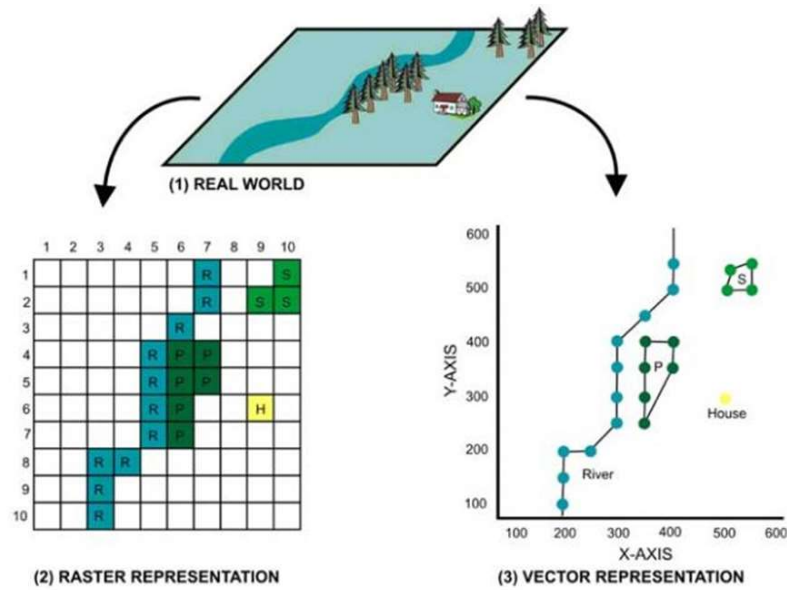


Figure 2.2: Raster and Vector data representation

cover type. This methodology is therefore not appropriate where such information is not available and road surfaces degenerate regularly, hindering standard road classification and speed assignment.

Different works have shown that there is no significant difference between vector-based and raster-based travel time analysis (Delamater et al., 2012; Mulrooney et al., 2017). Thus, the choice of the technique to employ depends on the circumstances, and phenomena to be analysed. Nonetheless, the prevailing situation in the study area requires vector-based analysis because:

- No available data on road characteristics, classification and travel speed for each road segment as required for raster technique;
- Vector-based analysis produces easy visualization to illustrate path and linear features;
- Vector-based analysis gives easy interpretation for non-GIS experts;
- Vector-based analysis is cost effective;
- It is easy to scale vector data down to an individual level; and
- File size is considerably smaller for vector data.

2.11 Summary

This chapter has discussed the theoretical framework for this thesis. The healthcare facilities to be located are identified as health centre facilities that are required to provide coverage to the entire population. Good access and total coverage are therefore important in locating such facilities. The facility location problems are proposed as special cases of the MCLP and PCLP covering-based models. A generalized Spatial-ABM approach to solving the problems is to create a model that will be flexible and adaptable using LNS meta-heuristic method such that: facilities can be located in suitable spatial locations, number of facilities is endogenously determined, types and costs (weight) of facilities are categorized based on capacity, and maximum covering distance is endogenously determined. The service distance D in MCLP is predetermined from the accessibility analysis of existing healthcare facilities. A selected facility can have additional facility within its catchment based on the population size within the catchment. The following chapters describe the methodology and implementation of the models.

Chapter 3

Materials and Methods

3.1 Introduction

Similar to other optimization models, spatial optimization models typically consist of objective functions, decision variables, and constraints conditions. However, in addition to these, a unique geospatial structure is integrated into the spatial optimization model. Consequently, the spatial properties, interdependence and relationship of the objective functions, variables and constraints identified in the theoretical framework discussed in Chapter 2 calls for a new synergy of Geographic Information Systems (GIS) and Agent-Based Model (ABM) in spatial optimization modelling for locating centre healthcare facilities (HCFs). The methodology proposed in this thesis provides an advantage over the traditional practices by combining agents characteristics with the spatial components of the problem in a flexible and adaptable process.

This chapter describes the various datasets required, the data acquisition procedure and the methodology used for creating the models for achieving the aims and objectives of this research. Two types of models were developed:

1. Travel time accessibility model to define the catchment of existing HCFs based on travel time between the HCF location and patient's location.
2. Optimization and location-allocation model for placing new HCFs at locations that will provide maximum coverage at reduced travelling cost to demand, and reduced establishment cost to policy makers.

The outputs of the first model serve as inputs for the second model. A solution to the facility location problems are based on the following objectives:

1. maximize demand coverage within a desired service distance by locating healthcare services to ensure maximum percentage of coverage
2. minimize the total number of HCFs that are required to give coverage to all demands

3. minimize the costs of establishing a HCF if the allocated population is less than a specified population threshold
4. minimize the number of uncovered demand outside the desired minimum travel distance or time coverage
5. minimize the maximum travel distance or travel time between demand location and HCF location
6. maximize covered demand while minimizing the maximum distance from uncovered demand to the closest HCF to ensure total coverage of healthcare services

The proposed framework consists of five major parts, which are:

1. **Spatial data collection:** Spatial data collection involves obtaining spatial data which includes, road network data, existing HCF locations, demographic data, water body dataset, institutional facilities locations such as schools, markets, and religious facilities.
2. **Spatial database creation and data processing:** Database creation and data processing consist of linking spatial data to their geographic locations and attributes; and data extraction and representation as shapefile format to serve as inputs into agent-based modelling environment.
3. **Travel-time accessibility modelling:** The travel-time Spatial-ABM simulation includes simulation of road layers and agents, calculation of travel time and travel distance of simulated patient agents to existing HCFs and obtaining accessibility maps.
4. **Spatial optimization or location-allocation modelling:** Optimization of the location-allocation problem of new HCFs.
5. **Visualization and analysis:** Finally, the outputs from each geosimulation serve as inputs in the GIS platform where they are visualised and analysed.

3.1.1 Research Design

The research design to answer the research questions and meet the objectives of the proposed thesis is presented in Table 3.1.

Further details of these processes are described in the following sections.

Table 3.1: Research design

| Objective | Strategy | Spatial data | Method | Analysis |
|--|---|---|--|---|
| (i) Relationship of travel time or distance with health coverage; How road structure relates to catchment and accessibility | Service area of existing HCF; Area covered by existing HCFs; Is accessibility to HCF the same? | Road network; Existing HCF location; Administrative boundary | Spatial agent-based travel-time model: Input data; Simulate: dwellings and patients; patients traverse road network from HCF to dwellings at specified travel speed and time threshold | Join travel time threshold locations in GIS; Determine HCF service area/distance and area of coverage |
| (ii) Optimize HCF locations for uncovered population | Maximize coverage of HCF within service distance in (i); Determine number of HCF; Ensure: HCFs are not on forbidden locations and are grouped with other institutional facilities | Administrative and coverage maps, average service distance, demographic, Land use/Land cover other institutional facilities (schools, religious, markets) | Spatial agent-based location-allocation model: Simulate adaptable, and non-static HCF agents whose selected locations are governed by spatial rules: distance, adjacency, containment, pattern and intersection ; Locate HCFs within permissible space | Spatial pattern: Test for clustering and spatial autocorrelation of HCF locations with neighbouring geographic features |
| (iii) Ensure total coverage; minimize maximum coverage distance | Link uncovered population in (ii) to the closest HCF | Existing HCF location Optimized HCF in (ii) | Simulate link agents to connect uncovered demand to closest HCF; Obtain the maximum link distance | Minimum maximum covering distance; non-spatial correlation of coverage with HCF locations and maximum distance |
| (iv) Evaluate under-served, over-served, and adequate HCFs; Availability of other institutional facilities (schools, religious, markets) | Determine capacity of HCF with facility/population ratio; Understand impact of population on costs of establishing HCF | Demographic; optimized HCF in (ii); other institutional facilities (schools, religious, markets) | Spatial agent-based location-allocation model: Adapt agents to capacity; Classify agents to low-cost (below) or standard (within) capacity threshold; within institutional facilities | Sort HCF to standard and low-cost, close to institutional facilities or not |

3.2 Spatial Data

All geographic data for these analyses will form map layers in the GIS environment and will be imported to the ABM software. Analyses will be based on travel time/distance, population-facility ratio, and a land-use/land cover dataset. The simulation model shall optimize HCF locations with considerations both for the healthcare policy makers and the population in need of healthcare.

Spatial data can be obtained from primary (survey in terrain using Global Positioning System (GPS) or secondary (scanning and vectorising hard copy maps, or by processing data captured by remote sensing techniques) sources. The choice of data is guided by the World Health Organization (WHO) specifications for decisions on proposed facility location, few of the specifications are adopted. According to WHO

(2018a), definition of catchment area for a HCF may consider:

- Politico-administrative boundary.
- Geographical boundaries that represent natural physical barriers to people's movement.
- Time boundaries that ensures the that HCFs are accessible to the population within the shortest time or distance.

Decisions on locating HCF should consider:

- 15-30 minutes travelling time or about 25km radius in an area with good roads and adequate transportation means.
- grouping HCFs with other institutional facilities, such as religious (church), educational (school), tribal (cultural) and commercial (market) centres.
- avoiding zones liable to dangers of flooding or lowest point of the area.
- avoiding area close to pollution of any kind, including air, noise, water and land pollution.
- servicing HCFs by public utilities such as water (may be well or bore hole), sewage and storm-water disposal, electricity (may be generators), gas and telephone.

The last two factors are not considered for this research due to time and resource constraints. The determinants of location liable to flood would also entail a flood-risk analysis which is also out of the scope of this work. However, water body datasets that ensure that HCF is not located inside water body such as rivers, swamps and lagoons are incorporated as part of the suitability analysis. HCF locations are also constrained within the administrative or geographic boundary of the region of focus.

Although there are no determinants for good road and adequate means of transportation classification, but comparatively the under-developed nations are in dearth of good roads and transportation means. Consequently, the spatial accessibility measure shall be based on travelling time with a model that is flexible for user specification of the travel-time threshold.

3.2.1 Coordinate Reference System

It is important that spatial data to be used in geospatial analysis is referenced to a coordinate system that is based on reference ellipsoid and geoid that approximate the shape of earth's surface. The geoid is the vertical reference datum, while the ellipsoid is the horizontal reference surface (Figure 3.1). A reference system serves to locate the position of an object in space. The reference ellipsoid (also called spheroid) is defined

by a set of initial values and parameters on which measurements in the system are based (shape, size and orientation).

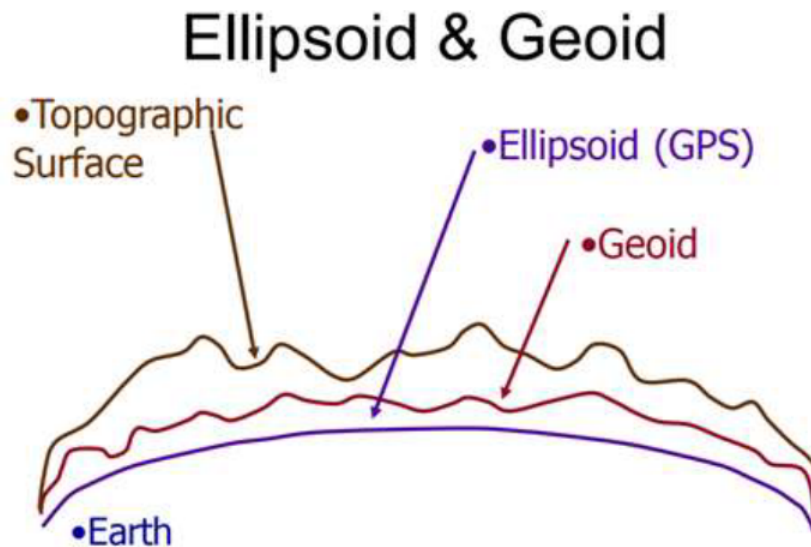


Figure 3.1: Geoid and Ellipsoid Reference Surface

There are different reference surfaces or ellipsoid of varying parameters that are used to approximate the earth's shape. While there are local ellipsoids that are adopted by different countries or geographic regions, there are also global reference ellipsoids among which is the World Geodetic System 1984 (WGS 84) that the Global Positioning System (GPS) uses, and is regarded as being equivalent to Global Reference System 1980 (GRS 80). WGS 84 was developed and is continually being maintained by the National Geospatial Intelligence Agency (NGA). It provides the common framework for all geospatial information.

Regardless of the type of coordinates being used or adopted, a suitable origin point with specified coordinate axes and the directions of the axes with respect to the earth are required. A pair of distances (arc or linear) from the reference point of an orthogonal system may be used to represent the coordinates. This concept is generally referred to as the Terrestrial Reference System (TRS). The basic coordinate systems used in GIS are geographic coordinate systems (GCS) and projected coordinate systems (PCS). Geographic coordinates use a three-dimensional spherical datum surface and are usually specified in degrees, minutes, and seconds of arc as lines of latitude and longitude. Projected coordinates on the hand project the geographic coordinates onto a two-dimensional plane surface and can be specified in linear units such as metres or feet. Due to the problems and clumsiness of using gridded curved lines to plot locations on flat maps; including calculating distances, directions and areas with spherical coordinates as compared with plane coordinates, the military and cartography officials in Europe and the United States came up with the Universal Transverse Mercator (UTM) coordinate system (Huisman and De By, 2009).

The UTM divides the globe into 60 vertical zones numbered from 1 to 60. Each zone has 6° longitude width and vertical lines that run parallel to the central meridian of the zone, while the horizontal lines run parallel to the equator. Although these vertical UTM lines are not specifically parallel to the equator, neither are they parallel to each other, due to a minor distortion by the flattening. Figure 3.2 shows world UTM zones with Nigeria falling between zone 31N and 33N of the UTM grid, and the study area is in the zone 31N region.

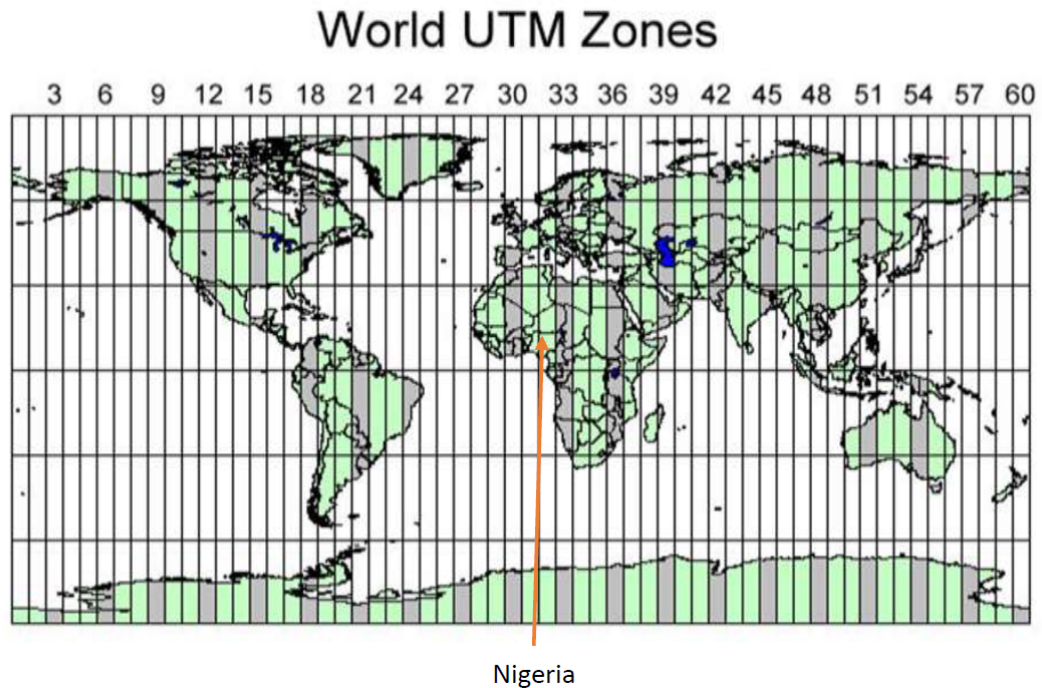


Figure 3.2: World Universal Transverse Mercator (UTM) zones

The two types of UTM are the Civilian UTM Grid Reference System and the Military Grid Reference System (MGRS), having different labelling systems.

A projected coordinate system (PCS) is employed when smaller distances are involved using map projection. Map projection involves the systematic projection of the earth's spherical surface maps onto a two-dimensional plane surface. Examples of map projections are Cylindrical Equal Area, Equidistant Conic, Albers Equal Area Conic, Azimuthal and Transverse Mercator.

Spatial data can be combined from different countries and different sources by data transformation from one reference system to another. These capabilities are within different open and proprietary GIS software. Maps that are used for projects, applications, analysis or services are expected to have a uniform and consistent coordinate reference system. This research uses the global WGS 84 reference ellipsoid which is a universal framework for GIS applications. The spatial datasets are projected on the plane based on the UTM in the zone 31N region where the study area falls.

3.2.2 Spatial Data Acquisition

The data requirement and data collection methods for both vector and raster data needed for the proposed models are provided in this section.

1. Vector data

- **Point features:** The point features are obtainable with the GPS and geocoding processes. GPS is a satellite-based navigation system that determines the location of a user, based on a system of triangulation calculations using information transmitted from satellites. The information is converted to position, velocity, and time estimates by the GPS receiver to provide its location and the location of the transmitting satellite.

The process of transforming an address into a coordinate is called geocoding and is offered by most desktop GIS systems. It involves passing the address of a place to a geocoding service that stores data about the location of addresses. This service is provided by the ArcGIS World Geocoding Service. The geocoder indicates a point on the map to represent the geographic location of the place of interest. Located points are converted to point features and stored as shapefiles. The find tool in the ArcGIS World Geocoding Service can also generate locations of specific places within the map view. This service is especially useful where a large number of point features are required.

The point features for this research are:

- i **Healthcare Facility Locations:** These are the geographic coordinates of existing HCFs.
 - ii **Institutional Facilities Locations:** These are geographic locations of schools, religious facilities and markets.
- **Line features:** Line features can be digitized or obtained from secondary sources. Line features are formed by joining at least two point features.

The line feature for this research include:

- i **Road network dataset:** The network analysis that measures travel time along road paths in GIS require some high level of sophistication, including good quality and costly input data. Road network data will need to be accurately digitized to ensure that there is no gap between links that should be connected by nodes. Database will also be created to include attributes such as road length, the coordinates of nodes that form the shape of the paths, and the lengths of lines or links to the nodes. This procedure will not only be costly, but will be time consuming as well when human and technology resources are scarce. A consequential effect of this is a setback on research and development in

resource and technology constrained settings such as Nigeria.

Although, free routing software can produce travel time estimates, they depend on on-line road dataset which are not updated and do not cover all of the study area, especially the rural areas that are characterised with haphazard roads and foot paths. Considering the limited time and cost available for this research, an alternative approach is considered that will not only reduce the resources spent at cleaning up and updating the available road data, but can be adopted and sustained. As a simplified method of acquiring the network data, Google Earth image is added as a basemap to the study area in the GIS where the road outlines are traced and converted to vector features. These features are merged with the existing road dataset. The dataset is then integrated in the ABM which does not require the level of refinement and details expected for the travel time estimates in GIS.

- **Polygon/Polyline features** Area features can be obtained by joining point features to form a closed polygon. They are formed by joining a sequence of points by starting and closing on the starting point. The polygon features required are:
 - i **Administrative and Geographical Boundaries:**
 - ii **Water body:** The water body dataset comprises of swamp, rivers, lagoon and the Atlantic Ocean. Same steps for acquiring the road network data are also required for polygon features.
 - iii **Committed spaces:** These polygon features that are restricted to certain uses and have no population within their boundaries. These features include airport, military cantonments and stadiums.

2. Raster data

The raster data for this research is:

i **Demographic map:**

Demographic data can be in raster format as population density maps representing population per unit area. It is also possible to have demographic data in vector format obtainable as point locations representing population centroid or census blocks. The free population density map - Gridded Population of the World version four (GPWv4), available as raster images provided by the Socioeconomic Data and Applications Center (SEDAC) (Center for International Earth Science Information Network - CIESIN - Columbia University, 2016) is proposed for this work. This service gives the estimates of the world population density.

The models created for this work are based on vector data. Therefore, dataset that are obtained in raster format require to be converted to vector data, which is done using the conversion tools available in GIS software.

3.2.3 Database Creation and Data Processing

Spatial data has to be linked to its attribute data before further analysis and processing is done. A database is created to link the data to its attributes such as names, X, Y coordinates, area, length, and addresses.

Data processing is done to project and transform all data to the same coordinate system. This process is known as georeferencing. To project data to a coordinate reference system, the projection tool in the GIS is used. The required transformation and coordinate system is specified, and a projected or transformed data will be obtained. Data processing also entails extracting data to the extent that is required, converting raster data to vector data, and creating different map layers.

The different datasets form different spatial layers such as road network, population density, schools and religious facilities. The catchment layer is the polygon layer created from the results of the travel-time model defining the service area of each HCF. The feature representing the area with no healthcare service coverage is another polygon layer that is created using the geoprocessing tool to clip the administrative boundary feature to the catchment feature.

3.3 Model Preparation

The NetLogo ABM (Wilensky, 1999) software proposed for this research offers a disaggregated analysis that healthcare providers and policy makers at the lower level of governance can utilize to improve healthcare services at community level. The simple user-interface makes it possible for non-technical users to operate, with tabs provided for model codes and documentation.

The different agents and their various characteristics that will be required for the proposed techniques are described in this section.

3.3.1 Agents and Environments Representation

The environment in the NetLogo is the space where agents move, operate and interact with each other or with the environment itself. The environment can be geographic space where agents operate in a real-world environment. Hence the agents have georeferenced locations (Crooks and Heppenstall, 2012). Agents in NetLogo do not interact directly with the GIS data, rather, the agents are created from the spatial data before they can carry out instructions and interact.

The NetLogo world consists of four agents:

- *Patches*. The NetLogo environment is made up of gridded space or world. Each square in the world is an agent called patch and the agent's location in the world

is defined by the NetLogo 2D coordinate system, which can be converted to geographic coordinates.

- *Turtles*. These are agents that can move around or stand on patches.
- *Links*. The links are agents that connect turtles.
- *Observer*. The observer is the agent that oversees what goes on in the world. While other turtles have locations, the observer does not have.

Similar to resolution of pixels or map grid size in GIS, the positional accuracy of the model is a function of the patch resolution. A set of agents is called an agentset. Figure 3.3 shows an illustration of the NetLogo environment, the two different turtles - turtle 1 and turtle 2 on the patches are connected by the link, while the observer oversees what is happening in the world.

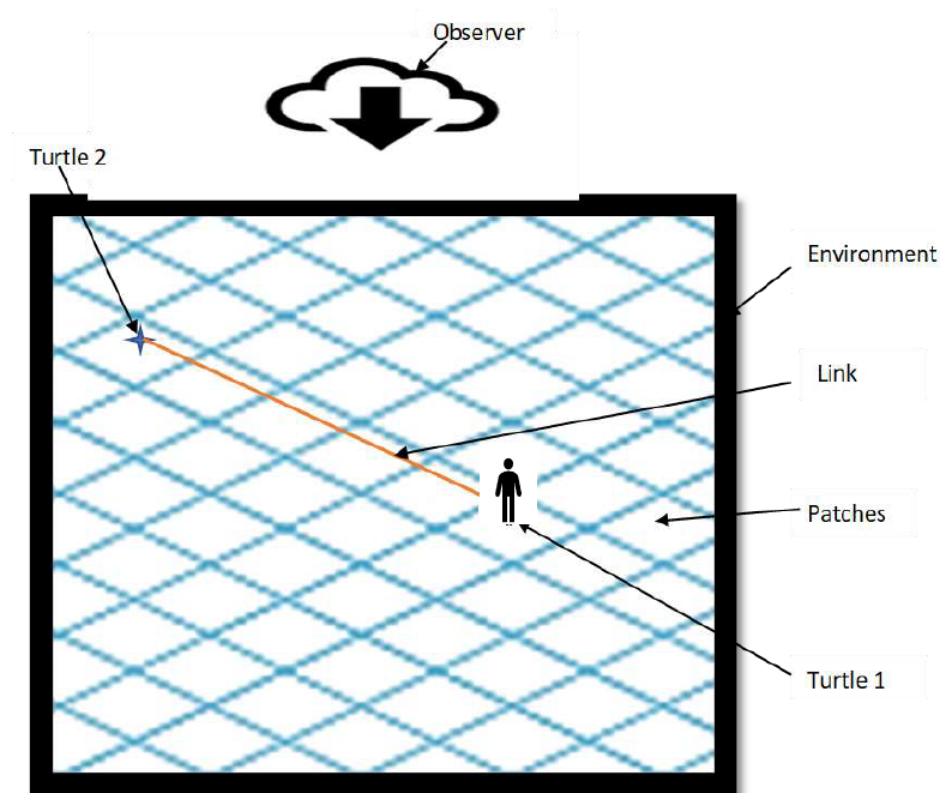


Figure 3.3: NetLogo environment

Each agent in an ABM has its own attributes or variables, character and behaviour. The foundations of an agent's behaviour and relationship with themselves and the environment are the rules they possess that dictate their actions. A given rule can be applied to a single agent or an agentset based on whether a condition is satisfied or not to perform certain actions. The actions may be performed simultaneously or individually after a time-step or after an agent has performed an action. Examples of rules that influence agent's actions are given below:

Conditional rule: The decision for an agent to retain its location may depend on if there is no other agent on that same location. That is there is no other competing interest. This is applicable in achieving the constraint of not having more than one HCF on a location.

Relationship rule: Agent's relationship may be with themselves or the environment and may be reactive or goal-directed.

Reactive relationship - This is activated by external stimulus such as the action of another agent. An agent may also be asked to move in a specified direction by the observer. This is applicable when a patient is travelling from HCF location towards the direction of its dwelling.

Goal-directed relationship - An agent may be seeking to accomplish a certain goal such as creating links with another agents to assess the maximum or minimum distance between itself and an agent. This is a very useful attribute in the optimization problem where demands create links with the closest HCF in the neighbourhood in order to fulfil the objective of reducing the maximum distance a patient has to travel to receive treatment.

The application of the varying agents' characteristics to spatial properties are discussed further in Section 3.6.

For the ABM, the administrative boundary map is used to represent the environment in a form that references the model to a location in the geographical space. The models comprise different layers which separately represent the road network as well as other features such as water body, demography, HCF and other amenities such as school, market and religious facilities.

3.3.2 Agents and Spatial Features Required for a Catchment Definition

For catchment or service area definition of each HCF, the travel time of patient is modelled where patients move along the road network from the HCF location to their dwellings. For patients to get to a healthcare facility, they are expected to navigate along a road or transport network. The environment must include the healthcare facility, buildings where patient agents dwell, and a road network through which the agents navigate. The agents involved in this model are listed below.

1. **Healthcare facility layer:** Each HCF feature is represented by a HCF agent. The agent adopts the attributes of the feature it represents such as name and geographic location.
2. **Patient layer:** Generating a simulation of every patient that attends a HCF is a huge task. Instead, a few numbers of people can be simulated to represent patients that have to travel to receive treatment at the healthcare facility.
3. **Residential houses layer:** The address of a patient who visits a HCF can be

used to locate where such patient comes from to receive treatment. This will require patient-level data and consent would need to be sought or anonymity of the patient is established. Considerable data is involved and individual house location will need to be represented to determine the time the patient travels to the HCF. Another way is to aggregate the patients' locations to population centroid or census blocks as obtainable in structured city planning. Unfortunately, the building patterns and addressing system in Nigeria is not well structured, neither is population centroid or census block defined. In such context, houses can be simulated to represent patient houses. This is especially helpful where there is no proper documentation to guarantee adequate patient level data. Simulating virtual houses also conceals the identity of patients. This research therefore simulates buildings to represent patient houses. HCF catchment is then defined with proximity of patient from HCF to their dwellings. The simulated houses are randomly placed in the environment to form a spatial layer of house locations as point features in the model.

4. **Road network layer:** To create a road network layer, two types of agents are required – turtle that represent the nodes, and lines that represent the edges connecting the nodes. A road network along which patients are to travel needs to be connected to indicate a change in direction if there is one. In ideal situations, roads should be classified according to the speed-limit, and such classification should be an attribute in the GIS database. The assumption in this model is that such detailed information is not readily available and so the travelling speed is constant and can be specified by the user on the user interface.

3.3.3 Agents and Spatial Features Required for Facility Location Model

1. **Healthcare facility layer:** The existing HCF agents are created as defined in the catchment definition model. Additional HCFs are virtual HCF agents simulated with the ability to change location. In this agent's simulation, occupancy on a location depends on demographic value, locational suitability and competing interest - measured by presence of another HCF agent on the same location or within a distance-limit.
2. **Catchment layer:** They are patches created from the intersection of the catchment feature with patches in the environment. The catchment of each facility is represented by a patch-set that is spatially connected. The cluster of patches have the same attributes.
3. **No-coverage layer:** The feature representing the area outside existing healthcare service coverage is termed *no-coverage* in this research. The no-coverage layer is another set of patches similar to the catchment layer created from the intersection of the no-coverage feature with patches.
4. **Demography layer:** The demographic raster dataset is a continuous dataset

from where a shapefile is created and the population density values are assigned to the vector dataset based on the area of coverage. The dataset is imported into the ABM environment where a reverse conversion is created by assigning the population density value to the patches intersected by the feature. A model population density is then calculated with the size of the patch. Instead of the number of people per square km in the real world, it becomes the number of people per patch (real world size). Seeing that creating a simulation of the number of persons occupying a patch will be quite clumsy, the patch is the agent having the number of people as its attribute.

5. **Water body layer:** They are patches created from the intersection of the water body feature with patches in the environment. The water body is represented by a spatially connected patch-set representing water feature. The cluster of patches have the same attributes and have no population.
6. **Committed layer:** Other than the water body, this is another layer that contains features that are not expected to have residents within them. These include the airport, stadium and army cantonment.
7. **Institutional facilities layer – school, markets, and religious facilities:** These are static agents simulated from the GIS layer. HCFs locations are expected to be in proximity to any of these amenities. Although there's no proximity value defined by WHO, however, to demonstrate how this is integrated into the proposed model for this research, the condition is assumed fulfilled if any one of such amenities is within the catchment of a HCF.
8. **Ancillary agents:** For the objective function of minimizing the distance between the farthest population that is not covered with MCLP and and their closest HCF, a turtle agent is simulated on each patch that is not within any facility catchment to represent population on the patch. Another link agent is simulated to connect the population representative to the closest HCF. The HCF could be a newly created HCF from optimization, or an existing HCF from the spatial data.

3.4 Travel-Time Catchment Model

The catchment area of a healthcare facility defines its coverage of service and this can be defined by having knowledge of the geographic extent or boundary of a threshold distance or time it takes patients to travel to the healthcare facility. When a distance measure or buffering technique is used, it is taken that the journey is made in a straight line. Such measure is usually adopted in analysis that are constrained by accurate data on the road network, especially in the low and middle income countries. However, the reality is that there is a connection of lines that indicates a change in bearing through which navigation is done by the patients. In the case of providing healthcare for an entire community, a catchment should be defined by a network travel-time measure

of movement of patients. Such measure is employed in this research for defining the catchment of existing HCFs. However an optional straight-line measure is incorporated into the model to reveal the difference between the two measures and provide decision makers the flexibility of comparison, and choosing a method that suits them.

The purpose of travel-time measurement in this research is to determine:

- an estimated value of area that is uncovered with healthcare services
- the average travel-distance that people in a region travel to HCFs; this is the desired service distance
- use the estimated uncovered area and average travel-distance as input to the location-allocation model

There are different open source and proprietary application software and models available for determining service areas. While some require the road network to be pre-created by the user, some depend on on-line maps such as Open Street Map (OSM) that depend on the internet. Internet-based applications are not appropriate and cannot be sustained in countries such as Nigeria where internet access and coverage is quite low, with costly and unreliable network that is basically provided via mobile network (Idiegbeyan-Ose et al., 2016). In addition, as a developing country, there are continuous construction and maintenance phases on road infrastructure by the government and by individuals who at times have to construct paths to their houses (Oshodi, 2010). The on-line tools do not capture such developments and may therefore not provide reliable travel scenarios.

The NetLogo ABM software is an open source software that does not need the internet to function and can be run on external storage devices, with the required spatial data stored in the NetLogo directory. NetLogo also has the capability of storing the output data as point, line or polygon features in GIS shapefiles and not just a list of coordinates. This serves as an advantage over existing application that require expert knowledge to process the output data. Unlike the available on-line and sophisticated tools that automatically create the service areas, the user of the NetLogo model can visualize the process to better understand the phenomenon that is modelled.

For the proposed model, travel-time and distance is measured from the HCF to patients dwelling. This model creates virtual patients at the HCFs to travel to their houses of residence, so that the catchment boundary can be defined by the agent's location on the patch that represents the end of the travel-time threshold. A user-defined travel-time limit and speed limit option is included in the model interface. A patient would need to choose his destination from one of the virtual houses simulated randomly in the environment. The patient then moves towards the destination by navigating through the road network using the A-star shortest path algorithm. The A-star algorithm is a heuristic optimal path finding problem that chooses the next node from a starting node, based on the cost from that starting node plus a proximity or distance

estimate to the destination, using a function:

$$f(n) = g(n) + h(n) \quad (3.1)$$

where:

$f(n)$ = total cost or distance of the node

$g(n)$ = the exact cost or distance between the starting node to any node n

$h(n)$ = the heuristic: estimated cost or distance from node n to the end node

The heuristic h is estimated with Euclidean distance given in equation (3.2). The A-star finds the node n that has the lowest total cost $f(n)$. The A-star algorithm is given in Algorithm 3.1. Two lists are created at initialization: OPEN list that consists of nodes that have not been explored; CLOSED list that consists of nodes that have been explored (Zhou, 2016).

Algorithm 3.1 A-star shortest path search Algorithm

```

1: Initialize the OPEN list with  $f(startNode) = h(startNode)$ 
2: create an empty CLOSED list
3: while the OPEN list is not empty destination is not reached do
4:   find the node with the least  $f$  from the OPEN list and set it as  $f(currentNode)$ 
5:   if  $currentNode$  is  $endNode$  then
6:     the solution is found; break
7:   else
8:     add  $currentNode$  to the CLOSED list and obtain  $g(successorNode)$  for all neighbours of  $currentNode$ 
9:     for each  $successorNode$  of  $currentNode$  do
10:      set  $g(successorNode) = g(currentNode)$ 
11:      if  $successorNode$  has  $g$  value less than  $currentNode$  and is in the CLOSED list then
12:        replace  $successorNode$  with the new, lower  $g$  value
13:         $currentNode$  is now successor's predecessor
14:      else if  $currentNode$  has  $g$  value less than  $successorNode$  and is in the OPEN list then
15:        replace  $successorNode$  with the new, lower  $g$  value
16:        change successor's predecessor to the  $currentNode$ 
17:      else if the  $successorNode$  is not in the two lists then
18:        add the  $successorNode$  to the OPEN list
19:      end if
20:    end for
21:  end if
22: end while

```

The estimated distance to destination is calculated using the Euclidean distance:

$$L = [(X_i - X_j)^2 + (Y_i - Y_j)^2]^{\frac{1}{2}} \quad (3.2)$$

For Euclidean distance:

L is the hypotenuse of a right-angled triangle.

X_i, Y_i is the coordinate of the point i , indicating the beginning of a line $i - j$ whose distance is to be calculated.

X_j, Y_j is the coordinate of the point j , indicating the end of the line $i - j$.

As the agent moves along the selected travel path, its positional coordinates are updated. The total distance it has travelled is the total distance on each link computed by:

$$D_n = \sum_{i=1}^n L \quad (3.3)$$

$$D_n = D_{n-1} + [(X_n - X_{(n-1)})^2 + (Y_n - Y_{(n-1)})^2]^{\frac{1}{2}} \quad (3.4)$$

and travel-time is:

$$T_n = T_{n-1} + \left(\frac{speed}{D_n}\right) \quad (3.5)$$

$$Speed = \frac{distance}{time} \quad (3.6)$$

where:

n = the number of times the agent has moved in a time step

D_n = the horizontal distance the patient has moved from the start of the route section to the n th location.

T_n = the travel time of the patient from the start of the route section to the n th location.

At initial location (x_0, y_0) of the agent (Figure 3.4), both the distance D and time T have zero values. When the agent moves to location 1 (x_1, y_1) , D_1 is equal to L_1 as obtained from equation (3.4), and T_1 is subsequently obtained from equation (3.5). As

the agent moves to the next node, the distance value, L it has covered is added to the previous value of D and the travel time T is successively calculated until it reaches its destination or stops at the time threshold at location n .

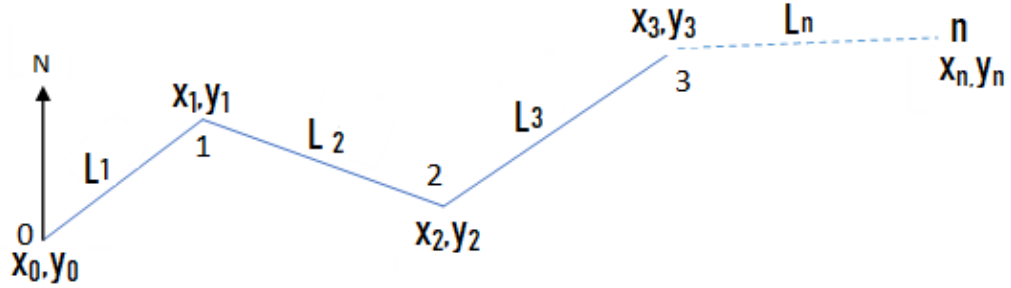


Figure 3.4: Travel distance covered

This calculation is made possible by the gridded ABM environment. The distance is scaled to real world distance using a scale factor that is determined from a known distance in the real world, drawn in GIS and projected onto the ABM. The travel-time model flow chart is shown in Figure 3.5.

For network measures, each patient can move towards any randomly chosen destination via the shortest path chosen using the A-star search algorithm (Zhou, 2016). As a patient travels towards a destination, the travel time and distance are calculated simultaneously. If a patient gets to its destination before the threshold time, it stops any further movement. If not, it continues moving until the time-limit is reached. The patient's location is then established and saved as a point vector feature representing the catchment limit of the facility which can be exported to GIS, together with their attributes (travel time, travel distance, patient number, model (x, y) coordinates and directions) for visualisation and further analysis. The travel time and distances are shown in the model interface. The user has the flexibility of viewing the locations reached on or before the time-limit. The straight-line measure does not require a chosen destination, rather the agents can move in any straight direction and stop only when the time-limit is reached.

Agents whose locations form the boundary of HCF catchment are used for further analysis. The next section describes steps involved to determine the catchment area estimate and create the features of the area outside existing healthcare service coverage which is termed *no-coverage* feature.

3.4.1 Catchment Area Estimate

The output point features shapefiles from the ABM simulation are viewed in the GIS where a minimum bounding geometric circle is drawn round them to define the catch-

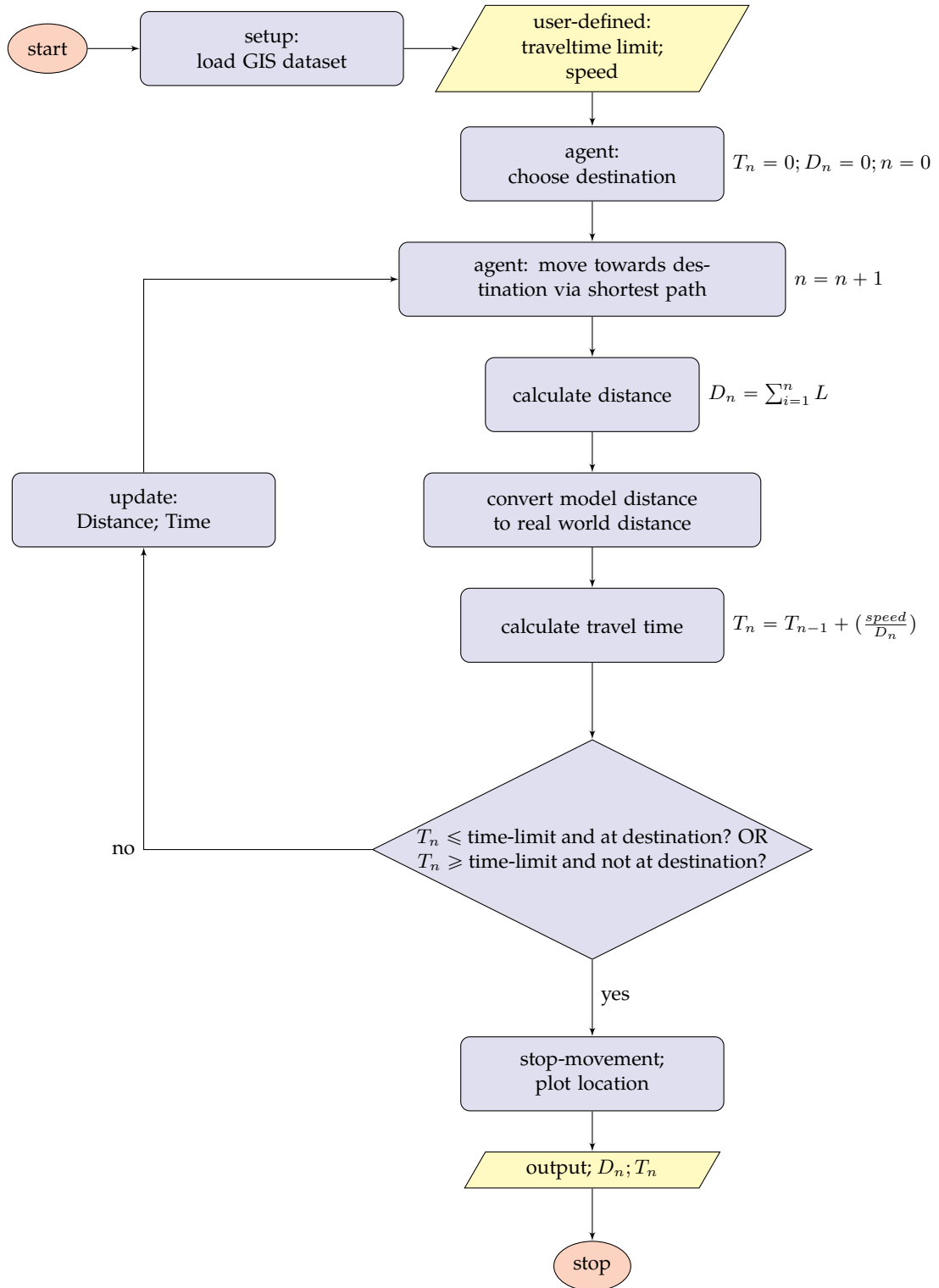


Figure 3.5: Travel-Time Flow Chart

ment. The minimum bounding circle (MBC) is the smallest circle that contains all the set of points in the plane that indicates the specified travel time limit of agents. This is especially useful in the HCF location problem to minimize the farthest distance that patients will travel to receive care. The centre of the circle is not the HCF location but is derived from systematically moving its centre towards the centre of the bisector of two chosen points enclosed by the circle, while reducing the radius. The MBC centre is thus the point that is at the minimum distance from all the points that are on or within the circle (Banik et al., 2014).

The choice of the MBC approximations is based upon the argument that circles are not sensitive to orientation, therefore a feature's MBC is unique and is not affected by its topological transformations (Brinkhoff et al., 1993; Safar and Shahabi, 1999). In addition, the MBC has a low storage requirement compared to other conservative convex approximations such as convex hull and minimum bounding rectangle (Brinkhoff et al., 1993). Using this feature makes it possible to derive the radius of the circle which minimizes the maximum travel distance of population to an HCF from their residence and is an improvement on the circular buffer method that assumes that every demand point on the circumference of a circle is centred at the HCF location and equidistant from such location.

Figure 3.6 shows an illustration of the point features representing the end of a number of agents' journeys at the time limit, and the MBC encompassing the points. It can be seen from the figure that the straight-line circle has its centre at the HCF location while the network bounding circle centre reveals different distance from the HCF location. This is where this work differs from the traditional GIS circular buffering and the sophisticated GIS network analyst tool. The buffer tool in GIS software provides the feature for drawing a circular polygon centred at the facility at a specified radius that defines the maximum distance a patient will need to travel for a healthcare service. Although generating a network service area with an irregular polygon boundary using the capability of the network analyst in GIS is regarded to be a more accurate method, it requires building a specific Network Dataset from the road map to form accurately defined topologically connected road network dataset that may consist of millions of lines and nodes. If the road map is not well digitized, an inaccurate service area is produced. Such techniques may be difficult to implement in many developing nations because of its technical and financial requirements.

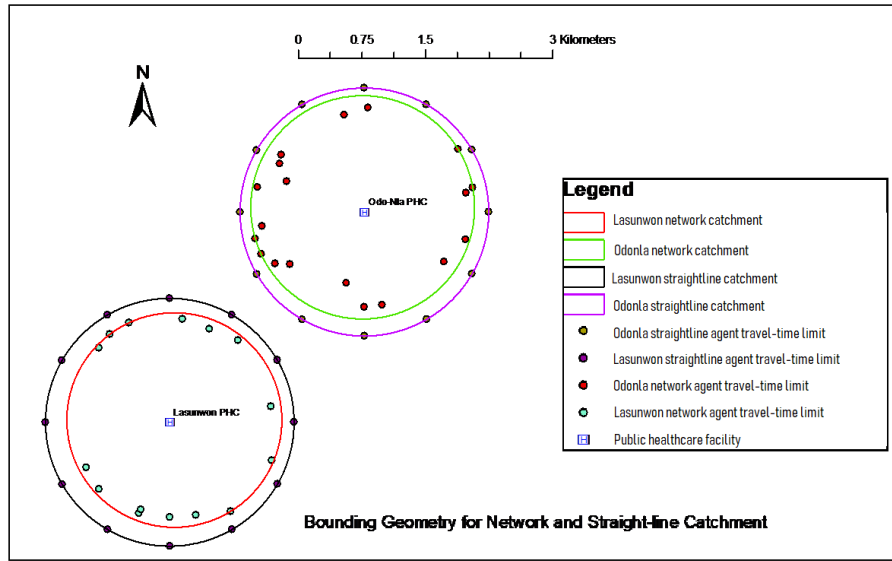


Figure 3.6: Minimum Bounding Circle and Circular Buffer

The MBC polygon are further merged and dissolved to form the healthcare service coverage feature whose area is obtained from the GIS geometry calculation tool.

3.4.2 No-Coverage Layer Creation

The no-coverage feature is extracted from the administrative boundary feature through the clip geoprocessing technique. The coverage feature is clipped with the administrative boundary so that the no-coverage feature can be extracted. The geometric area values are used for clipping the features and a resulting feature is the no-coverage feature that is saved as a shapefile.

That is:

$$NCA = AA - CA \quad (3.7)$$

where:

NCA = no-coverage area/feature

AA = administrative area/feature

CA = coverage area/feature

Figure 3.7 provides an illustration of no-coverage feature extraction from the administrative and the coverage features.

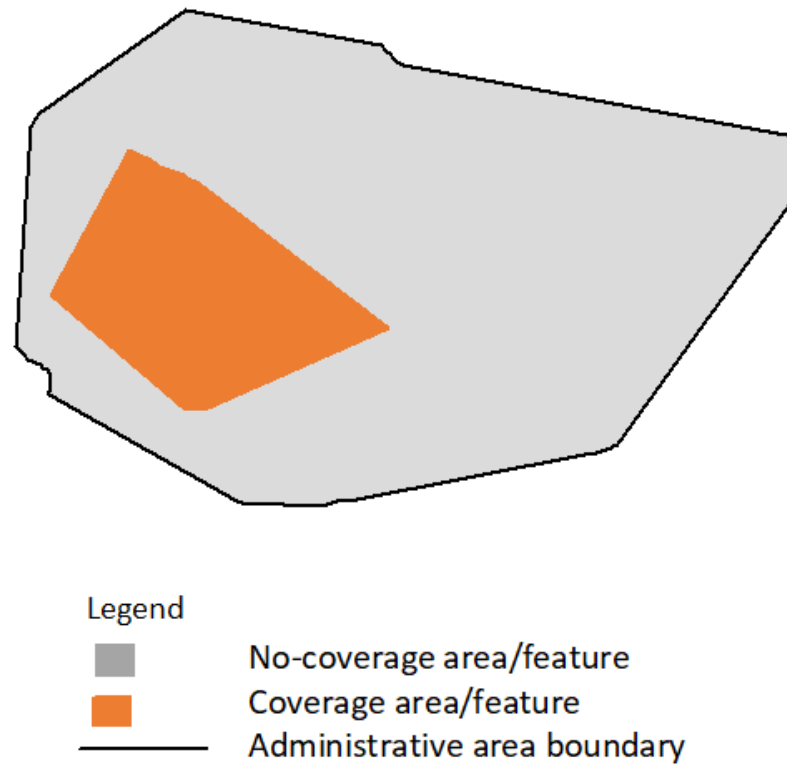


Figure 3.7: No-coverage feature extraction from administrative and coverage features

3.5 Location-Allocation Model

The traditional Maximal Covering Location Problem (MCLP) assumes that each facility has an infinite carrying capacity. This assumption is not the case in reality and has led to different versions of the MCLP which is generally referred to as capacitated facility location problem. Other than the primary objective of coverage of service within a maximum time or distance threshold, capacity constraint is included in the solution model of the facility location problem of this thesis. This is to account for capacity of the HCF which may be assessed based on the number of demands that the HCF can accommodate. When demands are assigned to a facility as its capacity, the location problem has been extended to include allocation problem. The two problems can be solved together as a location-allocation problem to include two factors. These factors are the location of the facilities and the allocation of demand to the facilities, and they are regarded as having a direct impact on a system's operating cost (Haghani, 1996). Inappropriate locations of public HCFs may increase resource allocation costs such as health workers and healthcare infrastructure.

The following factors are considered for the location-allocation problem:

HCF capacity

Capacitated MCLP models have been built on the traditional MCLP of Church and ReVelle (1974) which does not account for service capacity. For example, Haghani (1996) developed two models that include capacity constraints for a maximum covering location problem. The models assign uncovered demands to facilities and also consider minimum and maximum allowable utilization levels. In their models, the first solution uses a greedy adding heuristic, while the second is based on Lagrangian relaxation. A great number of other works have included budget constraints in the solution to facility location problems. For example, Salari (2014) considered the capacity and the budget allocation for constructing a facility and the service level of the facility to the demands being served, using both integer linear programming and local search algorithm. Zhang et al. (2016) incorporated the cost of building a healthcare facility using a genetic-algorithm based multi-objective optimization approach.

The concept of capacity and budget are introduced to the proposed facility location model using the global facility density benchmark. One of the tracer indicators for measuring healthcare service availability and readiness by WHO is the density of HCFs in a region. Other indicators include health worker and bed density. The global benchmark for facility density is given as one facility per 5,000 population or two facilities per 10,000 population (WHO, 2013). MCLP is solved with constraint on the maximum number of population to serve:

- A lower band and upper band capacity constraint is considered for each candidate HCF
- If the demand within the service area of a HCF exceeds the upper bound, additional HCF is placed within the service area. Each HCF is distinguished from the other based on the demand capacity.
- If the demand within the service area of a facility is below the lower bound, a HCF of lower establishment cost is proposed as a cost-cutting solution. Such facility can have less number of human and infrastructure resources. Therefore if a service area has more than one HCF, the facilities may not require the same establishment cost. For example optimized locations covering less than the maximum required capacity can have a community-based health workers as suggested by WHO (2017).

Budget flexibility

In addition to reducing establishment cost based on HCF carrying capacity, this thesis introduces a flexible budget constraint that considers the budget allocation. Rather than have a maximal coverage of one HCF to 5,000 population, the decision maker can relax the capacity of the HCF. The global benchmark is regarded as 100% coverage, but a capacity relaxation is achieved by reducing the expected coverage percentage.

For example, 50% coverage indicates a facility density of 1 HCF to 10,000 population. The HCF type is determined by the size of the population within the catchment of the candidate HCF.

Site suitability

Different solutions to the location-allocation problems have been suggested by introducing suitability analysis in the model. This is usually done with spatial analysis using GIS to evaluate the suitability of sites based on proximity to some defined environmental, cultural, economic or social factors. This work introduces suitability constraints in the proposed model. Unsuitable locations include sites outside the administrative boundary of the region of focus or inside water bodies. Suitability map layers are included to consider if a proposed location is suitable or not for placing a HCF.

Service distance

Distance or time relaxation is considered in defining the catchment of facilities. Although the uncovered population is determined based on network travel-time to existing HCFs, the new HCFs determine their catchment with average Euclidean distance which is derived from the catchments of existing HCFs within the region.

Another relaxation to the facility location problem is the service distance for the solution of the p -centre Location Problems (PCLP) so that the maximum distance of every demand to its closest facility is minimized by assigning the populations that are not covered by the MCLP to the new or existing HCFs that are closest to them. The covering problem therefore has two major objectives and two sub-models are included in the solution model.

3.5.1 Solution Models

As discussed in Chapter 2, the facility location problem for this thesis is formulated as Capacitated Mandatory Closeness Maximum Location Problem (CMC-MCLP). Many researchers have formulated and solved the location-allocation problems as mathematical optimization problems. Problems with a single objective and few variables can be effectively solved by a linear programming technique to provide an optimal solution of the problem. Such solutions may not be feasible in a two-stage multi-criteria and complex real-life situations proposed in this research. Therefore, a meta-heuristic optimization technique based on Spatial Agent-Based modelling is used to obtain feasible solutions to serve as a decision support system for health service coverage. Most of the models in the literature have regarded and treated capacitated MCLP as linear and deterministic, but in reality, the problem is complex and has probabilistic and stochastic characteristics that require alternative solutions (Farahani et al., 2014). Agents' characteristics in ABM can be utilised to simplify the problem.

Previous models are based on locating predetermined or restricted number of facilities. Some models have candidate facility locations from which the required number of facilities are selected and demands assigned to them, as is usually the case with GIS standalone location-allocation models. These do not proffer the flexibility of exploring different location possibilities. It is possible to have more than the required number of facilities at a lower establishment cost if the demand within the catchment of the facilities is low compared to the capacity benchmark. Fewer facilities may also provide the required level of coverage. The number of facilities to be located is governed by the spatial structure of the environment and distribution of population within the geographic space. The decisions regarding maximal covering of public HCFs should be based on answering the questions of how much services should be located to address the needs of the entire population, and where the services should be located following an efficient distribution pattern. The mandatory closeness decision is focused on which of the varying feasible solutions should be selected or not to be selected.

The MCLP solution model for this research has a large number of potential discrete locations which can be in the demand locations or anywhere else in the gridded environment, except on water or other land use/land cover feature that HCFs are not expected to be located, while the mandatory closeness that is based on p -Centre Location Problem (PCLP) has set of fixed number of possible HCFs, based on existing and newly selected HCFs.

In the capacitated MCLP model, an approximate number of facilities p required to cover the population is the value of:

$$p = RP/BMP \quad (3.8)$$

where:

RP = total population of the region

BMP = global benchmark population value

The mathematical formulations and meta-heuristic algorithms are provided in the following sections.

3.5.2 Formulation of Capacitated Location-Allocation Model

The mathematical formulations of the location-allocation model with the objective function to maximize the total population coverage within the desired covering distance D can be formulated using the following notations:

I = set of demand locations or nodes;

J = set of potential facility sites j ;

i = indices of demand locations or nodes;

j = indices of potential facility sites;

K = set of potential facility sites to be established at lower cost, $K \subseteq J$;

k = indices of potential facility sites to be established at lower cost;

g = indices of potential facility sites not at lower cost;

h_i = total sum of population located within the service area of facility;

p = total number of facilities required;

d_{ij} = the maximum distance or time between demand i and facility j ;

D_{ij} = the maximum acceptable travel distance or time between demand i and facility j that a demand point is considered covered;

$d_{ij} \leq D_{ij}$;

B_j = suitability of candidate site based on land cover or land use;

N_i = set of all candidate facilities which can cover demand point $\forall i \in I$, $N_i = \{j \in J : d_{ij} \leq D_{ij}\}$;

C_j = capacity of facility j in terms of facility to population density;

q = total number of lower-cost facilities to be established

X_j = binary decision variable indicating if the facility is located at point j , 0 otherwise;

Z_i = decision variable is 1 if node i is covered, 0 otherwise;

y_j = location decision variable is 1 if facility is established at point j , 0 otherwise;

The objective function of the model is:

$$\text{Maximize } \sum_i h_i Z_i \quad (3.9)$$

Subject to:

$$Z_i \leq \sum_{j \in N_i} X_j \quad \forall i \in I, \quad (3.10)$$

$$\sum_{j \in J} X_j \geq p \quad (3.11)$$

$$\sum_{j \in J} X_j \leq p \quad (3.12)$$

$$\sum h_{ij}X_j \leq C_j y_j \quad \forall j \in P, \quad (3.13)$$

$$\sum_{j \in N_i} X_j \leq 1 \quad \forall i \in I, \quad (3.14)$$

$$\sum_{k \in J} X_k = q, \quad (3.15)$$

$$K^c = \{g : g \in J, g \notin K\} \quad (3.16)$$

$$X_j \in \{0, 1\} \quad \forall j \in J \quad (3.17)$$

$$Z_i \in \{0, 1\} \quad \forall i \in I \quad (3.18)$$

$$y_j \in \{0, 1\} \quad \forall j \in J \quad (3.19)$$

The objective (3.9) maximizes the total covered demands. Constraints (3.10) are the constraints that describe the relationship between the coverage and location variables, and state that demand node is covered if it is covered by at least one facility within a distance less than or equal to D . Constraint (3.11) states that the number of facilities to be established is more than or equal to the approximate required p for the region based on facility density. Constraint (3.12) states that the number of facilities to be established is less than or equal to the approximate required p for the region based on facility density. Constraints (3.13) require that total demand allocated to facility j is not more than the capacity. Constraints (3.14) state that no location has more than one facility. Constraints (3.15) and constraints (3.16) limit the cost of establishing facilities. Constraints (3.17), (3.18) and (3.19) are integrality constraints.

3.5.3 Formulation of Mandatory Coverage Model

In the second sub-model, the isolated or uncovered populations in the first covering sub-model is assigned to the closest facility to their location. In this case however, the capacity of those facilities are no longer considered. Unlike the maximal covering

model that endogenously determines the locations of facilities p , the locations of facilities for mandatory coverage model are fixed. The potential facilities are the total number of existing facilities and the newly located facilities. It is therefore likely that an uncovered population may be closer to an existing facility than a new facility. It is assumed that patients attending public HCF will usually attend the closest facilities to them. This implies that the maximum travel distance S between an uncovered demand and its closest facility can be less than the maximum acceptable covering distance D , which is the radius of the new facility catchment, depending on the spatial distribution of the demand and the supply.

The mandatory coverage model is formulated using the following notation:

I = set of uncovered demand locations or nodes i ,

J = set of existing and newly established facility sites j ,

p = total number of existing and newly established facilities to be allocated,

d_{ij} = the maximum distance between demand i and facility j

D_{ij} = the maximum acceptable coverage distance between demand i and facility j from capacitated MCLP

S_{ij} = the maximum travel distance between uncovered demand i and facility j
 $D_{ij} \leq S_{ij}$

N_i = set of all candidate facilities which can cover uncovered demand point $\forall i \in I$, $N_i = \{j \in J : d_{ij} \leq D_{ij}\}$

X_j = binary decision variable indicating if the facility is located at point j or not,

Z_i = decision variable is 1 if node i is covered; 0 otherwise

The objective function of the model is given by:

$$\text{Minimize } W \quad (3.20)$$

Subject to:

$$\sum_{j \in N_i} y_{ij} = 1 \quad \forall i \in I \quad (3.21)$$

$$\sum_{j \in J} x_j \leq P \quad (3.22)$$

$$y_{ij} \leq x_j \quad \forall i \in I, \forall j \in N_i \quad (3.23)$$

$$W \geq \sum_{j \in N_i} d_{ij} y_{ij} \quad \forall i \in I \quad (3.24)$$

$$x_j \in \{0, 1\} \quad \forall j \in J \quad (3.25)$$

$$y_{ij} \in \{0, 1\} \quad \forall i \in I, \forall j \in N_i \quad (3.26)$$

$$W \geq 0 \quad (3.27)$$

The objective (3.20) minimizes the maximum distance or time between an uncovered demand point and the nearest facility allocated to it. Constraints (3.21) ensure that each demand point is covered by only one facility. Constraints (3.22) specify that the total number of facilities that uncovered demand is allocated to is within the total number of fixed facilities. Constraints (3.23) indicate that demand locations are only covered by established facilities. Constraints (3.24) enforce that W is the maximum distance from demand location to facility (W is the auxiliary variable used to determine the maximum distance). Constraints (3.25) (3.26) and (3.27) are domain constraints.

While all population could not be covered within the maximum covering distance D_{ij} in the MCLP, some existing facilities may be within distance D_{ij} of the uncovered population. Since an uncovered demand can be allocated to a newly located facility or an existing facility that is closest to its location i , it therefore implies that the maximum distance S_{ij} between the facilities N_i that can cover the uncovered demand can be less or greater than D_{ij} . Additional constraints (3.28) and (3.29) are set to incorporate this:

Subject to:

$$S_{ij} \geq D_{ij} \quad (3.28)$$

$$S_{ij} \leq D_{ij} \quad (3.29)$$

Constraints (3.28) states that the maximum distance from an uncovered demand to the closest facility is greater than or equal to the maximum acceptable coverage distance. Constraints (3.29) states that the maximum distance from an uncovered demand to the closest facility is less than or equal to the maximum acceptable coverage distance.

The difficulty in achieving a solution to the covering problem may be addressed by solving a relaxed form of the problem through the application of fewer constraints to achieve a rapid solution. In this case, distance measure is based on Euclidean metric.

3.6 Spatial and Agent Characteristics

The topological relationships or properties of spatial objects and agents required for the objective functions and constraints of the spatial optimization are distance, connectivity or contiguity, adjacency, containment, intersection and pattern.

Distance: For this research, the distance that a patient would need to travel from a residence location to the closest HCF location is measured in Euclidean distances. A facility agent is capable of identifying locations within a distance radius from itself. The two ends of the radius have their coordinate pairs that define their positions.

Connectivity: Agent representing a demand node is connected with its closest facility via a link agent. The link is another agent that represents the distance. Each link has a unique identification and can identify the facility and demand to which it is connected. It can also determine its distance in the ABM unit which is also convertible to real world distance.

Adjacency: When two features are next to each other or share an edge, they are said to be adjacent. For example when the polygon features of two HCF catchments share a boundary.

Containment: A containment relationship is established when an object is completely within another object. The containment condition is applied using the NetLogo primitive of 'gis:contained-by?'. For example,

- A HCF is contained by the catchment feature;
- All population that are totally enclosed in the catchment of a facility are within the service area of that facility. The population value of all patches within the boundary of a catchment represent the population within that service area;
- A facility agent is not permitted to be contained by a water feature. Agents will assess its location and if it discovers that it is contained by water feature, it will eliminate itself from the system;
- In satisfying the condition of proximity to institutional facility such as a school and a religious facility, a HCF will satisfy that condition if an institutional facility agent is enclosed by its catchment feature. In this case, if an HCF agent does not

satisfy the condition of having amenities within its catchment, it does not exit the system. Instead it changes its status to indicate that there is a population to serve, but no institutional facility within reach of that population. This condition also creates an additional awareness that the community is isolated from healthcare services and other amenities.

Intersection: Intersection is established where two objects exist on the same location at the same time. For two polygon features, the area of intersection is the section of overlap of the two features. Catchment overlap can be assessed with the intersection of patches of neighbouring service areas. A level of boundary overlap is permitted to avoid gap between catchments such that some population will be out of coverage. A facility can only be located within a certain distance outside the boundary of another facility service area to avoid conflict of interest or responsibility. The patches in the ABM environment that are enclosed within the distance radius can independently distinguish themselves from other patches with unique properties. An exception for this condition is if the population within the catchment of a HCF is high and another facility is required. In this case the new HCF's catchment can overlap the existing one since they are meant to serve the set of population within the travel-time threshold.

The different features representing the existing coverage of healthcare is distinguished from the uncovered region by using a NetLogo primitive that gives the patches intersected by the two features different characteristics that differentiates one feature from the other. This assessment ensures that facilities are located in only uncovered areas. This is also applied to the allocation of uncovered population in the MCLP where the locations of coverage are distinguished from the uncovered to ensure that only the uncovered population is connected to the closest facility. One of the conditions that must be satisfied is that only one HCF is located on a position in space.

Pattern: Pattern refers to how spatial objects are distributed or organized in space. Spatial pattern can be random, clustered, or dispersed. In the covering location problem, the spatial configuration of service facilities will depend on the settlement pattern of the demand, and land use/land cover characteristics of the region. An overall coverage of service is maximized by ensuring that HCF agents have neighbouring facilities within a certain radius, and if more than one HCF is required within the catchment of a selected HCF, the selected HCF agent reproduces the additional required number of HCFs. If the number of population to be served by the newly produced HCF agent is fewer than the threshold population, the agent changes its status to reveal that it can only serve a population below the threshold carrying capacity. This is also true for HCF agents that are located to serve fewer populations.

3.7 Location-Allocation Procedure

The procedure for the location-allocation process are two fold.

1. First the spatial datasets are prepared in GIS to be integrated into the ABM. These datasets are:
 - a) the administrative boundary of the region where HCFs are to be placed
 - b) the locations of existing HCF
 - c) the no-coverage feature obtained from the accessibility measure
 - d) the population density map
 - e) point features of schools, markets and religious facilities
 - f) water body dataset such as rivers, lakes, lagoon, ocean and swamp that are forbidden locations for HCFs
 - g) boundary map of committed zones that are also forbidden locations for HCFs which may be airports, stadiums, or military cantonments
2. The second step is to prepare the model. Since NetLogo does not relate with spatial data directly, all the spatial data will be converted to agents - turtles, links and patches. The turtles are point features, links are line features, and patches are area features. The projection and transformation data is uploaded to the ABM so that all analyses will be in the same coordinate system. In the ABM:
 - a) The GIS features will be re-drawn in the model so that agents can be created from them.
 - b) Each agent will adopt the properties of the spatial data it represents.
 - c) The patches that intersect the demographic feature will adopt the population density value of the grid that intersects it. This is achieved using the "gis:apply-coverage" primitive.
 - d) The area property value of the no-coverage polygon feature adopted by the ABM is divided by the total number of patches contained in the no-coverage feature. This gives the area of a square patch. The population per patch is therefore the population density multiplied by the patch area.
 - e) The intersection of different area features with patches are assigned different colours to distinguish them from other area features. This colour becomes the attribute of the patch that is unique to the feature it represents.
 - f) Where any forbidden zone intersects a patch, the population value within that patch is set to zero to indicate that no population resides there.
 - g) Population in a set of patches is the total population of each patch that represents the area of interest.

3.8 Large Neighbourhood Search for Location-Allocation

Several heuristics or meta-heuristics algorithms exist to solve covering problems, such as Tabu Search, Simulation Annealing (SA), Genetic Algorithm (GA), and Greedy

Randomized Adaptive Search Procedure (GRASP) (De Smith et al., 2018). Although such algorithms may or may not proffer globally optimal solutions, they are simple and easy to implement and can be used to quickly obtain a feasible solution for the covering problem. The Large Neighbourhood Search (LNS) meta-heuristic (described in Section 2.7) has been shown to find near optimal and sometimes optimal solutions to optimization problems very quickly and it is simple to implement. It also has the advantage of easily recovering from a poor initial solution (Bruglieri et al., 2018; Demir et al., 2012). The LNS meta-heuristics offer a local search that makes it possible to choose a locally optimum solution at each step of the optimization which may eventually lead to a globally optimum solution. This algorithm is therefore employed for this research.

An approximate solution is first found and further improvements are made with local search in the search space of the neighbourhood by using a destroy and repair technique. The search space is the space of all possible solutions for the specific problem, and the neighbourhood space is the available space after a relocation of a current solution. For this research, search algorithm ensures:

1. full or partial coverage of HCFs within a desired service distance with MCLP
2. full coverage with PCLP within endogenous maximum distance
3. the population assigned to a HCF does not exceed the HCF's threshold carrying capacity
4. budget constraints are considered by distinguishing HCFs into standard and low-cost facility
5. suitability conditions are met for locating HCFs.

The LNS algorithm takes an initial feasible solution S as input, and computes an initial best solution S_b in the neighbourhood of S through a destroy and repair method. S_b is returned as the current solution S . A temporary solution S_t can be the solution S_b or be discarded, depending on if it is accepted or not (Algorithm 2.1). Different solutions are obtained with either local search or greedy function. The local search explores the current solution to provide a better viable solution.

At initialization, the value of a number of facilities to cover the population that is not covered is first obtained from facility to population density. If a fraction is obtained, the value is rounded up to be the initial value. This value becomes the approximate required number of facilities that will provide coverage for the uncovered demand. These locations will be randomly placed within the no-coverage boundary. The optimal or near-optimal location of potential facilities will be found based on a local search, that considers maximal distance that is expected to be between a facility and the population, population within newly defined catchment, closeness of potential facility to one another, and closeness to amenities.

Further elimination and addition strategies will be done by considering the number of HCFs within the catchment of each HCF. The catchment will be an average distance computed from the bounding circles from the result of the initial catchment model in that zone. The population within the catchment will be calculated and compared with the number of agents from the HCF-population ratio. Additional HCF will be created within the catchment if required. Another suitability analysis is the HCF proximity to other amenities such as schools, religious facilities and markets, which are expected to be within the catchment. No facility is allowed within the catchment of an existing real-life facility. The objective is to locate a HCF on a suitable place and allocate a population to the HCF. Hence the name location-allocation.

A final solution is returned after all suitability analysis have been done. The output of the MCLP provides the percentage of the population that cannot reach a facility within the threshold distance or time. The relaxed form assigns the set of uncovered population to the closest facility which also includes existing facilities within the region, by relaxing the maximum distance between demand and supply.

Three different sets of candidate facility locations that will be returned are those that meet all conditions, those that are without amenities in their catchment, and those whose population size is below the threshold carrying capacity.

The LNS algorithm is broken down to six phases for the sake of this research. These are:

- Initialization phase
- Construction phase
- Destruction phase
- Repair phase
- Sorting phase
- Improvement and Stopping phase

These phases are described in the next sections. Figure 3.8 shows the flow chart of the phases.

3.8.1 Initialization

At initialization, the required number of facilities is obtained from a facility to population ratio. An estimate number of facility agents to be simulated can be the approximate required for the population or greater than the required. Let $N = \{1, 2, 3, \dots, n\}$ be the set of initial facility locations for the under-served population, and $M = \{1, 2, 3, \dots, m\}$ be the initial value of facility locations to be selected. An initial value of facility locations can be selected to be $m > n$ or $m = n$ or $m < n$. The number of HCF agents will

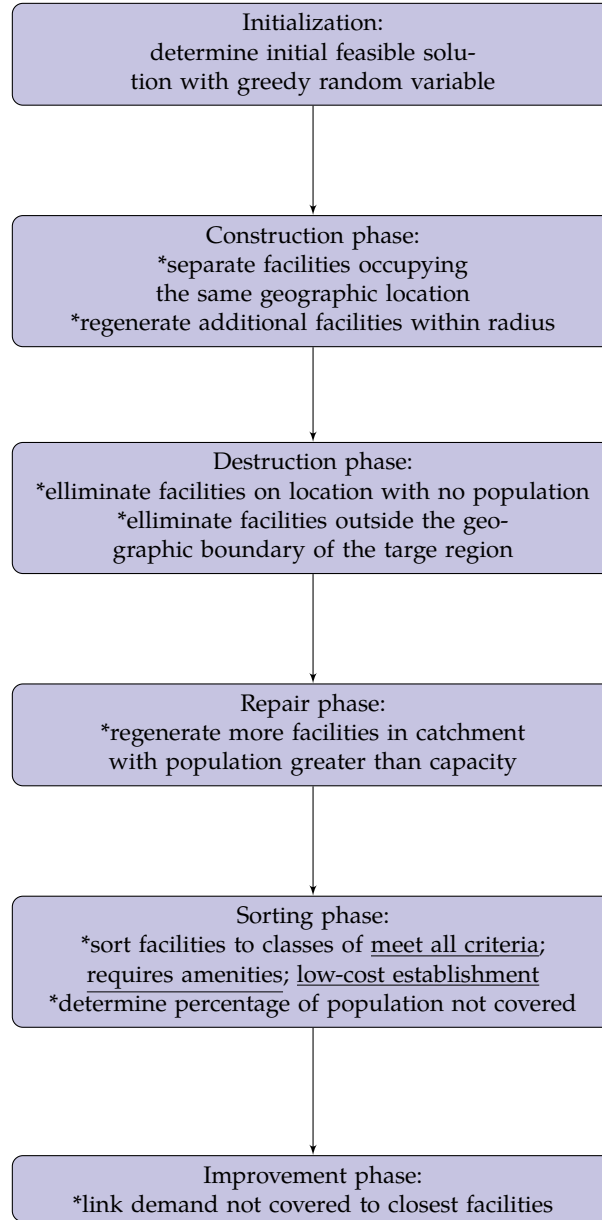


Figure 3.8: LNS Phases

eventually change using the principle of subtraction and addition to remove or add facilities as they meet certain criteria.

The population density map is imported into the ABM and the square patches adopt the number of population per size of the patch. The population will be calculated by summing up the population in all patches within the no-coverage dataset that represents the under-served area, obtained from the results of the catchment model (Equation 3.30). The total population on all patches that intersect the no-coverage feature are the uncovered demands. It is assumed that people will not reside in water and committed parcels such as airports and stadiums. Therefore, where such spatial feature intersects a patch, the population density property of the patch is assigned a

zero value.

$$\text{Population within a geographic area} = \sum_{i=n}^n PP \quad (3.30)$$

Where:

PP = total population in a patch that intersects no-coverage feature

n = number of patches

M agentset of the required number of facilities are created and randomly placed to occupy any location within the uncovered area. The required number of facilities will depend on the %coverage chosen by the user based on a budget constraint. To effectively manage and utilize available resources, policy makers can choose from different alternatives to the coverage of healthcare services by manipulating the carrying capacity of the HCF. The optimal carrying capacity suggested by (WHO, 2013) is one HCF to 5,000 population. This is taken to be 100% setting in this research. The facility density can be increased to 7,500 giving a 75% setting for ratio 1/7,500, or 10,000 to give 50% setting for ratio 1/10,000. The ratio is a choice to be made by the decision maker or user as an opportunity to minimize the number of facilities to be built. This is a user-defined variable in the model.

This forms the initial feasible solution.

3.8.2 Construction Phase

In the construction phase, three processes are involved. The first process is to place the facilities created at initialization phase randomly at different locations to form the initial feasible solution.

At initial potential HCF placement at different locations, there are no conflicts of interest and HCF can be located anywhere without considering if any other agent is occupying the same location. Since the HCFs are randomly placed, there is a possibility of some occupying the same location. If such situation occurs, one of them takes a decision to remain on the location and asks others to move until they each find a location that is not occupied by another HCF.

In the second process, the constraints of having not more than one facility in a location is enforced by separating any overlapping facilities. Each of the newly located facilities will examine its location to find out if there is any other facility sharing the same location with it. If this is so, this facility retains its position and the other facility or facilities are instructed to move to another location within the no-coverage area where no other facility is located. This forms a new feasible solution.

In the third process, if the configuration of facility placement is such that there is no other facility outside the catchment boundary of a HCF, within a neighbourhood

of at most three-quarters of the catchment radius, it implies some demands will be outside health coverage. Further densification is done by creating additional facilities within the specified distance neighbourhood of the affected HCFs. The choice of three-quarters outside the catchment is to reduce catchment overlap while still ensuring that as much demand is covered as possible. This forms another new feasible solution.

The algorithm for this processes is shown in Algorithm 3.2.

Algorithm 3.2 LNS Algorithm Construction Phase

```

1: randomly place facilities  $S$ 
2: for all facilities on  $j$  do
3:   if there is another facility on same location then
4:     tell others to relocate to another location without a facility
5:   end if
6: end for
7:  $S_b = S$ ; make this initial feasible solution
8: for all facilities  $S$  do
9:   create service area
10:  if there is no facility in neighbourhood up to three-quarter of the catchment radius then
11:    regenerate another facility in that radius
12:     $S_b = S$ ;
13:  end if
14: end for
15: return  $S_b$ 

```

3.8.3 Destruction Phase

The destruction phase consists principally of removing facilities from the existing solution. Such agent takes a decision to leave the population of HCFs through a process called agent death. That is the agent ceases to exist. The removed set of facilities do not have to be stored in the memory, because agents can be created and re-inserted. The selection of the destruction operators is governed by suitability constraint variables. At initial potential HCF placement, HCF can be located anywhere in the search space and the suitability of the location is not considered. The destruction is based upon the following considerations:

1. *Outside boundary removal:* The neighbourhood of the specified boundary is searched by each agent within the search space. For this removal operation, elimination is done for HCF from the initial solution that is not contained within the polygon boundary of the no-coverage dataset and the administrative boundary of the target region. As HCFs are placed on the centre of patches that intersect the polygon feature, some agents may fall outside the expected boundary. An illustration is shown in Figure 3.9. When agents are created, they are placed at the centre of

the patch. An agent may actually be on the patches (shaded) that intersect the polygon boundary, however it is outside the polygon boundary.

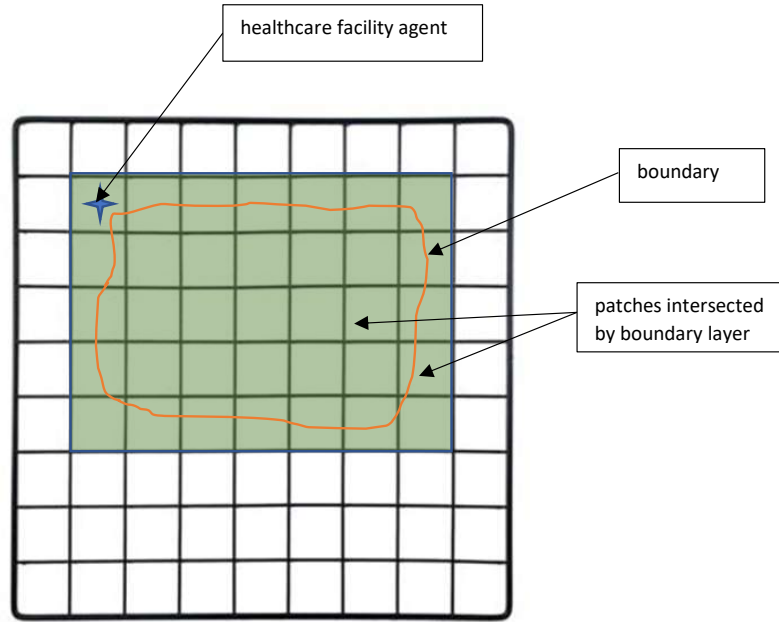


Figure 3.9: Patches and Agent location

2. *Uninhibited location removal:* In another destruction exercise, a suitability analysis is done by each HCF to evaluate if it is on a water body or on a committed land space such as airport, stadium and military cantonment regarded as not habitable. HCF that is on such uninhabited location is removed from the current solution.
3. *Neighbourhood-based destruction:* In this process, each facility checks within the neighbourhood of its catchment and instructs any other facility within that neighbourhood to cease to exist, including facilities outside a distance less than three-quarters of its catchment radius. This is to reduce catchment overlap. The search space is a circular neighbourhood that is defined by the average radius of the minimum bounding circle and is a relaxation on the network travel-distance.

An algorithm of the destruction procedure is presented in Algorithm 3.3. At initialisation of this phase, a feasible solution of the immediate past procedure serves as input. Another improved solution is returned following a partial destruction.

Algorithm 3.3 LNS Algorithm Destruction Phase

```

1: Violation of geographical location constraint
2: for all facilities  $S$  do
3:   if your location is outside neighbourhood boundary then
4:     destroy
5:   end if
6: end for
7:  $S_b = S$ ;
8: for all facilities  $S$  do
9:   if your location is on a location with no population then
10:    destroy
11:   end if
12: end for
13:  $S_b = S$ ;
14: return  $S_b$ 

```

3.8.4 Repair Phase

In this phase the partially destroyed solution will be repaired by enforcing the demand constraints. New facilities are added to the existing list of facilities by agents "giving birth" to other agents. An agent can reproduce another agent that will inherit all of its parent's properties, but a different identity. The number of new facilities to be created may not be the same number that was destroyed. This depends on the additional "births" required to meet the objectives.

The first insertion will enforce capacity constraints in the main algorithm. Each facility will evaluate the population within its catchment using a summing up of the population on the patches inside its catchment. If the population is more than the expected capacity and the excess population value is more than a predefined percentage, additional facilities will be generated inside that catchment. The newly created facilities will be randomly placed within the catchment of the parent facility. The target percentage is introduced to reduce cost. For example, if there are 8,000 population in the service area of a facility and the capacity is 5,000, one new facility will be created if the population is in excess by over 50%. This relaxation is introduced to minimize the number of under-served facilities. It is a variable factor, depending on the budget capacity.

The newly inserted facilities search their locational suitability. If there be any violation of a geographical locational constraint, they move to another location, but still within the catchment of their parent facility.

A second repair is to identify the facilities whose population size is below the target value. These set of facilities are retained to serve their community, but will be termed low-cost facilities. They will be recommended to be established with reduced

resource cost.

The repair phase therefore serves a cost-control objective. The repair algorithm is shown in Algorithm 3.4.

Algorithm 3.4 LNS Algorithm Repair Phase

```

1: Capacity constraint
2: for all facilities  $S$  do
3:   calculate demand in your catchment
4:   if demand is larger than your capacity then
5:     regenerate (additional facilities = required facility - 1) within your
       catchment
6:     for all facilities generated do
7:       if you are on same geographical location with another then
8:         relocate within the catchment
9:       else if your location is on location with no population then
10:        relocate within the catchment
11:      else
12:        stay within the catchment
13:      end if
14:    end for
15:   else if demand is below target then
16:     form a list of low-cost facilities
17:   else
18:     all conditions are met by you
19:   end if
20: end for
21:  $S_b = S$ ;
22: return  $S_b$ 

```

3.8.5 Sorting Phase

This sorting phase sorts the selected facilities by the facilities with other institutional facilities in their catchments, facilities serving population below the threshold population capacity, and facilities with no institutional facilities within their catchments. It separates low-cost facilities, where the cost is defined as the cost of establishing a facility in relation to the population to be served by the facility.

Algorithm 3.5 LNS Algorithm Sorting Phase

```

1: Sort facilities by close amenities and establishment cost
2: for all facilities  $S$  do
3:   if you have no amenities within your catchment then
4:     form a list of "no-amenities" facilities
5:   end if
6: end for
7: for all facilities  $S$  do
8:   if facility status is low-cost then
9:     change state
10:  end if
11: end for
12:  $S_b = S$ ;
13: for all facilities  $S$  do
14:   if your location is on location with no population then
15:     leave the solution
16:   else
17:     remain on your location
18:   end if
19: end for
20:  $S_b = S$ ;
21: return  $S_b$ 
22: calculate percentage with no-coverage

```

3.8.6 Improvement and Stopping Phase - PCLP

The capacity and distance constraints are relaxed in this phase. Uncovered populations are linked to any closest facilities to them. The percentage population uncovered will be calculated from:

$$UP\% = \left(\frac{UP}{IP}\right) * 100\% \quad (3.31)$$

Where:

UP = total population not within the catchment of the new facilities

IP = initial total population that intersects the no-coverage feature

The improvement and stopping phase algorithm is shown Algorithm 3.6.

Algorithm 3.6 LNS Algorithm PCLP Phase

```

1: Improvement and Stopping based on PCLP
2: for all demands  $i$  with no coverage do
3:   create links with facilities  $S + \text{existing facility}$ 
4:   if more than one facility is the closest then
5:     create link with one of closest facilities
6:   end if
7: end for
8:  $S_b = S$ ;
9: return  $S_b$ 

```

3.9 Visualization and Analysis

All results can be viewed in the NetLogo window and output platform. It is possible to view the model process and agents' movements as the model runs. The speed setting on the NetLogo interface can be set to watch the process slowly or at a higher speed. Visualisation of model results in GIS is made possible through the available feature in NetLogo that allows agents to be saved as points, lines or polygon features as shapefiles by specifying file names and locations. The spatial data adopts all attributes of the corresponding agents in the ABM. The files are then open in GIS software for visualisation and other analysis such as overlay, integration and extraction.

3.10 Summary

This chapter has discussed the spatial data requirement and their acquisition method, including how the datasets are represented. Further explanation was provided on the types of agents and their characteristics for the Spatial-ABM centre facility location problem. The methodology and algorithms for locating HCFs and allocating demands to HCFs were also explained. The next chapter discusses the testing and implementation procedure of the models.

Chapter 4

Model Testing and Application to Existing Case

The previous chapters have discussed the techniques and processes of accessibility and covering location problems. The different data acquisition processes and data structures have been highlighted. In addition, the varying spatial datasets and characteristics of agents required to achieve the aim of this research were identified.

It is absolutely important that a model is verified and validated before using it to answer any research question, for policy applications or to support decision making. This can be achieved through different test processes on the internal behaviour and output of the model (Abdou et al., 2012; Fagiolo et al., 2017). The varying test approaches applied on the travel-time and facility location models developed for this research are described in this chapter. Using the road network data of Lagos State, the study area, the travel-time model results are compared with the results from Geographic Information Systems (GIS) and Google Maps web mapping service. The procedures outlined for the Spatial Agent-Based Model (ABM) optimization are tested using the Swain 55-node dataset (adapted from ReVelle et al. (2007)), which is extensively utilized in testing covering location algorithms.

4.1 Description of the 55-Node Dataset

The Swain 55-node dataset consists of the geographic (X,Y) coordinates and population values of 55 location nodes (Figure 4.1) that are spatially distributed, representing locations of demands that can be used to test the covering facility location model. The total population on all the nodes is 6400. In applying the Large Neighbourhood Search (LNS) meta-heuristics Spatial-ABM algorithm to the 55-node dataset, the geographic locations of the nodes were plotted in GIS and a polygon feature was created to encompass the nodes, representing a geographic boundary beyond which no healthcare facility (HCF) can be placed. These features were exported to the ABM with the population attribute values of the nodes. A scale factor of a GIS distance to ABM distance

was applied to the model from the distance in the map. It is assumed that the boundary feature, including population locations are areas with no healthcare coverage covered with virtual regular grids or patches that are potential discrete locations for HCFs. A patch adopts the population value of any node that it contains. This indicates that any patch with no nodes has zero population. The total population within the catchment of a HCF represents the demands that have access to that HCF services.

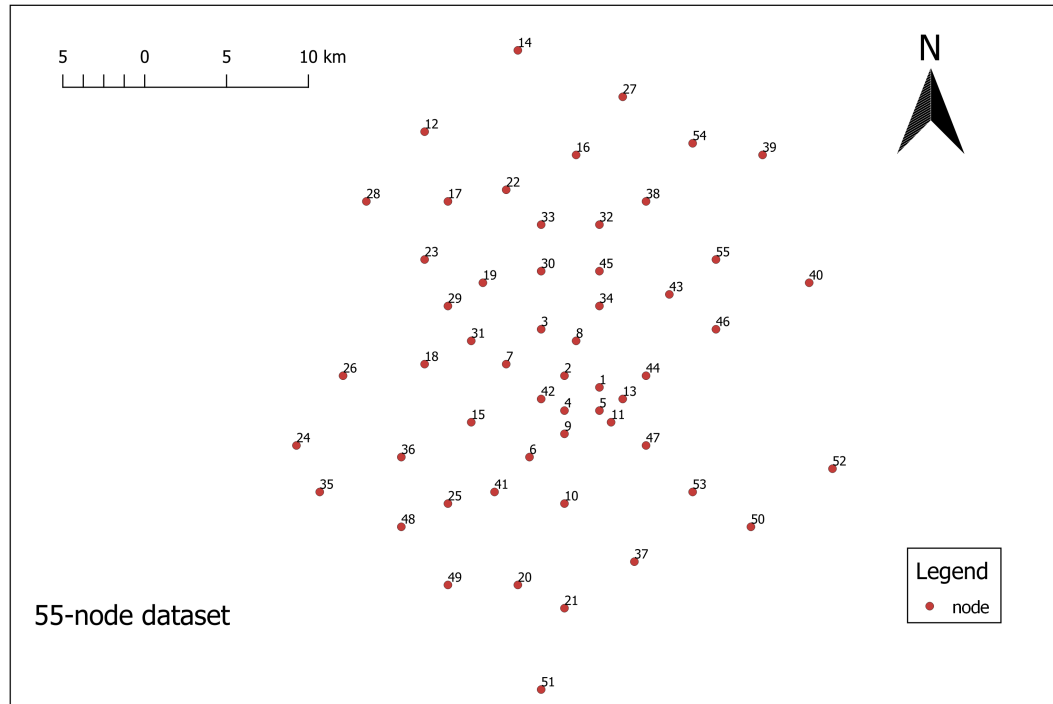


Figure 4.1: 55-node map (ReVelle et al., 2007)

4.2 Model Verification and Validation

The information produced by a simulation model is credible and most valuable if it can be verified and validated. Verification is performed to ensure that the model is programmed correctly and does not contain errors in the code and the algorithm that make the model not perform as expected. Validation on the other hand tests if the model correctly represents and reproduces the real-world system, and meets the intended result requirements (Macal, 2005).

Varying techniques exist for model verification and validation, depending on the purpose for which it is created or the data that is available. This research has adopted the following techniques:

Parameter variability or sensitivity analysis This entails changing the values of the input parameters or the initial condition of the model to ascertain the effect on the behaviour of the agents and the model output.

Compare the results of several replication This is done to check the consistencies and robustness of the model.

Model-to-model comparison This compares the output of the model with other models. The reliability of the outputs from the model is confirmed through this process.

With these done, the model can be useful in addressing the problems, and provide accurate and reliable information about the phenomenon that is being studied.

The NetLogo ABM tool provides a feature called Behavior Space that allows several runs of a model. It offers parameters to be varied and combined differently for each model run. The results of the model execution can be compiled in a spreadsheet format for further analysis. This feature was utilized to test the performance of the model and examine the behaviour of the agents and their characteristics to confirm if they conform with what is expected. The experiments were performed on a desktop Intel (R) Core i5 CPU @ 1.85 GHZ with 1.98 GB RAM.

The tests are discussed in the following sections.

4.3 Travel-Time Model Test

The tests on the travel-time model were based on:

- the agent's travel-time and geographic location within the model using a network measure;
- the estimated catchment area of the HCFs with a straight-line measure; and
- model-to-model comparison of the agent's network travel-time and geographical location with Google Maps on-line results.

The travel specifications were 30 minutes travel-time limit (World Health Organization, 1998) and 0.8m/s travel-speed (Olowofoyeku et al., 2019).

4.3.1 Verification on Travel Time-Limit and Geographic Location

This test used the Ikorodu Local Government Area (LGA) datasets in Lagos State, consisting of 20 healthcare facilities (HCFs). The verification tested the agents' network travel-time from the HCFs to their final locations based on repeated runs on varied destination inputs.

On each of the HCFs, 50 patient agents were simulated to travel along the road networks from the HCF to various randomly simulated destinations placed in the environment. The travel scenario of the 50 patients formed one model run. With five model runs, 250 patients were simulated on a HCF giving a total of 5,000 patients for the 20 HCFs in the LGA. A new set of destination agents was simulated for each run to vary their locations. Each patient took an independent decision to select a destination

and travelled towards its chosen destination (described in section 3.4). Locations of the patients at the end of the journey indicated that they had reached the travel-time limit or they got to their destinations before the threshold time. These locations were saved in a GIS shapefile format.

The test confirms:

1. *if there is similar behaviour in agents that select the same destination or travel path*
2. *if agents stop within the time-limit, hence distance-limit*
3. *the variation of time spent by agents to complete their journey*

The procedure was repeated twice. In the first procedure (designated network1 in Table 4.1), 4,443 patient agent locations were at the end of the time threshold while the second procedure (designated network2 in Table 4.1) had 4,588 locations out of the 5,000 simulated patients. Other patient agents reached their destinations before the time-limit.

It was observed that:

- Some patients reached their destinations before the time threshold. Therefore, the number of locations returned at the travel-time limit was less than the number of patients that was simulated.
- Some patients' destination choice was similar. Such patients exhibited the same behaviour and travelled on the same paths. Hence, their travel-time was the same and they stopped on the same location. There is an indication that the same paths could lead to other destinations within the same region as well.

Table 4.1 shows the spread between the travel-time and travel-distance in each procedure, including the standard deviation. The results show that overall, the maximum travel-time by the agents was 34.85 minutes and the minimum was 30.03 minutes, corresponding to maximum distance of 1672.89 m and minimum distance of 1441.42 m.

Table 4.1: Travel-Time Distance and Time Experiment Result

| Procedure | No of agents | | Time (minutes) | | | Distance (metres) | | |
|-----------|--------------|--------|----------------|-------|---------|-------------------|---------|---------|
| | input | output | Max | Min | Std dev | Max | Min | Std dev |
| Network1 | 5000 | 4443 | 34.85 | 30.03 | 0.884 | 1672.89 | 1441.42 | 42.46 |
| Network2 | 5000 | 4588 | 34.26 | 30.03 | 0.808 | 1644.47 | 1441.53 | 38.80 |

4.3.2 Estimated Catchment Area Test with Circular Buffer

This test also used the Ikorodu LGA to compare the ABM HCFs catchment area with the traditional GIS buffer based on the straight-line measure. The agents' behaviour was based on: (i) distance-specified rule (ii) time-specified rule.

- *Distance-specified rule*

First, a set of patient agents was created on each HCF to move forward at a specified distance radius of 1,440m (distance covered for 30 minutes at 0.8m/s travel-speed) from the HCF, similar to drawing a buffer circle in GIS (as-the-crow-flies). Each agent randomly chose the direction to go, covered the user-specified distance in a single step on a straight line and stopped thereafter. The final point locations of the agents were saved as a shapefile in ABM.

The Minimum Bounding Circle (MBC) (described in section 3.4.1) encompassed the agents' locations in the GIS to define the catchment. Another catchment feature was created by drawing a circular buffer centred on each HCF using the same 1,440m distance radius.

- *Time-specified rule*

For the second catchment test, the agents travelled along a straight-line in their chosen directions from each HCF. The difference is that a time-limit is specified for the agent to stop while moving a step at a time on the patches. The distance covered at each time step is calculated within the model similar to the network journey. The travel threshold was 30 minutes on a speed of 0.8m/s.

The agents' locations were found to be the same as the distance-specified procedure. The patch size was varied to observe the effect on the agent's journey.

Agents' geographical locations were spatially joined with MBC in the GIS. Based on the size of the patches, the covered distance by each agent was 1,505m. The value was used as buffer distance for catchment of HCFs in the GIS.

All catchment features were dissolved and clipped with the geographic boundary of the LGA, to extract the area and percentage coverage of the resulting polygon feature representing areas outside healthcare coverage (Olowofoyeku et al., 2019).

The results showed that each agent's travel-distance and travel-time were the same as expected. Since the agents move to the centre of the patch on each step taken, the covered distance is a function of the size of the patch. Therefore there is little variation in the travelled distance with respect to patch size.

Percentage gap between the ABM coverage and GIS coverage for the different specifications was computed from Equation 4.1:

$$gap\% = ((ABM_{value} - GIS_{value}) / GIS_{value}) * 100 \quad (4.1)$$

where:

ABM_{value} is the area or percentage coverage obtained with ABM

GIS_{value} is the area or percentage coverage obtained with GIS

The GIS and ABM catchment area values were found to be the same. The differ-

ences in coverage values were less than 0.02 km² and 0.02%. The little difference may be attributed to scale and projection errors, considering that small differences were observed in the diameter of the constructed MBCs instead of being equal values. The direct movement of agents on patches for the time-specified journey has 0.00% gap compared to a straight-line radius gap of 0.02%. The results are presented in Table 4.2.

Table 4.2: Comparison of ABM and GIS catchment area estimates using Euclidean distance equivalent for 30 minutes travel time

| Coverage | 1.440 km distance-specified | | | | 1.505 km time-specified | | | |
|-------------------------|-----------------------------|--------|------|-------|-------------------------|--------|------|-------|
| | ABM | GIS | diff | gap % | ABM | GIS | diff | gap % |
| Area (km ²) | 113.98 | 114.00 | 0.02 | 0.02 | 122.85 | 122.85 | 0.00 | 0.00 |
| % coverage | 29.15 | 29.16 | 0.01 | -0.03 | 31.42 | 31.42 | 0.00 | 0.00 |

The experiments revealed that the model scale or the size of the patches has impacts on the results. This is generally true for geographic representation of objects.

4.3.3 Validation With On-Line Travel-Time Application

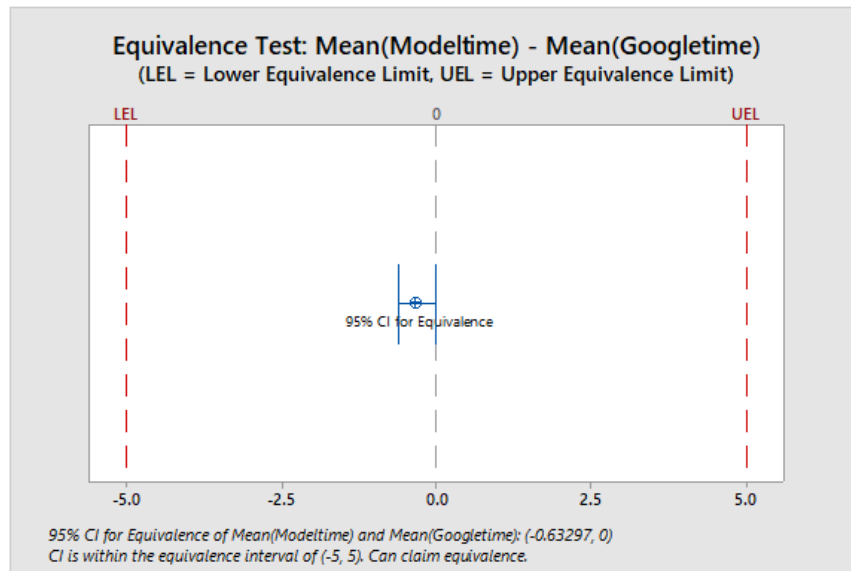
This test was carried out to compare the simulated network travel-times with the Google Maps on-line results. The chosen regions for comparison within Lagos State had different road network characteristics.

Agents were simulated for network travel measure with a 30-minute time threshold and their locations saved in a GIS shapefile format together with their travel-time and distance attributes. The shapefiles were imported into the GIS and additional attribute fields were created to contain generated real-world geographical coordinates (latitudes and longitudes) from the agents' locations. The coordinates were used to pass a request in the Google Distance Matrix Application Programming Interface (API) to obtain the walking travel-time and travel-distance from the target HCF as origin to the agents' positions as destinations.

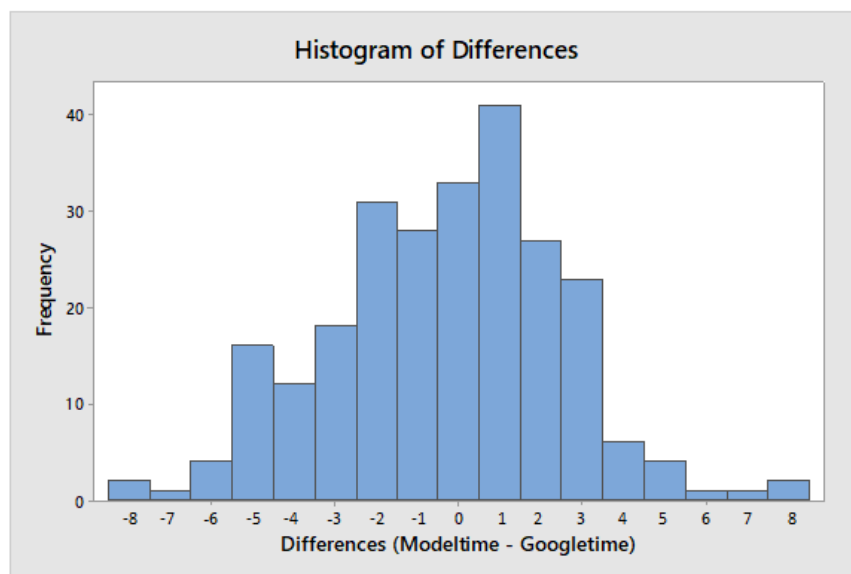
The travel speed calculated from the Google Distance Matrix API distance and time revealed that Google Maps uses variable speed. Therefore for a consistent comparison, the travel-speed for each set of data was used to obtain equivalent travel-time value to agree with the 48m/min analysis speed. Some outliers were observed in the on-line results. This is further discussed in Chapter 5.

A paired Equivalent Test was done to determine whether the means of the two modes of travel-time estimate are equivalent within ± 5 minutes. The results are presented in Figure 4.2 and Table 4.3. Without the outliers the equivalent tests confirmed equivalence and shows that there is no significant difference between the means. The more structured road networks showed good agreement with the paired travel-time estimates.

The Confidence Interval (CI) is within the equivalence interval, and the greater of the two P-Values is $<10^{-5}$. The models are therefore considered to be equivalent.



(a) equivalence test for paired spatial ABM and Google travel times



(b) Differences of paired spatial ABM and Google travel times

Figure 4.2: Travel time equivalence test with on-line road data: Googletime = travel time on Google Maps; Modeltime = travel time on Spatial-ABM

Table 4.3: Equivalence Test with Paired Data: Spatial-ABM travel time, Google Maps travel time

Method:

Test mean = mean of Spatial-ABM travel time

Reference mean = mean of Google Maps travel time

| Descriptive Statistics | | | |
|---------------------------|---|---------|------------------------|
| Variable | N | Mean | StDev |
| Modeltime | 250 | 31.178 | 1.160 |
| Googletime | 250 | 31.518 | 2.865 |
| Mean Difference | | | |
| Difference | StDev | SE | 95% CI for Equivalence |
| -0.340 | 2.799 | 0.177 | (-0.633, 0) |
| Test | | | |
| Null hypothesis: | Difference \leq -5 or Difference \geq 5 | | |
| Alternative hypothesis: | -5 < Difference < 5 | | |
| Alpha (α) level: | 0.05 | | |
| Null Hypothesis | DF | T-Value | P-Value |
| Difference \leq -5 | 249 | 26.317 | $<10^{-5}$ |
| Difference \geq 5 | 249 | -30.166 | $<10^{-5}$ |

With the results of these tests, it is confidently confirmed that the travel-time model is robust, consistent and reliable.

4.4 Spatial-ABM Optimization Model Verification

The verification on the optimization model involved:

- Parameter variability and comparison of results of several replications performed on the 55-node dataset and Lagos State dataset. As described in Section 3.8, the LNS spatial optimization algorithm returns a final set of endogenously selected facilities that is preceded by an initial feasible solution of the number of facilities. With these tests, the clear understanding of the effect that the initial p value has on the convergence of the model results can be revealed, in terms of: the number of facilities selected and the consistency of convergence.

4.4.1 Initial Number of Facilities Variation on 55-Node Dataset

At initialization, the 55 nodes were fixed potential facility locations. Additional potential HCFs were simulated and placed randomly in the environment so that the location of HCFs would not be restricted to the nodes. The number of facilities added were varied to verify the effect it has on the number of selected facilities, p . The candidate p -facility simulated were 50, 200, 500, and 600 in addition to the 55 fixed facilities. Using a service distance $D = 11$ and mandatory closeness distance $S = 15$, 200 model simulation runs were performed for each set of initialization to:

1. *endogenously determine the number of facilities that will ensure total coverage of health-care within D or S distance units*
2. *the consistency of the number of facilities selected*
3. *the effect of initial feasible solutions on the model output*

(The node locations in the 55-node dataset are defined by latitude and longitude geographical coordinates, therefore no distance unit is considered in the literature.)

The results show that for each initial p value in the 200 model runs, the selected number of facilities are between the range of five and eight as shown in the line plots in Figure 4.3a. The results indicate that:

- Maximum number of facilities selected from the different initialization input was 8, and the minimum was 5. The highest value was from the highest number facilities input - 600.
- The median for all sets is 6.0 and all means range between 6.2 and 6.3.
- 50% of the output values for all initialization is between 6 and 7.

The model is therefore consistent and robust. Although, the outputs absolutely agree, the more the number of initial facilities, the longer it takes for the model to converge.

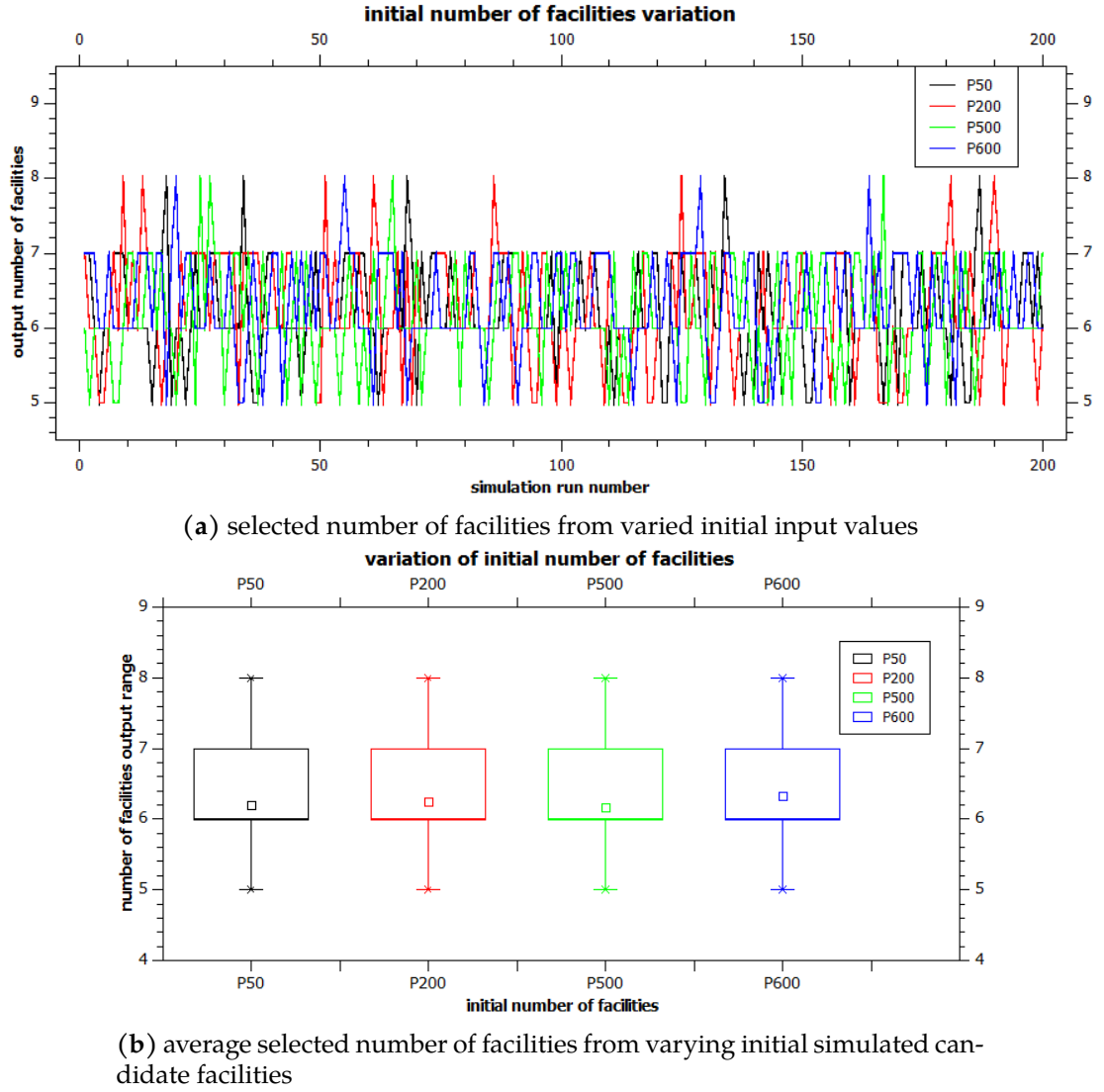


Figure 4.3: 55-node dataset: selected number of facilities from varying initial candidate facilities

4.4.2 Initial Number of Facilities Variation on Lagos State Dataset

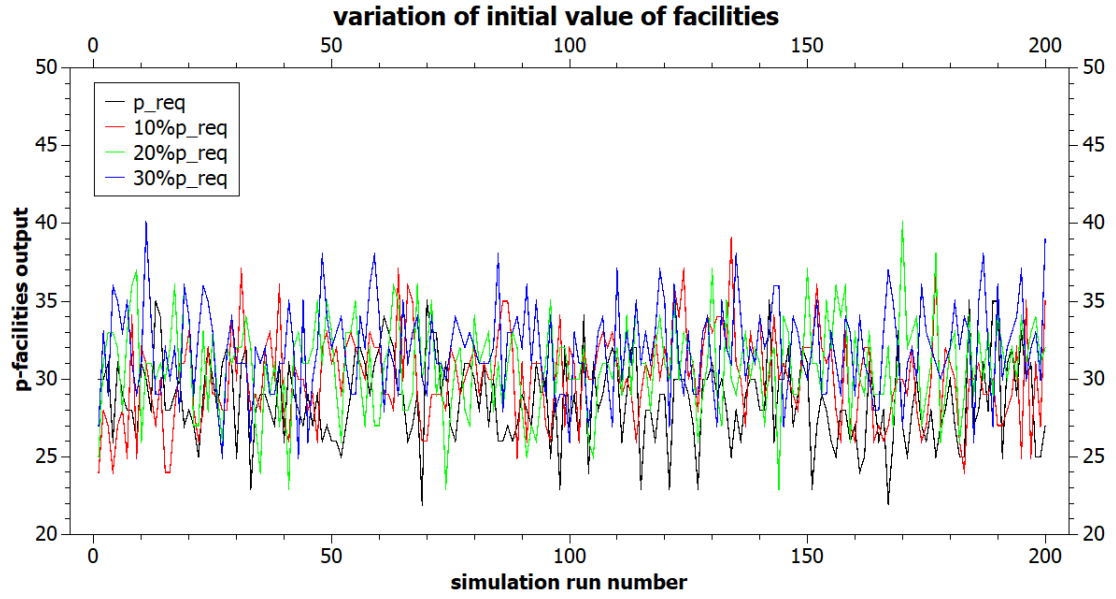
The same initial facilities variation test performed on the 55-node dataset was also carried out on the Lagos State dataset. This was necessitated by the rich geographic features that will enhance the considerations of additional criteria for solving the facility location problem. For example, facility location on water or an uninhabited zone is prohibitive, however, such spatial data is not available for the 55-node dataset. The population in the Lagos State dataset is also not represented on nodes, but as population density within the gridded plane. The initial value of p is the value obtained by dividing the population by the expected capacity of a facility, which for this research is 5,000. The number of facilities as initial feasible solution for the model had four variations by increasing the calculated value, p by 10%, up to 30% as: p , $(p + 10\%p)$, $(p + 20\%p)$, and $(p + 30\%p)$.

For example, if the calculated required value for p is 20, the initial number of facilities for the experiments will be varied as 20, 22, 24, 26 and each set will have 200 runs. For this experiment, the required number of HCFs for the population was 29.

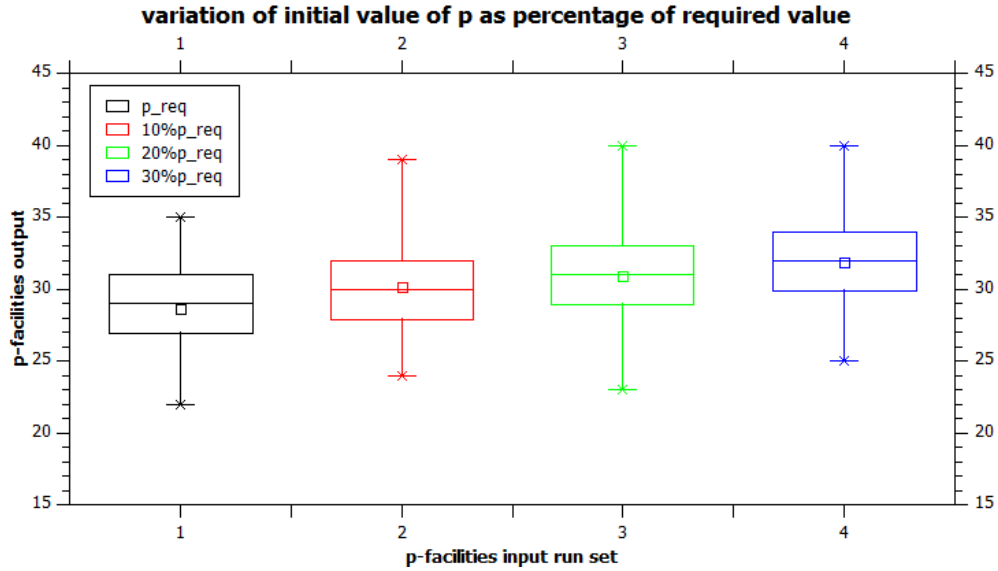
The line plots are shown in Figure 4.4. The set of runs that is initialized with the actual required number of facilities is designated 1. 2 is the set initialized with 10% of the required number of facilities, while 3 and 4 are initialized with 20% and 30% of the required number of facilities respectively.

The results show that the selected number of HCFs are between the range of 22 and 40 as shown in the line plots in Figure 4.4a. The box plots in Figure 4.4b shows that:

- for initialization with required p , minimum output of number of selected facilities is 22 and maximum is 35 with median value of 29.
- for initialization with additional 10% p , minimum output is 24 and maximum is 39 with median value of 30.
- for initialization with additional 20% p , minimum output is 23 and maximum is 40 with median value of 31.
- for initialization with additional 30% p , minimum output is 25 and maximum is 40 with median value of 32.
- The median of each set increases as the input value increases.
- 50% of the output values for initialization with additional 30% p is higher than the required number of facilities, p .



(a) selected facilities from varied initial value required for population



(b) average selected facilities from varied initial value required for population

Figure 4.4: Lagos State dataset: selected number of facilities from varying initial candidate facilities

The results are similar and consistent. However, in addition to the observations from experimenting with the 55-node dataset, the median value tends to be away from the required value of 29 (obtained from facility-population ratio) as the initial value increases.

4.4.3 Statistical Robustness of Coverage

The Maximal Covering Location Problem (MCLP) model covers a percentage of the population with a selected number of facilities within a specified service distance, D and returns the percentage of population that is not covered. The behaviour of the agents in this model was statistically tested with repeated simulations using four re-

gions in Lagos State - Badagri, Ikeja, Etiosa and Ibeju-Lekki, representing north, east, south and western parts of the state. For each region, four independent 100 simulations were carried out to produce 400 outputs and their results were statistically analysed. A box plot of each set was drawn to compare the output of percentage of population that is not covered (designated uncovered population in Figure 4.5) in the area of focus.

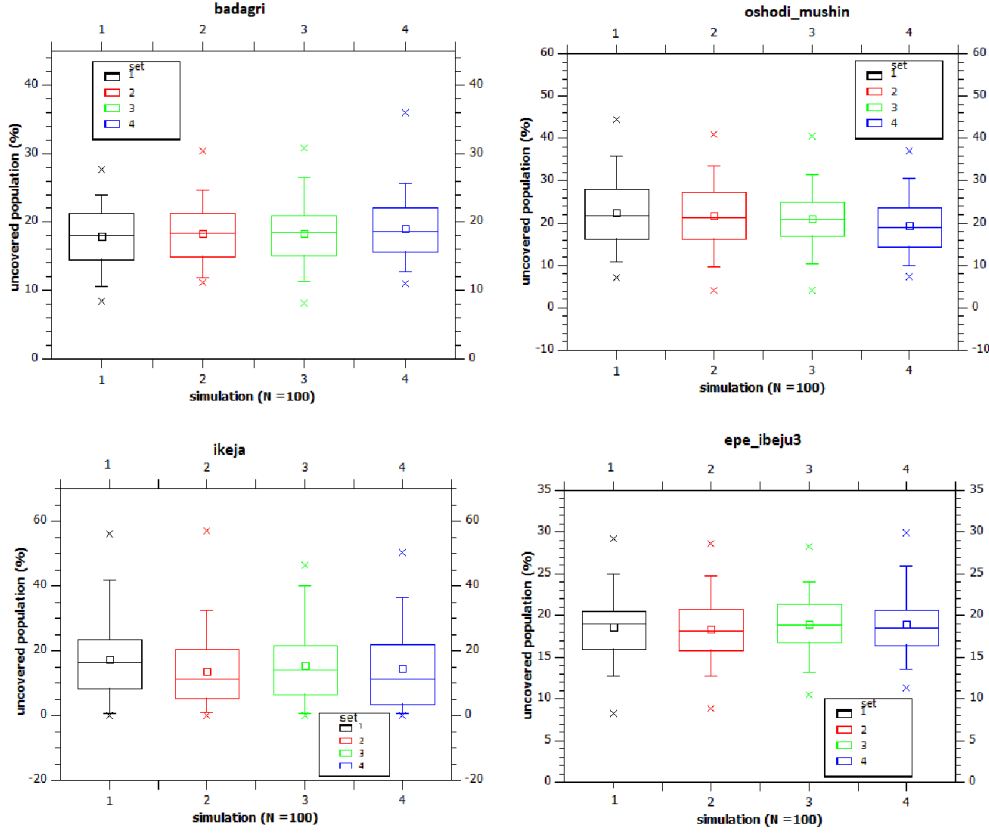


Figure 4.5: Coverage verification analysis

Having applied the covering model to different datasets with repeated runs, the results indicate that the model is robust and consistent in the selected number of HCFs and the population coverage.

The following observations were made from the experiments:

- Regardless of the value at initialization, the LNS algorithm is capable of recovering to produce expected outcomes.
- The initial value of p did not affect the final solution. The results show that the behaviour of the HCF agents at searching the neighbourhood and taking independent decisions to stay in suitable locations, including generating more HCF agents is reflected in the number of HCF agents selected. The graphs show clearly that the initial value of p did not affect the behaviour of the agents at replicating and finding suitable locations.
- Another interesting characteristic of the covering model is that a different number

of facilities, p was selected and coverage was not the same at each initialization. This reflects the randomness of the initial potential location and the spatial distribution of the population that is also a candidate facility location. It also confirms the flexibility and adaptability of the model as a useful decision-support tool.

The following section describes the validation process for the LNS meta-heuristic facility location model.

4.5 Location-Allocation Validation on 55-Node Dataset

The location-allocation model was validated on the 55-node dataset as a model-to-model comparison technique. In this process, the outputs of the model for this research are compared with optimal linear programming (Church and ReVelle, 1974; Church, 1984) and heuristic (Galvão and ReVelle, 1996) solutions from other published works.

The 55-node dataset has been used extensively for analysis of different facility location applications. It has significant indication in the works of Church and ReVelle (1974), Church (1984) and Galvão and ReVelle (1996) as a general form of MCLP analyses. While there is no benchmark for testing the new capacitated mandatory closeness MCLP proposed in this research, Church and ReVelle (1974) provided an additional optimal solution for 5-facility MCLP with mandatory closeness on the 55-node dataset which is also used for validating this research.

The 55-node Swain dataset comprises of population locations and size of the population at each node. For this research, the population values of Church and ReVelle (1974), Church (1984) and Galvão and ReVelle (1996) are multiplied by a factor of 10 based on ReVelle et al. (2007), bringing the total population at the nodes to 6400. Church and ReVelle (1974) solved the MCLP and MCLP with mandatory closeness constraints on the 55-node dataset using heuristics and linear programming with candidate facilities restricted to the nodes, while Galvão and ReVelle (1996) applied a Lagrangean heuristic for the MCLP with facilities also restricted to the nodes. Church (1984) considered placing facilities anywhere in a continuous space or an infinite plane. The Spatial-ABM proposed in this research locates facilities at the centre of a finite network of grids or patches within the boundary of the dataset, including the nodes. The centroid of each patch is therefore a potential facility location.

In applying the dataset to this research, the whole population is assumed not to be covered with healthcare services and endogenously determined p healthcare facilities will be sited for the population to be totally covered with healthcare services. A certain percentage of the population can be covered within a fixed service distance, D , while ensuring that the uncovered population will travel within a mandatory closeness distance, S . Of these values, the input parameters are D and S distance units, while the model returns the values of:

- p number of HCFs;

- population that is covered within p ;
- the maximum distance between the uncovered population and the closest of the HCFs.

The spatial datasets of the boundary and nodes were integrated into the ABM. Potential facility agents were created and placed randomly in the environment, in addition to the 55 candidate facility agents on the nodes. The agents are eliminated based on the suitability of locations and conflicting interests. For example, agents outside the boundary are eliminated from the set of facility agents. Other remaining agents search the neighbourhood of the specified distance radius, D and sums up the populations on the patches within its catchment. Link agents were created to connect facility agents to demand nodes within the specified covering distance radius. Another set of link agents were created to connect the demand nodes outside the MCLP covering distance, D to the closest selected facility within the mandatory closeness distance radius, S . The longest of these links is the maximum distance between the farthest demand node and its closest facility.

4.5.1 Decision Support for Number of HCFs and Service Distance

Following the service distance $D = 10$ and mandatory closeness distance $S = 15$ settings of Church and ReVelle (1974), different solutions that will give total coverage within the specified distances were obtained. A total coverage is desired within S distance units and if achievable, within D distance units. Therefore, any population outside the mandatory distance $S = 15$ is considered to be out of coverage.

With 200 model runs, a set of solutions was obtained with values of p ranging from $p = 6$ to $p = 11$. With a reduction of the map size, the values of p ranged from $p = 3$ to $p = 7$. This reveals the significance of scale in spatial data representation and analysis of location-allocation and the effect of aggregation of data. The larger the map, the more the number of patches, and hence the number of potential facility locations. The smaller scales returned fewer facilities than the larger scales. For solutions with 100% coverage within the specified distances, the maximum distances from the farthest population to the closest facility ranged from 9.99 to 14.86 distance units. Table 4.4 shows the range of p values and distance range for the scale settings.

Table 4.4: Range of number of healthcare facilities, p and service distance for 100% coverage

| range of p (number of facilities) | range of service distance (S distance units) |
|--|--|
| 3 to 11 | 9.90 to 14.86 |

This approach is an effective means to conclude that:

- MCLP can be solved for the dataset with number of facilities ranging from 3 to 11.

- The number of facilities that can ensure total coverage is between 3 and 11.
- Total coverage is achievable with distances ranging from 9.90 to 14.86 distance units.

As a major contribution in this research, the unique capability of the Spatial-ABM to give better understanding and an overview of the coverage requirements are revealed. This is a valuable decision support for determining the covering distance parameter and the number of facilities that can ensure healthcare coverage. This is one of the advantages of ABM over mathematical modelling where the first principle is used to generate a set of data that can create a general theory.

4.5.1.1 Population Coverage

Based on the previous findings, the values of D were set to 10, 11, 12, 13, 14 and 15 to test if total coverage is possible within these distances, still using mandatory closeness distance 15 as stated by Church and ReVelle (1974). As mentioned earlier, the number of facilities is not restricted. The parameter variation is specified in the Behaviour Space tool for multiple runs of the algorithm.

Table 4.5 shows the results of these varying parameters. The maximum distance that the farthest population will cover is provided, including population within the desired covering distance D and the mandatory coverage distance threshold $S = 15$. Table 4.6 summarises the percentage coverage for a specified D and the corresponding percentage coverage within S with p -facility. The comparison with other published research works are based on these simulation results.

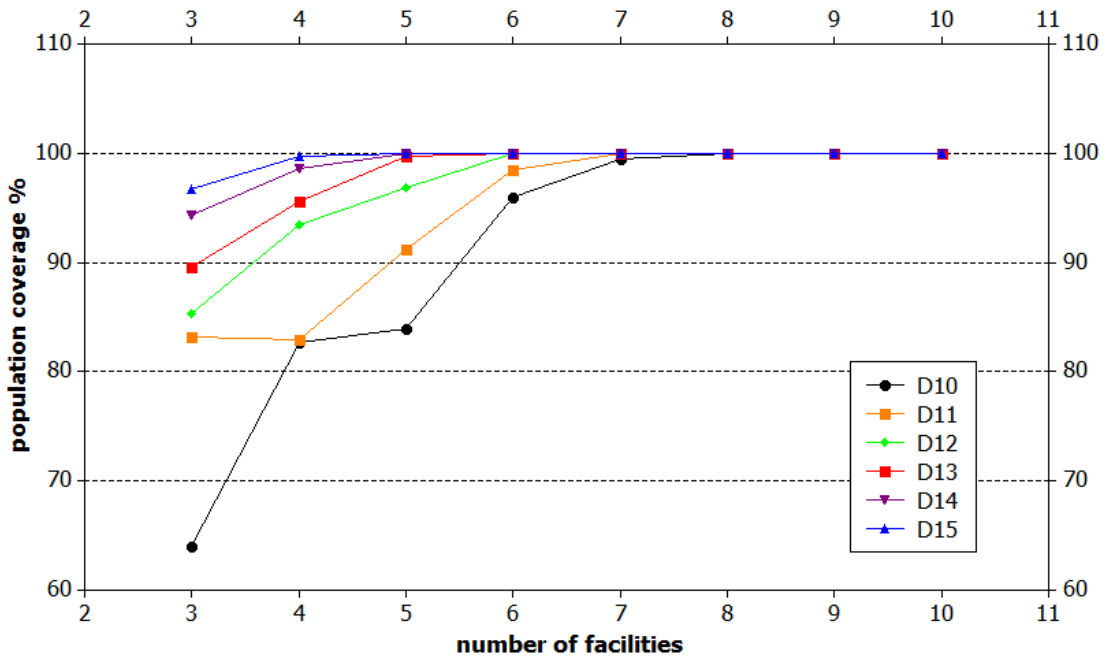


Figure 4.6: Percentage coverage by endogenous facility numbers for 55-node

Table 4.5: Results for Capacitated Mandatory Closeness-MCLP for 55-node

| D | p | coverage within D | | coverage outside D | | coverage within $S = 15$ | | Distance out of D | | |
|-----|-----|---------------------|-------------|----------------------|--------------|--------------------------|-------------|---------------------|--------------|--------------|
| | | max % | max pop | pop | % | % | pop | max dist | diff of S | diff of D |
| 10 | 3 | 64.06 | 4100 | 2300 | 35.94 | 88.59 | 5670 | 20.47 | 5.47 | 10.47 |
| | 4 | 82.03 | 5250 | 990 | 15.47 | 94.06 | 6020 | 22.50 | 7.50 | 12.50 |
| | 4 | 82.66 | 5290 | 1110 | 17.34 | 93.75 | 6000 | 18.49 | 3.49 | 8.49 |
| | 5 | 83.91 | 5370 | 1030 | 16.09 | 100.00 | 6400 | 14.60 | -0.40 | 4.60 |
| | 6 | 95.94 | 6140 | 260 | 4.06 | 100.00 | 6400 | 13.12 | -1.88 | 3.12 |
| | 7 | 99.53 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 8 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 9 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | | | | | | | | | | |
| 11 | 3 | 83.13 | 5320 | 1080 | 16.87 | 94.37 | 6040 | 20.59 | 5.59 | 9.59 |
| | 4 | 86.25 | 5520 | 720 | 11.25 | 93.59 | 5990 | 17.89 | 2.89 | 6.89 |
| | 5 | 91.25 | 5840 | 560 | 8.75 | 100.00 | 6400 | 14.66 | -0.34 | 3.66 |
| | 6 | 98.44 | 6300 | 100 | 1.55 | 100.00 | 6400 | 13.51 | -1.49 | 2.51 |
| | 7 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 8 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| 12 | 3 | 85.31 | 5460 | 940 | 14.69 | 93.75 | 6000 | 22.51 | 7.51 | 10.51 |
| | 4 | 93.13 | 5960 | 440 | 6.88 | 98.59 | 6310 | 18.20 | 3.20 | 6.20 |
| | 4 | 93.44 | 5980 | 420 | 6.56 | 96.25 | 6160 | 17.26 | 2.26 | 5.26 |
| | 5 | 90.16 | 5770 | 630 | 9.84 | 100.00 | 6400 | 12.88 | -2.12 | 0.88 |
| | 5 | 96.88 | 6200 | 200 | 3.13 | 100.00 | 6400 | 14.44 | -0.56 | 2.44 |
| | 6 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 7 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | | | | | | | | | | |
| 13 | 3 | 89.53 | 5730 | 670 | 10.47 | 95 | 6080 | 22.42 | 7.42 | 9.42 |
| | 4 | 95.00 | 6080 | 200 | 3.13 | 98.13 | 6280 | 17.38 | 2.38 | 4.38 |
| | 4 | 95.63 | 6120 | 280 | 4.87 | 97.81 | 6260 | 22.84 | 7.84 | 9.84 |
| | 5 | 99.67 | 6380 | 20 | 0.31 | 100.00 | 6400 | 14.25 | -0.75 | 1.25 |
| | 6 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 7 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| 14 | 3 | 94.38 | 6040 | 360 | 5.63 | 96.72 | 6190 | 21.19 | 6.19 | 7.19 |
| | 4 | 98.59 | 6310 | 90 | 1.41 | 98.59 | 6310 | 16.00 | 1.00 | 2.00 |
| | 5 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 6 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| 15 | 3 | 96.72 | 6190 | 210 | 3.28 | 96.72 | 6190 | 19.35 | 4.35 | 4.35 |
| | 4 | 98.91 | 6330 | 70 | 1.09 | 98.91 | 6330 | 16.24 | 1.24 | 1.24 |
| | 4 | 99.69 | 6380 | 20 | 0.31 | 99.69 | 6380 | 18.91 | 3.91 | 3.91 |
| | 5 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | 6 | 100.00 | 6400 | - | - | 100.00 | 6400 | - | - | - |
| | | | | | | | | | | |

Table 4.6: Percentage coverage within D and corresponding coverage within S

| coverage within D (%) | | | | | | | | | |
|-------------------------|-------|-------|-------|-------|-------|------|------|-------|--|
| D | $p3$ | $p4$ | $p5$ | $p6$ | $p7$ | $p8$ | $p9$ | $p10$ | |
| $D10$ | 64.06 | 82.66 | 83.91 | 95.94 | 99.53 | 100 | 100 | 100 | |
| $D11$ | 83.13 | 86.25 | 91.25 | 98.44 | 100 | 100 | 100 | 100 | |
| $D12$ | 85.31 | 93.44 | 96.88 | 100 | 100 | 100 | 100 | 100 | |
| $D13$ | 89.53 | 95.63 | 99.67 | 100 | 100 | 100 | 100 | 100 | |
| $D14$ | 94.38 | 98.59 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D15$ | 96.72 | 99.69 | 100 | 100 | 100 | 100 | 100 | 100 | |

| coverage within S (%) | | | | | | | | | |
|-------------------------|-------|-------|------|------|------|------|------|-------|--|
| D | $p3$ | $p4$ | $p5$ | $p6$ | $p7$ | $p8$ | $p9$ | $p10$ | |
| $D10$ | 88.59 | 93.75 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D11$ | 94.37 | 96.41 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D12$ | 93.75 | 96.25 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D13$ | 95.00 | 97.81 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D14$ | 96.72 | 98.59 | 100 | 100 | 100 | 100 | 100 | 100 | |
| $D15$ | 96.72 | 99.69 | 100 | 100 | 100 | 100 | 100 | 100 | |

In support of decision-making process, the results show that:

1. 100% coverage is achievable within the range of D for certain p -values.
2. Certain percentage coverage is achievable within D to be complemented with 100% coverage within S for a particular p -facility.
3. 3-facility and 4-facility solutions indicate that few populations will be out of coverage of S for all the service distances $D = 10$ to $D = 15$.
4. The value of D that will ensure total coverage with a fixed number of facilities p can be endogenously determined. For example, instead of fixing S to 15 for the 5-facility solution of Church and ReVelle (1974), the same coverage is achievable with $D = 14$. Alternatively S can be chosen to be closer to D as revealed in a 5-facility solution for $D = 12$ where coverage within D is 90.16% and the remaining 9.84% is within a distance of 12.88 to a health service.
5. Rather than have a percentage of the population covered within 10 and others covered within 15, a service distance can be chosen to cover all the populations.

4.5.2 Comparison With Other Works

The results in Table 4.5 are used to compare existing published research works with the method proposed within this thesis. A desirable solution is one that has as much as possible population covered within D , and the uncovered population within D totally covered within S . The table indicates if the farthest population will travel within or beyond S .

Church and ReVelle (1974) provides an optimal 5-facility solution for MCLP with mandatory closeness constraints of $D = 10$ and $S = 15$. While these parameters were fixed in their work, this thesis has endogenously indicated that the uncovered population can be covered within 14.60 distance units with five facilities (highlighted in red). Table 4.7 compares the optimal solution of Church and ReVelle (1974) with the results for this thesis. This research has produced a better solution with increased population coverage of 5370 (83.91%) within D compared to the optimal solution of 3540 (55.31%) population coverage. The graphic solution for the optimal 5-facility configuration is shown in Figure 4.7 (Church and ReVelle, 1974) and Figure 4.8 shows the Spatial-ABM 5-facility graphic output.

Table 4.7: Five facilities solutions for MCLP with mandatory closeness constants

| | Optimal solution (Church and ReVelle, 1974) | Spatial ABM solution |
|------------------------------|--|------------------------------|
| | facilities on nodes | facilities in discrete space |
| Demand within 10 | 3540 (55.31%) | 5370 (83.91%) |
| Demand outside 10, inside 15 | 2860 (44.69%) | 1030 (16.09%) |
| Demand outside 15 | 0 | 0 |

It is a clear indication that by placing facilities in the centroid of any of the finite set of patches within the boundary without restricting facilities to the nodes, better solutions are obtainable with the Spatial-ABM covering model.

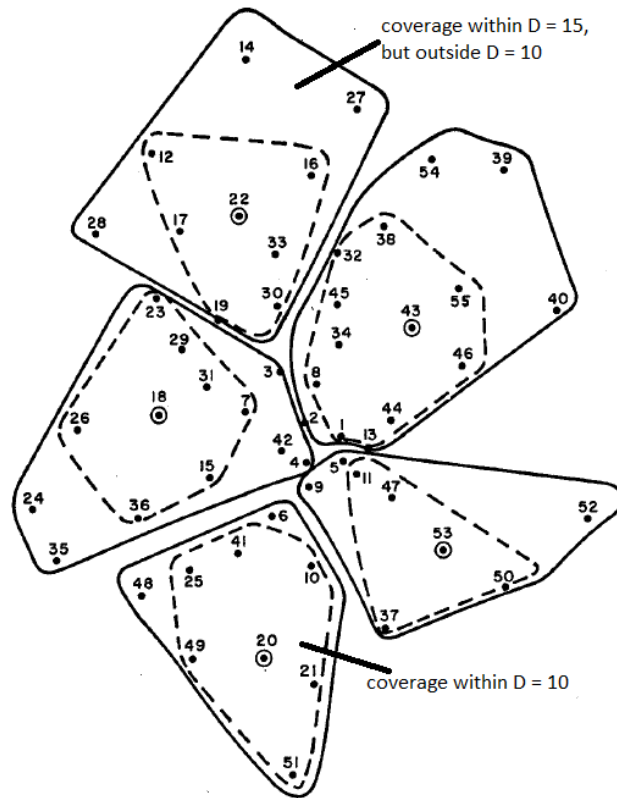


Figure 4.7: Optimal solution for mandatory closeness constraints for 55-node (Adapted from Church and ReVelle (1974))

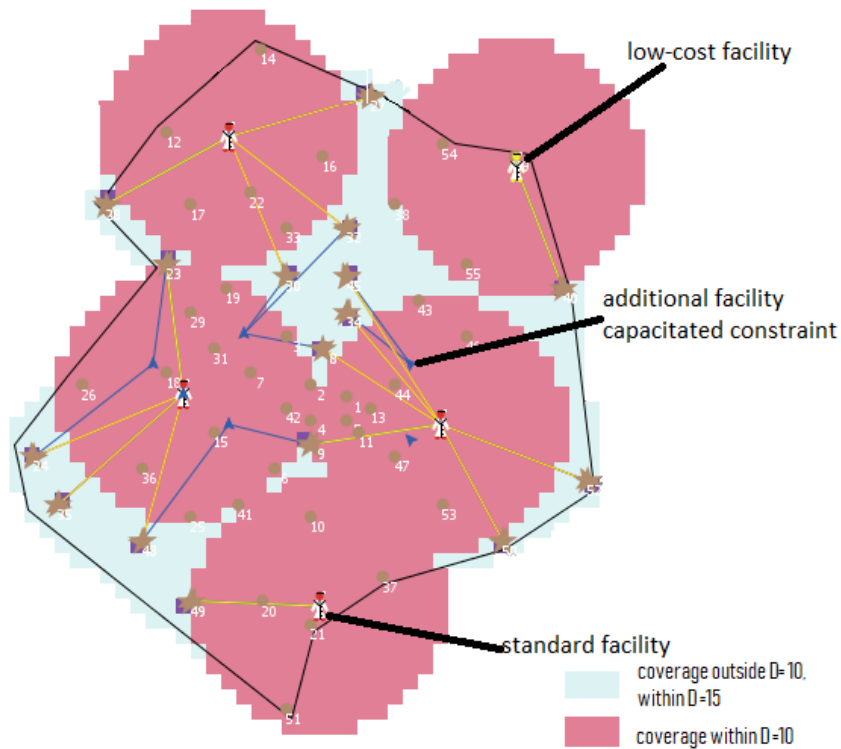


Figure 4.8: Spatial-ABM solution for MCLP with mandatory closeness constraints for 55-node

4.5.2.1 Distance Consideration

The decision maker may prefer some level of equity and not want large populations outside of D . Table 4.5 proves that this strategy can be achieved still using five facilities by increasing D which in turn increases the coverage. For example, with $D = 12$ distance units, 5770 (90.16%) population is covered within D and 630 (9.84%) out of D will travel within 12.88 to get to a facility. Further increase in D leads to an increased coverage up to 14 and 15 distance units that gives 100% coverage within the service distance.

With $D = 14$, the Location Set Covering Problem (LSCP) (discussed in Section 2.4.2) is solved by this model. The LSCP achieves total coverage within a specified covering distance with as few number of facilities as possible. The four-facility solution to the MCLP is an improvement to the LSCP five-facility solution of previous works (Toregas et al., 1971; Church and ReVelle, 1974).

The distance that the farthest population outside D will travel to its closest facility was not considered by previous works (Church, 1984; Galvão and ReVelle, 1996), and so cannot be compared with this model. However, the solutions with total coverage by any of the works will be considered. Table 4.8 compares the 100% coverage solutions with service distance values of 10 and 12 optimal solutions of these previous works. Church (1984) considered locating facilities in an infinite continuous plane, Church (1984); Galvão and ReVelle (1996) provided solutions for facilities restricted to the nodes, while this thesis locates facilities in a gridded discrete plane.

Table 4.8: Comparison of total coverage (%) for facilities restricted to the nodes, facilities placed in continuous plane and Spatial-ABM facilities located at discrete plane.

| D | p | Church and ReV- elle (1974) | Galvão and ReV- elle (1996) | Church (1984) | Spatial-ABM |
|----------------|-----|--------------------------------|--------------------------------|------------------|---------------|
| | | facilities at nodes | | planar | gridded plane |
| | | % | % | % | % |
| 10 | 7 | 98.91 | 98.91 | 100.00 | 99.53 |
| | 8 | 99.69 | 99.69 | 100.00 | 100.00 |
| | 9 | 100.00 | * | 100.00 | 100.00 |
| | 10 | 100.00 | * | | 100.00 |
| | 11 | 100.00 | * | | |
| 12 | 6 | * | 99.69 | * | 100.00 |
| | 7 | * | 100.00 | * | 100.00 |
| | 8 | * | 100.00 | * | 100.00 |
| | 9 | * | | * | 100.00 |
| | 10 | * | | * | 100.00 |
| *: no solution | | | | | |

Table 4.8 shows that this work agrees with previous works and also gives better solutions in some instances. For example, with $D = 10$, eight facilities are required

for total coverage in the planar and Spatial-ABM methods, in contrast to nine facilities required for placing facilities on nodes. The Spatial-ABM methods provides greater coverage of 99.53% with seven facilities compared to 98.91% coverage for restricting facilities at nodes. However the planar method has 100% coverage. With $D = 12$, six facilities are required for total coverage compared to the optimal *seven*-facility solution of Galvão and ReVelle (1996).

The results for four-facility and five-facility solutions highlighted in Table 4.5 also show that an increase in population coverage with D using p -facility does not necessarily imply that the uncovered population with D will travel greater distance to their closest facility than when the coverage is reduced. For example, the four-facility solutions for $D = 10$ and $D = 12$ indicates a lower maximum distance with an increase in coverage. On the other hand, the four-facility solutions for $D = 13$ and $D = 15$; and five-facility solutions for $D = 12$ show an increase in maximum distance for increased population coverage.

4.5.2.2 Facility Capacity in Cost of Establishment

The determining factors for total coverage are not only p and S , but the costs of establishing the HCFs based on the size of the population that is covered by each HCF. The cost objective serves a valuable role in healthcare facility coverage. According to the WHO standard (World Health Organization, 1998), facility density is one HCF to 5,000 population or two HCFs to 10,000 population. To introduce this concept to the 55-node dataset, a facility density of 1 to 500 is assumed due to the small values of the population in the dataset. For the 6400 population, the required number of facilities is the round value of $6400/500$. A selected facility may therefore require additional facility or facilities within its catchment if the population within the catchment is greater than the value of $500 + (50\% \text{ of } 500)$. Where such is the case, the selected HCF hatches the additional number of HCFs and they are located within the catchment of that parent HCF. In the application to this research, it is assumed that if a facility has excess population that is less than 50% of the carrying capacity of 500, no additional facility will be needed. A facility is regarded as a low-cost facility if it covers less than 50% of the carrying capacity.

By comparing the cost of establishing fewer standard facilities to the cost of establishing more facilities driven by the number of low-cost facilities, the decision maker may consider establishing more facilities. The $D = 10$ and $S = 15$ five-facility solution with four standard facilities and one low-cost facility is shown in Figure 4.8. Also, the seven-facility solutions offer options such as five standard facilities and two low-cost facilities; four standard and three low-cost facilities; and three standard facilities and four low-cost facilities (Figure 4.9). The cost of establishing these seven facilities may be lower than providing five standard facilities with the five-facility solution. This is a novel and substantial contribution of this thesis. Although Church and ReVelle (1974) and Church (1984) suggest that the choice of placing fewer facilities may be based on

the decreasing marginal difference in coverage as the number of facilities increases. This is also revealed in Figure 4.6 and Table 4.5 where only little incremental coverage results from increasing the number of HCFs. For example, in the solution for $D = 10$ distance units, the decision maker may want to establish seven HCFs that will provide 99.53% coverage instead of eight HCFs that will give 100.00% coverage. However this thesis has also shown that more facilities may be established at a lower level of expenditure, based on the cost that is driven by the population size within a HCF catchment.

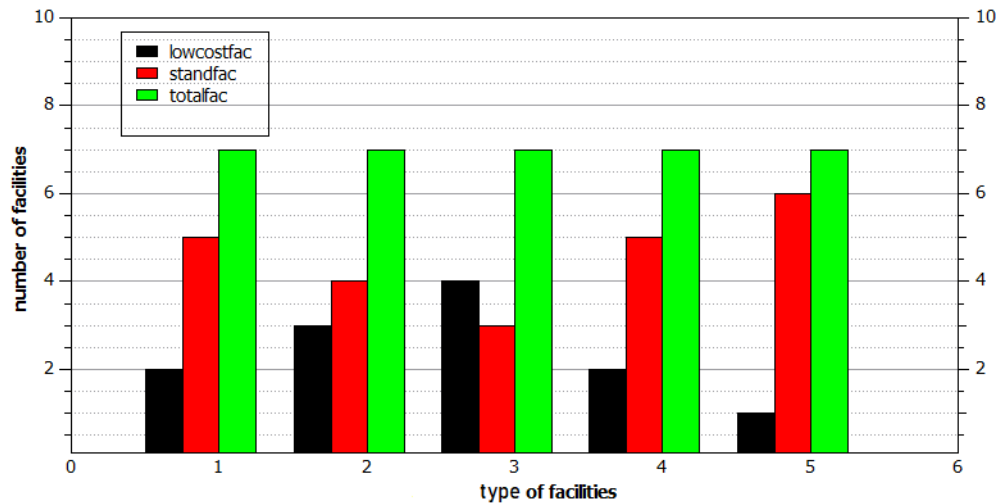


Figure 4.9: Alternative 7-facility solutions for 55-node dataset

4.5.3 Point Density With Facility Distribution

A point density analysis to calculate the population per unit area from the population of the nodes on the 55-node dataset was carried out to show how the population spreads over the region, and where the population is concentrated. An overlay analysis of the point density map, node locations and the HCF locations was done to examine the relationship of HCF distribution with population density.

Figure 4.10 shows one of the solutions with seven selected HCFs in the GIS. The overlay analysis shows that if population constraint is introduced into the facility location problem of the 55-node dataset, additional HCF will be required. In this case, eight extra HCFs added to the main seven HCFs selected are located where the population is concentrated. From the overlay analysis, it is confirmed that when added together, the population size of the nodes within the high-density area of the population density map is high and the HCF agents considered this in the choice of their locations.

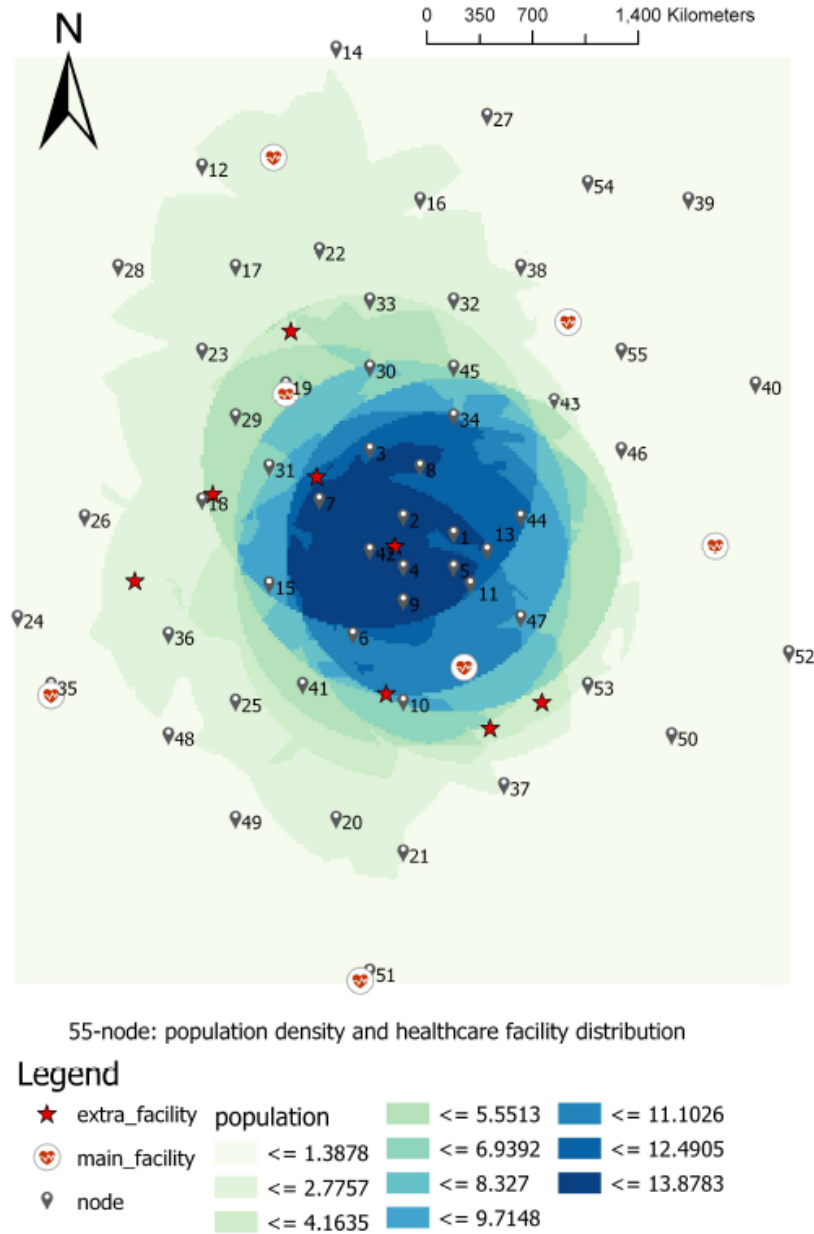


Figure 4.10: capacitated 7-facility solution: population density and facility distribution

4.5.4 Significance of the Model With 55-Node Dataset

Having applied the LNS algorithm to the 55-node dataset, the significant outcomes are as follows:

- The LNS algorithm shows adaptability and flexibility by providing information on the maximum distance that the farthest population will travel to its closest facility, and the range of distance that can provide complete coverage of health services. For the solutions with total coverage within S , the maximum distance that the farthest population outside D will travel to their closest facility indicates

the value of service distance that can provide total coverage. This ranges from 11 - 15 distance units.

- The results show that if 100% coverage is within the desired service distance, D , the maximum distance to the closest facility is not greater than D . However, if the coverage within D is less than 100%, the maximum distance is greater than D but less than the threshold value of S . Also, if the population is not totally covered within S , the maximum distance to the closest facility is greater than S . This shows consistency and reliability of the model.
- The value of mandatory closeness distance, S can be endogenously determined by the model and can be chosen to be close in value to the desired service distance D , while maintaining total coverage of population to be served. Also, non-restriction of the potential facility sites to the nodes and the number of facilities to be selected provides decision making flexibility of different p -facility results within a single model run.
- Agents are capable of interacting with each other and the environment. They obey specified spatial rules such as proximity, containment and adjacency. For example facility agents staying within the specified geographical boundary and also not occupying the same location: additional facilities are within the catchment of the selected facility with excess population; no location has more than one facility.
- Agents can change state. For example indicating a status of standard or low-cost HCF. The type of facilities (either standard or low-cost) depends on the population within the catchment of that facility.

4.6 Summary

This chapter has discussed the different experiments performed on the proposed models and has established robustness and consistency of the models. The verification and validation have revealed the capability of the travel-time and spatial optimization models to produce comparable results. The convergence of the Spatial-ABM optimization model from an initial simulated number of healthcare facilities has helped in deciding the input parameters. The results of the optimization test on a widely used dataset has demonstrated that the model can offer better and more reliable solutions than previous works, including substantial decision making information for improved healthcare coverage.

The experimentation of the models with previous known works reveals satisfactory outcomes and shows that the models are transferable. Therefore the next chapter describes the implementation of the models on Lagos State, Nigeria. To the best of our knowledge, no similar work exists in the study area.

Chapter 5

Implementation and Analysis

5.1 Introduction

The focus of this chapter is to discuss the implementation of the Spatial Agent-based travel-time and Large Scale Neighbourhood Search (LNS) meta-heuristic algorithms for optimizing healthcare facility locations on Lagos State in Nigeria, using the processes described in Chapter 3 of this thesis. This study uses Lagos State as a case study in order to examine and analyse the coverage of the existing healthcare services, reveal the uncovered population, and suggest locations for new public healthcare facilities (HCFs) and assign populations to the facilities. Lagos State is considered a robust analytical environment due to its heterogeneity, especially in spatial characteristics: population distribution across the state; rural and urban settlement; land-use and land-cover distribution that is characterised with swamps, rivers, lagoon and socio-economic uses; and social and cultural diversity. Although the smallest in size in Nigeria, Lagos State has the highest population. Different from the Swain 55-node dataset, the state has a large number of existing public HCFs and vast in other ancillary data that can further enhance the potential capabilities of the models. The irregular distribution of the population and road structure in the state also reveals some of the complexities of the system.

Each of the state's 20 Local Government Areas (LGAs) has different spatio-temporal characteristics in terms of population size, population distribution, land-use/land-cover and geometric area. For the effective exploration of the varying characteristics on the outcomes, a disaggregated analytical approach was employed. It is important that places with heterogeneous nature should not be represented as aggregated entity so that any localised problem will not be generalised. Agent-Based Model (ABM) offers a disaggregated modelling advantage that makes it possible to explore individual element, such as understanding how the structure of the road can influence a travel pattern.

5.2 Healthcare in Nigeria

Having approximately 2.7% of the world's population, Nigeria is ranked the seventh most populated country in the world and the most populated in Africa. With an estimated population of over 197 million populations in need of healthcare, Nigeria is yet to deliver adequate healthcare services for all the people. Since gaining independence in 1960, Nigeria has been unable to effectively address the public health challenges that has been responsible for the high mortality and morbidity rates in the country. The high cost for treatment and inability to access healthcare has contributed significantly to the high level of poverty in the lives of the people. The Nigerian health system has over the years been going through different reforms in order to address healthcare coverage and ease the burden of healthcare costs on the citizens, unfortunately, there has been no significant improvement (Aregbeshola and Khan, 2017; The World Bank, 2010). On the average, 2018 under-five mortality rate is estimated at 100 per 1,000 live births against the Sustainable Development Goals (SDGs) global target of 25 per 1,000 live births, and infant mortality rate is estimated at 65 per 1,000 live births. The 2015 maternal mortality ratio for Nigeria was estimated at 814 per 100,000 live births against the SDGs country target of less than 140 per 100,000 live births (United Nations Children's Fund (UNICEF), 2019).

Healthcare facilities in Nigeria are mainly classified into primary, secondary and tertiary facilities. The primary healthcare facilities (PHCs) are the responsibility of the Local Government in each State of Nigeria. Secondary HCFs and tertiary HCFs are the responsibility of the State and Federal Government respectively. The PHCs are the first level of contact for preventive care and uncomplicated health cases, while the secondary and tertiary HCFs attend to referred and complicated cases. Although the secondary and tertiary HCFs often manage health cases that are expected to be managed at the PHCs. Research shows that out of a total of 34,423 HCFs in Nigeria, the PHCs account for 88% while the secondary facilities are 12% and tertiary facilities are 0.25% (Makinde et al., 2018). Although Makinde et al. (2018) indicate that the government facilities are more than private facilities in the ratio of 67% to 33%, the non-functional public health system has encouraged and boosted the establishment of private hospitals which provide an average of 70% of the healthcare services (Nigerian Health Sector Market Study Report). The Southern part of the country has more private HCFs than public HCFs. On the contrary, there are more public HCFs than private HCFs in the Northern part.

The lack of access to the already over-burdened public HCFs, high costs of receiving care at the private HCFs, and the burden of receiving prompt medical attention at the government HCFs make people to either delay seeking medical care or not seeking care at all (Cremers et al., 2019; Ekpenyong et al., 2019). These also encourage visits to traditional clinics and self-medication (Oladigbolu et al., 2018) which sometimes complicates the health issues of the patients (Varga and Veale, 1997; Abasiubong et al., 2012). Consequently, some pregnant women give birth at home unattended to by med-

ical personnel. According to The World Bank data, pregnant women in Nigeria who gave birth unattended to by skilled health staff increased from 53% in 2011 to 60% in 2018 (The World Bank, 2018).

For over five decades, Nigeria has been undergoing different formative stages in public HCFs strengthening. The varying policies and strategies proposed to address coverage and treatment costs have not received significant uptake and stability due to implementation challenges, political imbalance, mismanagement of funds, corruption and low economic strength. The National Health Insurance Scheme (NHIS) was initially proposed in 1962 but did not take effect until 2005. The scheme is commissioned with the goal of enhancing access to quality and affordable healthcare for all Nigerians through the integration of private sector participation and equitable distribution of health facilities. Selected private hospitals are accredited for NHIS such that patients can receive care close to their homes. Unfortunately, this is yet to be realised as the scheme covers less than 5% of Nigeria population (Olatubi et al., 2018). Most enrollees on the NHIS are the formal private sectors and government employees, some of who are withdrawing from the scheme due to low quality of care and attitudes of private providers towards the NHIS patients. The National Health Policy (NHP) aimed to achieve health for all Nigerians was initially implemented in 1988 with subsequent reviews in 2004 and 2016. The first National Health Act (NHA) was signed into law in 2014 to form a legal framework for the NHP in order to improve access to healthcare and strengthen the National Health System with funds from the Basic Healthcare Provision Fund (BHPF). All the policies and strategies emphasise the importance of PHC which is placed under the National Primary Health Care Development Agency (NPHCDA) – a parastatal of the Federal Ministry of Health (FMH) to ensure the improvement of health for Nigerians for the attainment of Universal Health Coverage (UHC) and health-related SDGs.

The NHA allocates 50% of the BHPF through the NHIS to ensure access to basic minimum package of health services to Nigerians in eligible primary and secondary HCFs, and 45% to the NPHCDA for eligible PHC facility with 20% for provision of essential drugs, vaccines and consumables, 15% for upgrade and maintenance, and 10% for deployment of human resources. The FMH will manage the other 5% for national health emergency and response to epidemics through the National Council on Health. The general government expenditure on health is deemed to be very low at an average of 6% of total annual government expenditure (Hafez, 2018) against at least 15% targets pledged by African Union countries since 2001 (World Health Organization, 2011). The LGAs saddled with the responsibility of the PHCs are the least funded and their eligibility to the funds from the BHPF is a contribution of 25% towards the proposed intervention.

However, these interventions have not considered the impact of proximity to healthcare and expansion of the existing network of public HCFs. While there is much emphasis on equity and even distribution of HCFs, distance between patient and HCF locations, including spatial distribution of population play major roles in realising UHC.

In 2017, the Revitalization of PHC for UHC was initiated by the Federal Government as part of strategy to revitalize over 10,000 out of the existing PHCs. Although this may serve to enhance quality care, the need for more HCFs may be more effective in isolated regions. Spatial accessibility is one of the important keys to health development, utilization, and measures of health outcome. Research has shown that long travel time results in less utilization of HCFs. Spatial accessibility is influenced by good road infrastructures and transport systems. Despite the target of Nigeria to provide UHC for all by 2030, a substantially large population cannot access basic healthcare. Most Nigerians do not have private vehicles (Afolabi et al., 2017; Salau, 2015). Houses and infrastructures, including road paths to dwellings are usually self-provided unregulated which leads to an unstructured road network. Low government participation in the public transport system, including depraved road surfaces that causes traffic jams and damages to vehicles which incur a high cost of maintenance make transport fares expensive and unaffordable to most citizens (Wajuade, 2017). Most rural areas do not have motorable roads, while in the major cities such as Lagos, the transport demands highly exceeds the supply and capacity of the road infrastructure due to rapid development and overpopulation.

With Nigeria's new commitment to achieving UHC, it is important to integrate proximity to HCFs, spatial population distribution, and capacity of HCFs in terms of population size into healthcare policies and strategies in addition to financial and infrastructure indicator of disparity. Different studies have analysed the spatial distribution of the existing HCFs, however they are based on aggregated analysis of facility-population ratio. Inequality in public HCF distribution has been identified in these works with the Northern part of Nigeria having a higher number of public HCFs than the Southern part (Nwakeze and Kandala, 2011; Makinde et al., 2018). However, a more realistic accessibility measure will provide a better-informed decision on the healthcare coverage requirement. Undoubtedly, the UHC will be almost impossible to accomplish if spatial gap is not realistically identified in both HCF distribution and proximity to HCFs, and an appropriate location is determined for new HCFs in compliant with WHO guidelines of distance threshold and proximity to other institutional facilities such as schools. Certainly, the low budget allocation for the healthcare sector and scarce resources are inadequate for the required analyses considering the spatial data and other costly software and technology requirements. Effective management of the limited resources for planning the locations for new HCFs is required for successful coverage intervention as proposed in this thesis. Using a bottom-up approach, the local, state and federal governments can coordinate the healthcare delivery system for the well-being of all Nigerians and appropriate resource allocation in realisation of the SDGs. With the trend of increasing population, diseases and life expectancy in Nigeria, new healthcare facility locations are inevitable. Suitable strategy that will ensure access to prompt and adequate treatment is therefore essential for the entire Nigeria citizens and the economy.

5.3 Spatial Data and Agents

The spatial datasets required for the implementation of the proposed Adaptable Spatial Agent-Based Facility Location (ASABFL) modelling in this thesis were obtained and processed as described in Chapter 3. The spatial data formed the different agent inputs in the models. The data acquisition was highly influenced by non-availability of data from government sources. However, different alternative methods were utilized to obtain data for the implementation of the models as presented in Chapter 3. Such impediments are main characteristics of many developing countries which this research seeks to address so that alternative measures can be employed to support health interventions. Most schools in Lagos are privately owned and religious facilities are numerous. A street of about 2 km length may have up to five schools and religious facilities. Such facilities can also be located within residential houses. Hence, their vast on-line availability that makes it possible to use the geocode tool to determine their locations.

5.3.1 Healthcare Facility Locations

There was no database for geographic locations of HCFs from the relevant government ministries. However, the addresses of the HCFs were obtained and field work was carried out to obtain the geographic locations of the facilities as point objects using the Global Positioning System (GPS). A total of 177 existing public HCFs locations were obtained. Table 5.1 shows the geographic extent of each LGA and its number of existing public HCFs.

5.3.2 Road Network Data

A road map was obtained from the Lagos State government office. After an overlay analysis with the Google Earth image, the road network data was discovered not to be comprehensive in its coverage of the study area as required for this analysis, especially in the rural areas of the state. However the Lagos Metropolitan and the urban regions were fairly covered. Other observations were that most of the road segments were not well connected to support network analysis in Geographic Information Systems (GIS), and there was no speed classification. The available road data was updated on the Google Earth satellite image as described in Chapter 3.

All geographic data formed different map layers and were integrated into the ABM. Figure 5.1 shows the HCFs and water bodies which includes rivers, swamps and lagoon.

Table 5.1: Public Healthcare Facilities in Lagos State by LGA

| Total number of public healthcare facilities: 177 | | |
|---|---------------|-------------------|
| LGA | Number of HCF | Land Area (km sq) |
| Agege | 3 | 11.106 |
| Ajeromi | 5 | 12.230 |
| Alimosho | 23 | 183.627 |
| Amuwo-Odofin | 6 | 133.450 |
| Apapa | 2 | 26.443 |
| Badagry | 4 | 437.111 |
| Etiosa | 10 | 190.894 |
| Epe | 7 | 1178.179 |
| Ibeju-Lekki | 12 | 451.790 |
| Ifako-Ijaye | 9 | 26.391 |
| Ikeja | 14 | 45.784 |
| Ikorodu | 20 | 390.993 |
| Kosofe | 9 | 80.757 |
| Lagos Island | 10 | 8.591 |
| Lagos Mainland | 3 | 19.308 |
| Mushin | 10 | 17.337 |
| Ojo | 3 | 156.764 |
| Oshodi-Isolo | 8 | 44.387 |
| Shomolu | 10 | 11.457 |
| Surulere | 9 | 22.811 |

5.4 Access to Existing Healthcare Facilities

It is unrealistic to measure travel time with the assumption that everyone has access to public or privately owned vehicles. The accessibility measure for this research was therefore based on walking transport mode with a 30 minutes travel-time limit (World Health Organization, 1998). Considering the weakness of a person with illness, and may likely be a pregnant woman or an under-5 child at the time of visits to the HCF, this research assumes the average travel speed of 0.8 m/s (Olowofoyeku et al., 2019) as opposed to the recommended average walking speed of 48m/min or 0.9 m/s (Freiria et al., 2015). While 30 minutes travel-time may translate to several kilometres in the developed countries where high driving speed can be achieved due to good transport infrastructure, it only means a few kilometres in the under-developed countries or remote areas where people typically walk to HCFs. However the travel-time model provides the flexibility of varying the average speed of travelling, travel-time threshold and specifying if analysis should be with straight-line or network journey, which can be analysed simultaneously (Olowofoyeku, 2018). If the straight-line journey path is desired, no road data is required, instead the patient agents would travel in any straight direction. This flexibility is a novel innovation which GIS would not have been able to offer simultaneously.

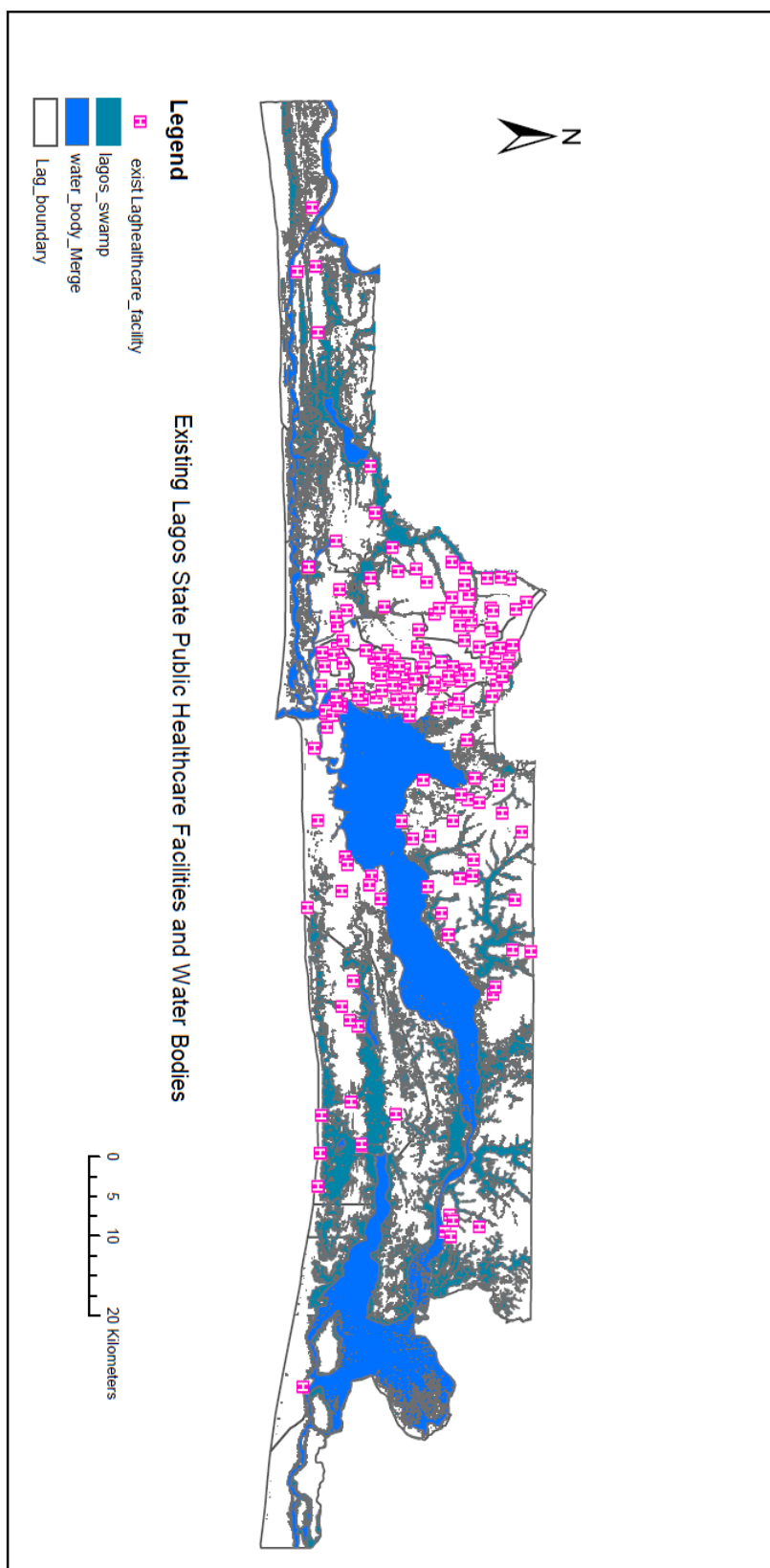


Figure 5.1: Existing Lagos State Public Healthcare Facility and Water Bodies

5.4.1 Catchment Definition With Network Travel-Time Analysis

The catchment area in this application was centred on the HCF location, based on network journey. The travel-time model flow chart is presented in Figure 5.2. The chart shows that at initialization, the user specifies the travel-speed, time-limit and journey mode. For a network analysis, destinations are created as patients' destinations, while for a straight-line analysis, patients only move towards their chosen direction in a straight path until the time-limit is reached.

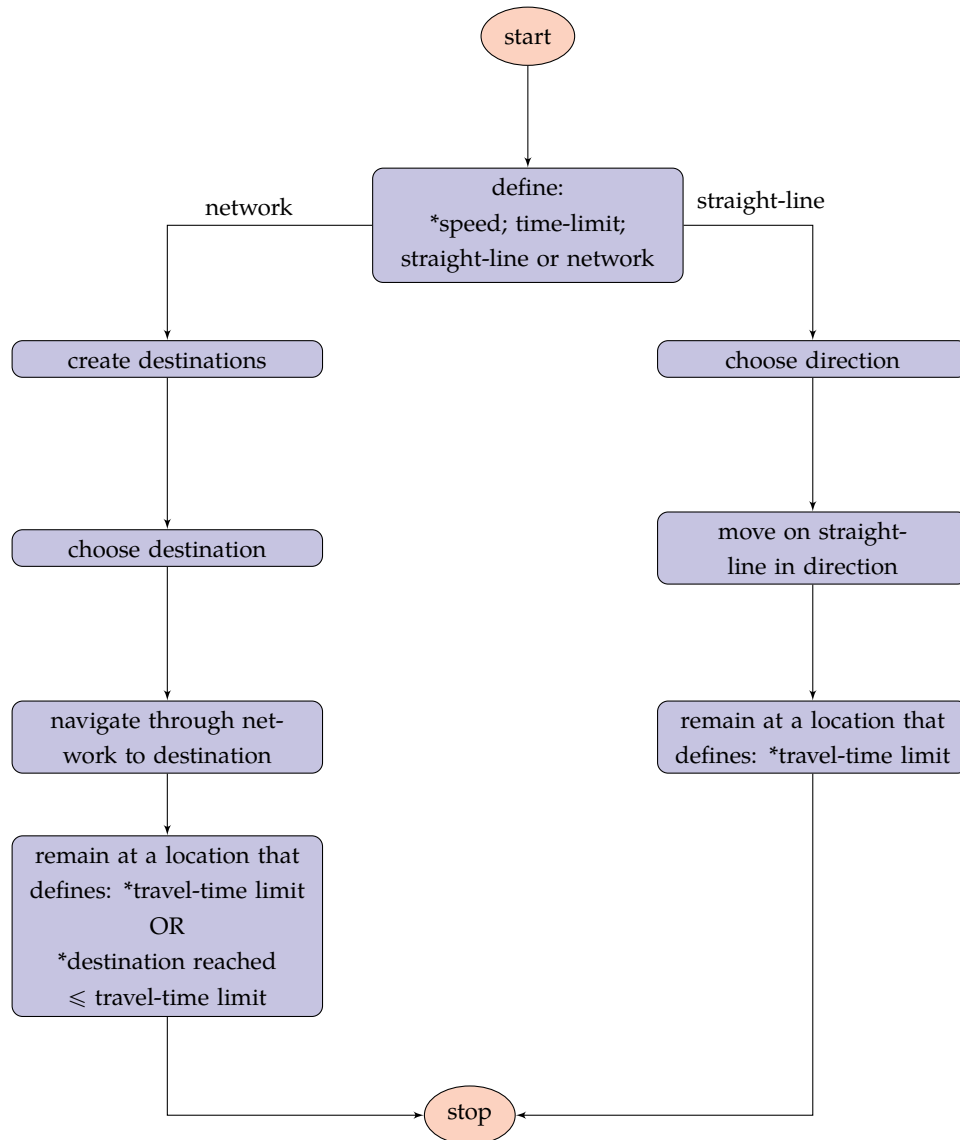


Figure 5.2: Agent Travel Flow Chart

5.4.2 Patient Agents' Travel Characteristics and Time-Limit Compliant

The spatial agent-based travel-time model has the advantage of providing additional information about the characteristics of the road network in relation to the catchment boundary definition. This can further help decision makers to visualize and understand the prevailing situation, unlike when catchment is automatically generated in

GIS. The system was better understood as the patient agents were visualised to only follow road link agents. Where there was no road connecting the destination and the patient's location, the patient agent ceased to move. Several agents can display different human characteristics as they demonstrate independent decisions in the choice of destinations and routes. This agent's behaviour proffers the possibility of identifying areas in need of road infrastructure.

The travel-time analysis confirms that a greater number of travelling patient agents defined HCF catchments within 30 minutes and 31.5 minutes, with no catchment defined by less than 30 minutes travel-time. Figure 5.3 presents the histogram of patient agents' travel-limit in defining HCF catchments in each region.

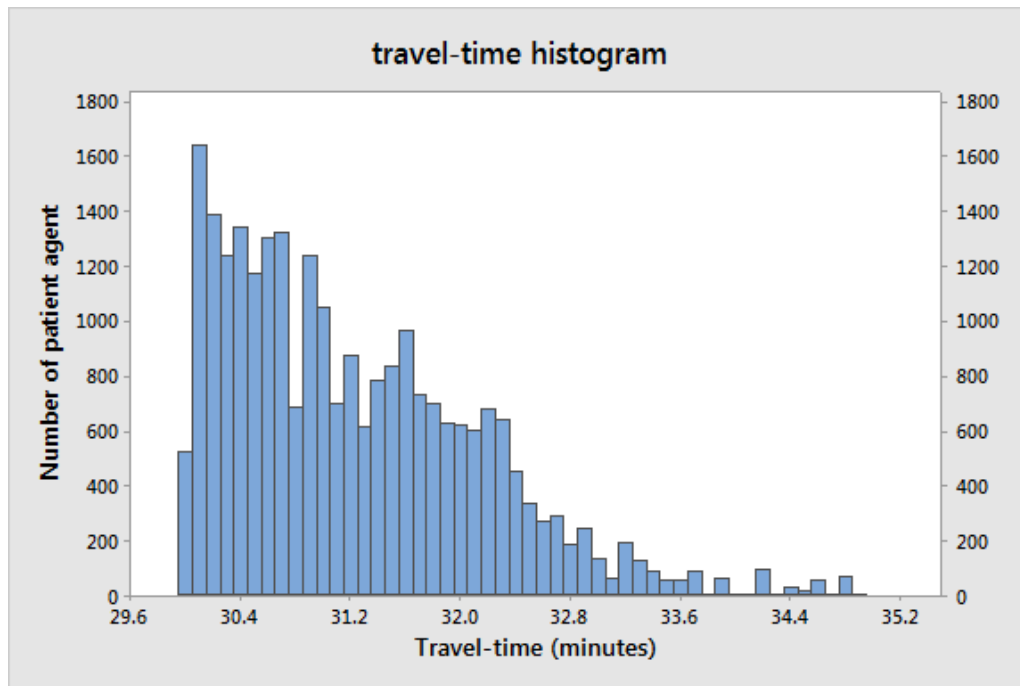


Figure 5.3: Histogram of travel-time of simulated patient agents

5.4.3 One-Sample Equivalent Tests on Travel-Time Threshold

The One-Sample equivalent tests were carried out to test if the resulting travel-time is equivalent to the target value of 30 minutes in all regions. These tests were based on lower equivalence level (LEL) of -5 and upper equivalence level (UEL) of +5 to test if the difference between the mean travel-time in a region and the target 30 minutes travel-time is ≤ -5 minutes or ≥ 5 minutes.

Null hypothesis: Difference ≤ -5 or ≥ 5

Alternative hypothesis: $-5 < \text{Difference} < 5$ α level: 0.05

The results from the statistical analysis reveal that the resulting travel-time estimation is within the threshold time. The difference in the mean of travel-times and

the 30 minutes time-limit in each region is less than 1.5 minutes. The confidence interval (CI) for each region is within the equivalence interval and all P-values are $<10^{-5}$. We therefore reject the null hypothesis and accept the alternative hypothesis for all the regions. Table 5.2 presents the summary of the statistical analyses.

5.4.4 Assessment of Road Dataset

The results of model validation with Google Maps in section 4.3 revealed outliers in certain areas. These extreme cases were looked into for likely causes. The Google Maps Distance Matrix Application Programming Interface (API) had no results returned in some cases due to non-availability of connecting road networks. Similarities were observed in areas with good connectivity as revealed in the validation. Areas with the outliers were observed to be rural regions with unstructured network. A set of results of such outliers from two origins to a number of destinations is presented in Table 5.3.

Figure 5.4 shows sample paths taken by a spatial agent in the proposed Spatial-ABM and Google Maps in the rural community of Ikorodu - one of the regions with outliers. The travel-time in the Spatial-ABM was 30.84 minutes at 48 m/min covering 1,480 m. The Google Maps travel-time on the other hand was 40 minutes for distance coverage of 3,300 m giving a travel speed of 82.5 m/min, which is equivalent to 68.75 minutes travel-time at 48 m/min. The visual analysis revealed that the paths taken in the two models were different. The Google Maps lacked comprehensive road networks as available for this research resulting in differing travel paths and longer travel pattern on the Google Maps. Such wide disparities are likely to have substantial impact on health outcomes. Analysis based on traditional data availability assumptions are therefore unreliable and can have a draw-back on healthcare planning (Olowofoyeku, 2018). This was also confirmed by United States Agency for International Development (USAID) (2016) at the implementation of the World Health Organization (WHO) free and open source AccessMod for physical accessibility measure that integrates Open-Street data (Ray and Ebener, 2008) in other parts of Nigeria.

The non-availability of road data in the study area for on-line service area analysis justifies the adoption of the budget-friendly methodology employed for this research. This is one of the areas where the ABM is effective as it provides a faster, easier and less-costly method of up-dating the spatial data. An example of the on-line road dataset and the updated version in the Badagri LGA is shown in Figure 5.5.

5.4.5 Service Area of Existing Healthcare Facilities

Accessibility to the existing HCFs based on travel-speed of 48m/min and 30 minutes travel-time threshold cannot be regarded as equal. The regions with well structured road networks, especially the metropolitan and urban areas have greater proximity, established by longer distances covered by patient agents within the time threshold.

Table 5.2: One Sample Equivalence Test against 30 minutes travel-time (minutes) limit

| Region | N | Mean | StDev | SE Mean | Null Hypothesis Difference($\leq -5, \geq 5$) | | | |
|------------------------|------|--------|--------|---------|---|----------------|---------------------|-----------------------------|
| | | | | | Diff | CI(-5, 5) | T-value(-5, 5) | P-value(-5, 5) |
| Apapa/Ajeromi_Ifelodun | 974 | 31.301 | 0.8627 | 0.0276 | 1.301 | (0.000, 1.346) | (227.929, -133.810) | ($<10^{-5}$, $<10^{-5}$) |
| Surulere_Mland | 681 | 31.124 | 0.8432 | 0.0323 | 1.124 | (0.000, 1.177) | (189.525, -119.951) | ($<10^{-5}$, $<10^{-5}$) |
| Etiosa | 969 | 31.431 | 0.9076 | 0.0292 | 1.431 | (0.000, 1.479) | (220.553, -122.408) | ($<10^{-5}$, $<10^{-5}$) |
| Alimosho | 4770 | 31.222 | 0.8297 | 0.0120 | 1.222 | (0.000, 1.241) | (517.924, -314.525) | ($<10^{-5}$, $<10^{-5}$) |
| Badagri | 2676 | 31.25 | 1.0644 | 0.0206 | 1.25 | (0.000, 1.284) | (303.762, -182.236) | ($<10^{-5}$, $<10^{-5}$) |
| Oshodi_Mushin | 875 | 31.184 | 0.8898 | 0.0301 | 1.184 | (0.000, 1.234) | (205.576, -126.846) | ($<10^{-5}$, $<10^{-5}$) |
| Shomolu_Kosofe | 4481 | 31.197 | 0.8610 | 0.0129 | 1.197 | (0.000, 1.218) | (481.794, -295.647) | ($<10^{-5}$, $<10^{-5}$) |
| Ikeja | 3009 | 31.168 | 0.8504 | 0.0155 | 1.168 | (0.000, 1.194) | (397.855, -247.160) | ($<10^{-5}$, $<10^{-5}$) |
| Ifako_Agege | 550 | 31.177 | 0.8725 | 0.0372 | 1.177 | (0.000, 1.238) | (166.015, -102.768) | ($<10^{-5}$, $<10^{-5}$) |
| Epe_Lekki | 1963 | 31.307 | 1.2311 | 0.0278 | 1.307 | (0.000, 1.353) | (226.983, -132.904) | ($<10^{-5}$, $<10^{-5}$) |
| Ikorodu | 4214 | 31.192 | 0.8958 | 0.0138 | 1.192 | (0.000, 1.215) | (448.727, -275.933) | ($<10^{-5}$, $<10^{-5}$) |

Table 5.3: Paired ABM and Google travel information: travel mode —walking

| ORIGIN | DESTINATION | | TIME (min) | | | DISTANCE (Km) | | |
|--|---|-----------|------------|---------|---------|---------------|--------|-------|
| | latitude | longitude | ABM | Google | diff | ABM | Google | diff |
| A | 6.487196 | 3.124613 | 32.36 | 682.02 | 649.66 | 1.55 | 32.74 | 31.18 |
| | 6.498837 | 3.106269 | 31.75 | 1408.08 | 1376.33 | 1.52 | 67.59 | 66.06 |
| | 6.494495 | 3.108965 | 30.51 | 732.42 | 701.90 | 1.46 | 35.16 | 33.69 |
| | 6.494658 | 3.13047 | 30.57 | ZERO | — | 1.47 | ZERO | — |
| | 6.499181 | 3.106696 | 31.17 | ZERO | — | 1.50 | ZERO | — |
| | 6.453073 | 3.257856 | 31.57 | 454.15 | 422.57 | 1.52 | 21.80 | 20.28 |
| | 6.456357 | 3.259963 | 30.15 | 456.94 | 426.79 | 1.45 | 21.93 | 20.49 |
| | 6.452874 | 3.25567 | 30.33 | 449.69 | 419.35 | 1.46 | 21.59 | 20.13 |
| | 6.493909 | 3.130991 | 32.49 | ZERO | — | 1.56 | ZERO | — |
| | 6.497221 | 3.130928 | 30.12 | 31.44 | 1.32 | 1.45 | 1.51 | 0.06 |
| B | 6.453073 | 3.257856 | 31.57 | 214.83 | 183.26 | 1.52 | 10.31 | 8.80 |
| | 6.456357 | 3.259963 | 30.15 | 217.63 | 187.48 | 1.45 | 10.45 | 9.00 |
| | 6.46416 | 3.246477 | 32.38 | ZERO | — | 1.55 | ZERO | — |
| | 6.452874 | 3.25567 | 30.33 | 210.38 | 180.04 | 1.46 | 10.10 | 8.64 |
| | 6.457173 | 3.249106 | 30.38 | ZERO | — | 1.46 | ZERO | — |
| | 6.452049 | 3.256048 | 30.66 | 212.83 | 182.17 | 1.47 | 10.22 | 8.74 |
| | 6.470804 | 3.25239 | 30.50 | 32.46 | 1.95 | 1.46 | 1.56 | 0.09 |
| | 6.46707 | 3.24825 | 31.01 | ZERO | — | 1.49 | ZERO | — |
| | 6.454565 | 3.252922 | 31.64 | 199.58 | 167.94 | 1.52 | 9.58 | 8.06 |
| | 6.468425 | 3.250944 | 30.21 | ZERO | — | 1.45 | ZERO | — |
| | 6.459389 | 3.248589 | 30.14 | ZERO | — | 1.45 | ZERO | — |
| | 6.46632 | 3.247958 | 31.52 | ZERO | — | 1.51 | ZERO | — |
| | A: latitude = 6.49785°N, longitude = 3.118662°E | | | | | | | |
| B: latitude = 6.462568°N, longitude = 3.258936°E | | | | | | | | |

The minimum bounding circles (MBC) encompassing the point features of the travel threshold in the GIS defined the catchment of the HCF from which the patient agents travelled. HCFs having access roads with less nodes and longer links have larger service area than roads with several nodes and shorter links. Indicating that unstructured road networks impede accessibility. Examples of such disparity are shown in Figure 5.6 where there is greater accessibility indicated by larger circles in Ikorodu (Figure 5.6a) than there are in the Epe and Ibeju_Lekki (Figure 5.6b) region having smaller circles. The catchment of each HCF is presented in Figure 5.7.

The dissolved MBC features were clipped with the Lagos State boundary map so that the portion that is not intersected by the coverage area indicates no health service coverage. This is the no-coverage layer presented in Figure 5.8.

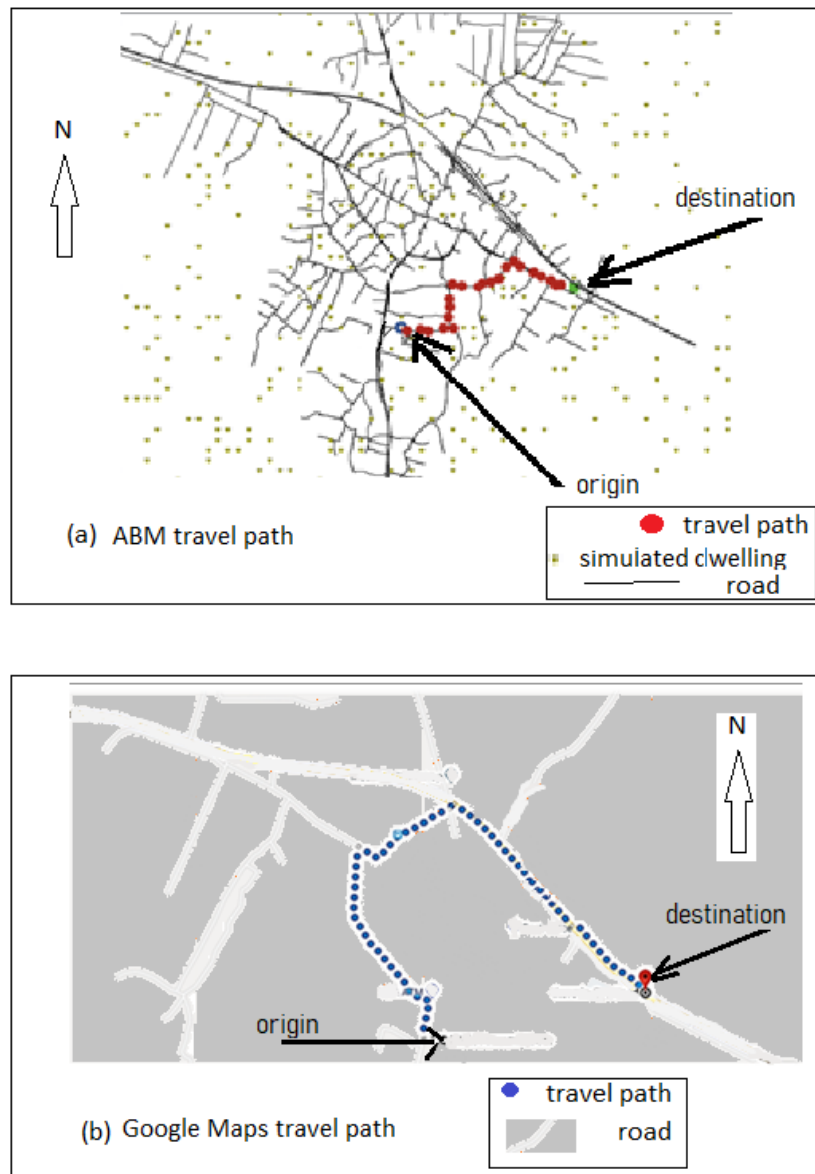
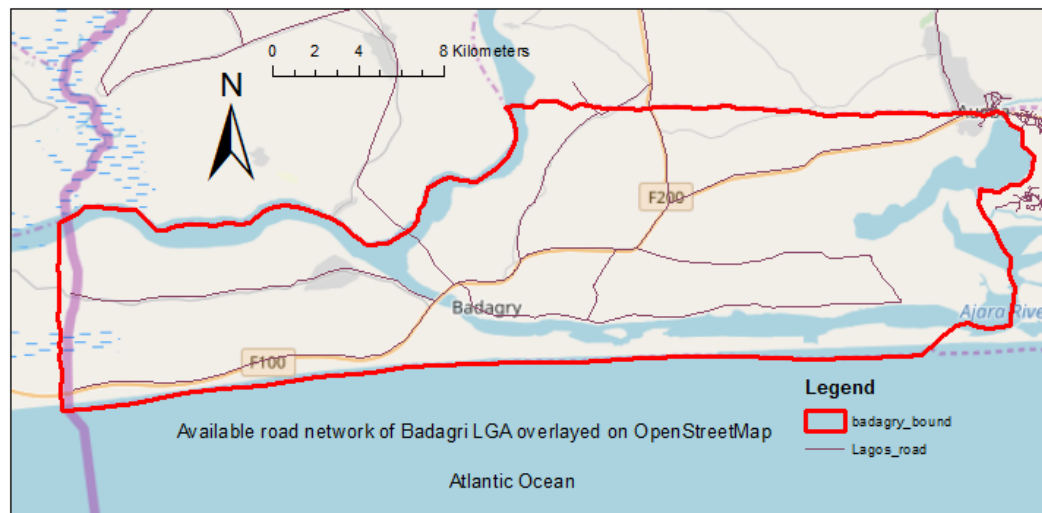
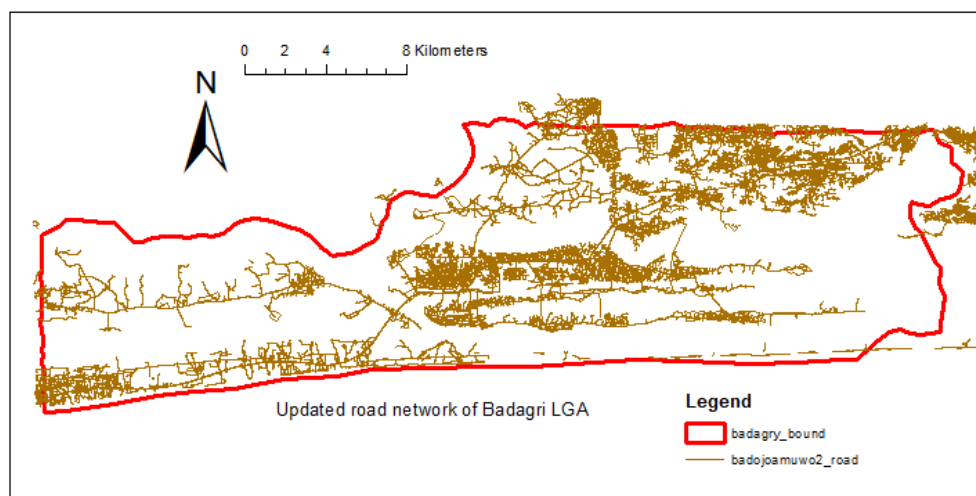


Figure 5.4: Visual comparison of road data and travel paths of agents with Google Maps



(a) Available on-line road network of Badagri LGA



(b) Updated road network of Badagri LGA

Figure 5.5: Sample road dataset update for Badagri LGA

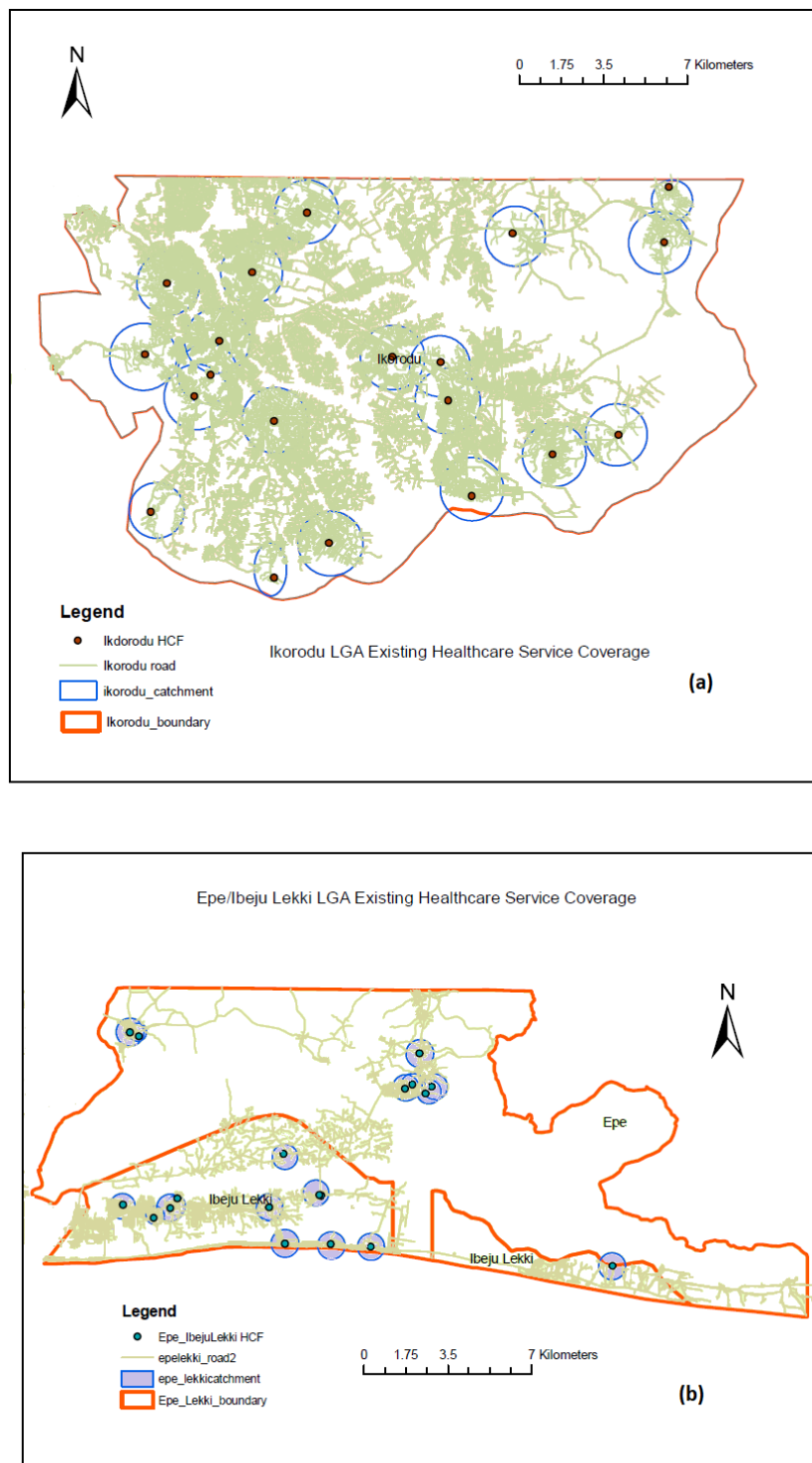


Figure 5.6: Variation in catchment area of healthcare facilities

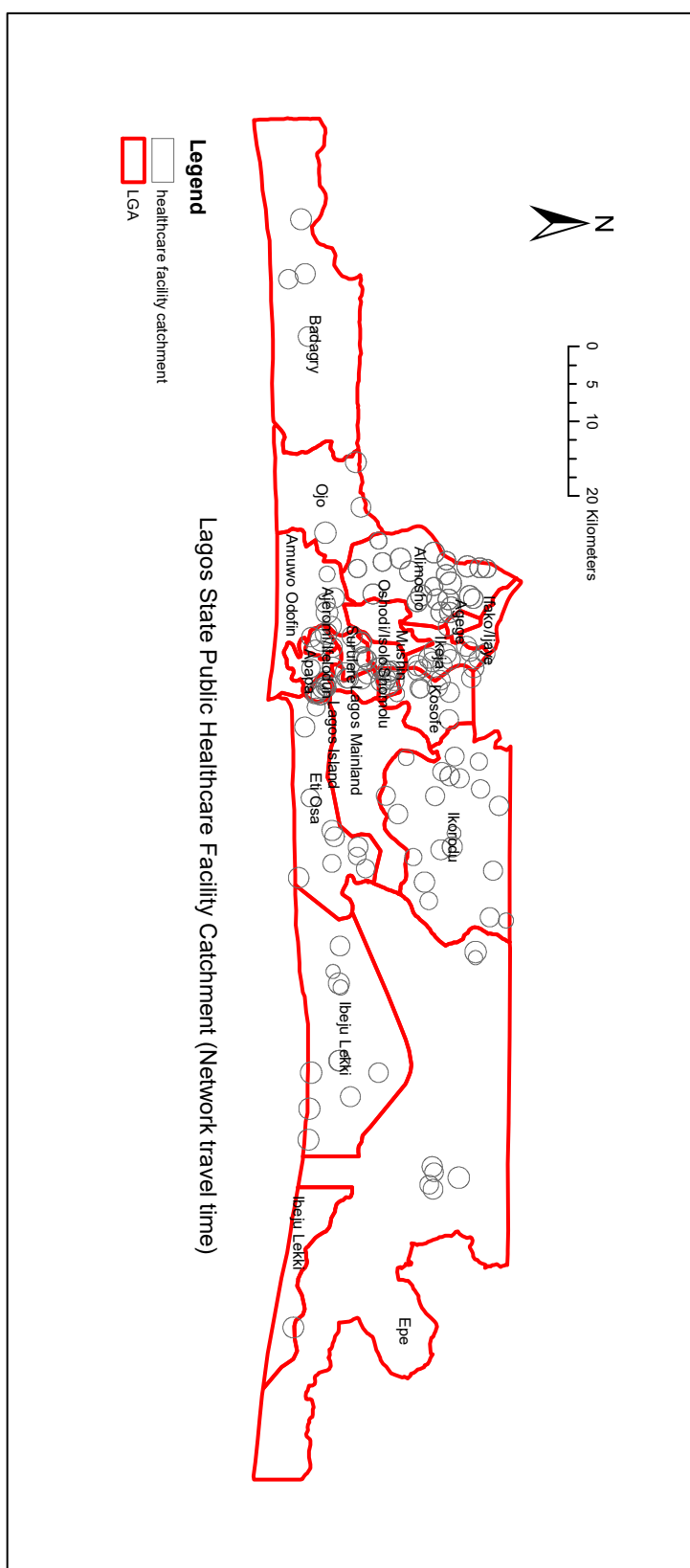


Figure 5.7: Lagos State Public Healthcare Facility Individual Catchment

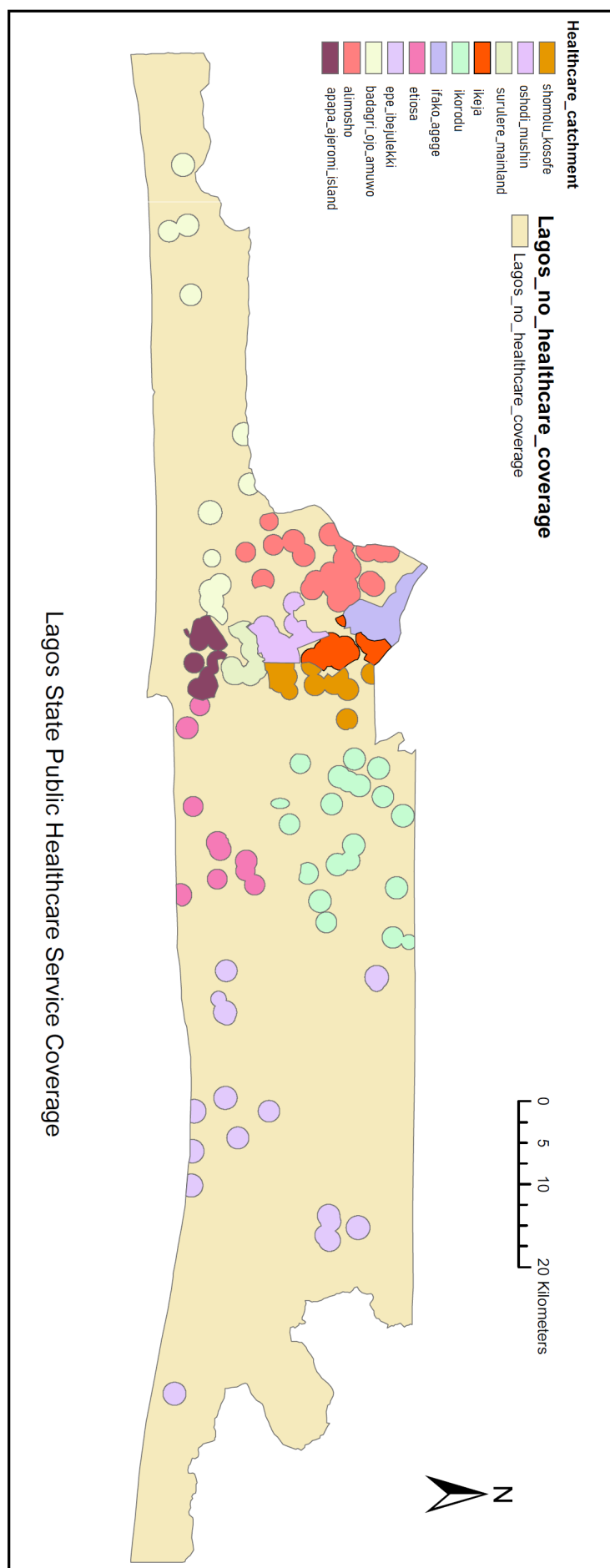


Figure 5.8: Lagos State Public Healthcare Service Coverage

5.5 Location-Allocation Implementation

Having established the areas in Lagos State that are not covered within the existing health service coverage, the next procedure is to fill this gap by determining the number of HCFs that will be required to cover the whole population within reasonable service distance, and finding suitable locations for them. The Large Neighborhood Search (LNS) location-allocation model was implemented to meet these objectives. This two-stage multi-criteria covering location problem formulated as Capacitated Mandatory Closeness Maximal Covering Location Problem (CMC-MCLP) has been described in Chapter 3 of this thesis.

The optimization objectives are:

1. Maximize coverage
2. Minimize the maximum distance that the farthest patient will travel to the closest HCF.

Subject to the following constraints:

1. Total demand allocated to a potential HCF is not more than the specified capacity.
2. Limit the cost of establishing HCFs.
3. Maximum distance between demand and its closest HCF is minimized.

Solving the problems involves applying relaxation to certain constraints:

- Assume average distance radius of the MBCs in the region of focus as catchment radius of the candidate HCF. In this case patient agents will not travel along the road network, instead Euclidean distance radius is used as the service distance based on the average travelling distance obtained from the travel-time analysis for that region.
- The population that is not covered by the Maximal Covering Location Problem (MCLP) solution is attached to its nearest HCF which may be an existing HCF or a new candidate HCF as a p -Centre Location Problem (PCLP) solution. The population's distance to such HCF may therefore be less or greater than the service distance for the MCLP, depending on the nearness of existing HCFs and other additional HCFs within the catchment of a selected HCF.

Decisions for potential HCF locations by the decision maker may be considered subject to the following priorities:

1. More populated communities may be given higher priorities than the less populated communities.

2. The population who live closer to HCF may be given lower priority than those living farther off.
3. The isolated communities may get higher consideration than those within the metropolis.
4. The decision maker can choose to give HCF close to other institutional facilities a higher priority or provide appropriate institutional facilities within the catchment of the HCF.

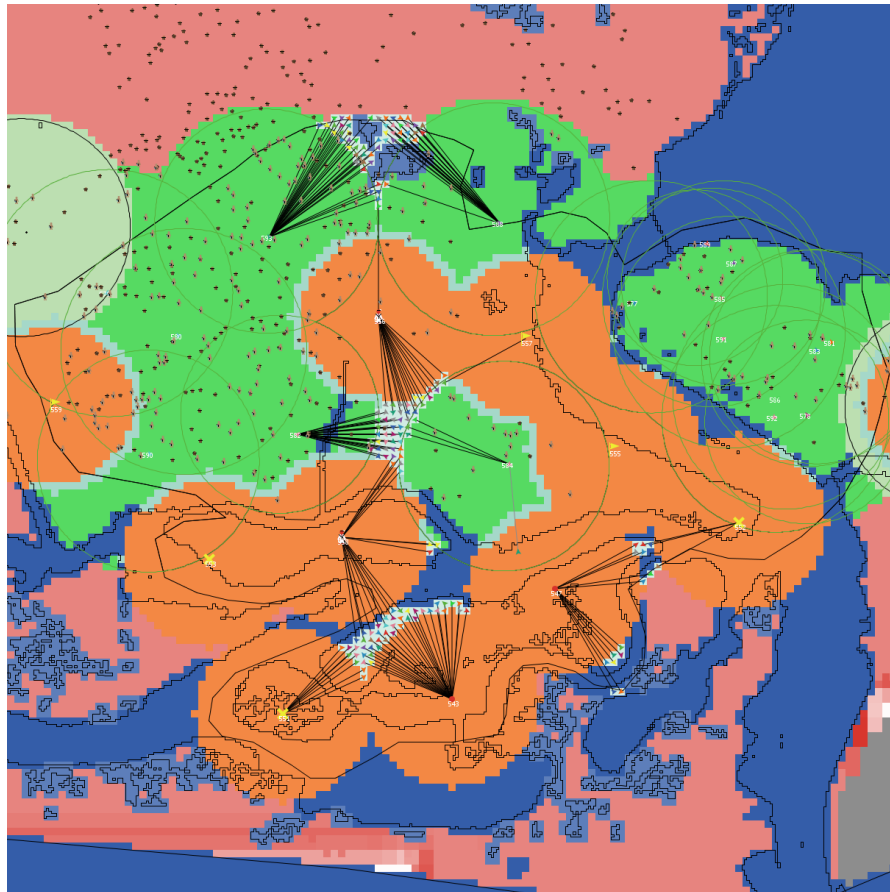
5.6 Optimization Results and Analyses

The results for the location-allocation simulation for placing new HCFs in the uncovered areas identified with the accessibility measure are presented in this section. The facilities generated by the optimization model were categorized into four. The sum of the different types of HCFs in any solution is the total number of new HCFs selected. The different HCFs are:

1. **standard HCFs:** These are HCFs that meet all requirements of proximity to institutional facilities such as schools, religious facilities and markets. They also meet the required population capacity.
2. **standard without utilities HCFs:** These set of HCFs have the required population capacity, but are not close to institutional facilities. They indicate that there is a large population in that area that require standard HCF, however there are no other institutional facilities.
3. **low-cost with utilities HCFs:** These set of HCFs have population below the capacity, but are close to other institutional facilities. They are located in less populated areas and a community health post or clinic is recommended to reduce the cost of establishment.
4. **low-cost without utilities HCFs:** These set of HCFs have population below the capacity, and are not close to other institutional facilities.









An example of a visual output for the Apapa_ajeromi/Island region is shown in Figure 5.9.

Considerations for the decision maker include: covering a larger population, providing additional institutional facilities in a community, establishing standard HCFs or/and low-cost HCFs.







(a) Apapa_Ajeromi/Island healthcare facilities

LEGEND

-  religious_market
-  school
-  uncovered population link to closest HCF
-  water body
-  new HCF catchment
-  existing HCF catchment
-  adjoining region catchment
-  background

HCF

-  standard with utility
-  standard no utility
-  low-cost with utility
-  low-cost no utility

(b) Legend

Figure 5.9: Sample model output view

In the Apapa_ajeromi/Island region solution presented in Figure 5.9, the desired coverage distance derived from the accessibility analysis is 12 m and the required number of facilities based on population is 5. However, from the DSABFL simulation proposed in this thesis, a total number of 10 new HCFs was selected comprising of:

- two standard
- two standard with no utility
- three low-cost with utility
- three low-cost with no utility

This is a clear revelation and as will be explained in later sections, that the number of required HCFs should not be explicitly based on the population size, but requires other considerations such as geographical distribution of demands.

5.7 Decision Support for Facility Locations

Decision examples are provided for the Apapa_Ajeromi/Island and Ikeja regions in this section, using the results of 100 simulations. The summary solution table consists of total HCFs and their frequencies, the maximum and minimum number of each type of HCFs, the maximum and minimum percentage of uncovered population, the maximum and minimum values of maximum distance, S , units between the uncovered population and the closest HCF.

For the LNS meta-heuristic, near optimal solution is sought and many solutions will fulfil the objective functions of the CMC-MCLP with different combination of p -facility and corresponding population coverage within D as feasible solutions. It is possible to select the desired solution sets from:

- A scatter plot of the percentage of population uncovered (UP%) in the MCLP and the maximum service distance, S between the patient and their closest facility.
- The pairs of S and UP% for certain p -facility within the left lower corner of the scatter plot are regarded to have maximal coverage and minimal maximum service distance.

It should be noted that these solutions have different corresponding number of selected HCFs.

We can deduce from this thesis that:

- The cost factor for location decision is not basically the total number of HCFs and population coverage within the desired distance, D or maximum distance, S .

- Other factors can be proximity to institutional facilities such as schools, markets and religious facilities.
- The decision maker may consider solutions that may not require the provision of such institutional facilities as a cost-effective measure.

Sample of decision procedures for the Apapa_Ajeromi/Island and Ikeja regions are presented using:

- Scatter plots of the maximum coverage distance S units and UP% for 100 runs.
- A table and graph of selected solutions for specific p -facility showing the type of HCFs that make up the solution.

On the graphs, the HCF types are represented as:

LCwithutility = low-cost with utility HCF

LCnoutilility = low-cost no utility HCF

stdwithutility = standard with utility HCF

stdnoutilility = standard no utility HCF

5.7.1 Apapa_Ajeromi/Island

The results of the Apapa_ajeromi/Island facility location shows that the p values that can give maximal coverage range between 6 and 12, with 9 to 10 HCFs having the higher frequencies. The summary solution is presented in Table 5.4.

Table 5.4: Summary of solutions for Apapa_Ajeromi/Island

| low-cost facilities | | | | standard facilities | | | | facilities | | uncov pop | | max dist | |
|---------------------|-----|------------|-----|---------------------|-----|--------------|-----|------------|------|-----------|-------|----------|-------|
| with utility | | no utility | | no utility | | with utility | | total | | (%) | | S (m) | |
| min | max | min | max | min | max | min | max | p | freq | min | max | min | max |
| 0 | 2 | 0 | 0 | 2 | 2 | 2 | 4 | 6 | 2 | 31.32 | 32.46 | 20.02 | 22.27 |
| 0 | 3 | 0 | 1 | 2 | 4 | 1 | 4 | 7 | 3 | 17.98 | 34.21 | 17.71 | 22.27 |
| 0 | 3 | 0 | 2 | 1 | 4 | 2 | 5 | 8 | 14 | 12.54 | 27.72 | 16.37 | 23.35 |
| 0 | 4 | 0 | 3 | 1 | 5 | 1 | 6 | 9 | 35 | 8.33 | 24.56 | 15.29 | 27.07 |
| 0 | 4 | 0 | 3 | 1 | 5 | 1 | 7 | 10 | 32 | 4.47 | 18.33 | 15.46 | 22.27 |
| 0 | 5 | 0 | 3 | 1 | 5 | 2 | 7 | 11 | 11 | 2.11 | 17.89 | 13.45 | 22.27 |
| 3 | 4 | 1 | 3 | 2 | 3 | 3 | 4 | 12 | 3 | 2.72 | 4.39 | 14.32 | 15.23 |

5.7.1.1 Paired maximum distance and percentage of uncovered population

Although five HCFs are required from the facility-population ratio, the output from the 100 simulations shows that there is no feasible solution with five HCFs. However, five or less than five standard HCFs were suggested with additional low-cost HCFs. Using the scatter plots of S and UP% for p -facility solutions in Figure 5.10, suggested alternatives are presented in Table 5.5.

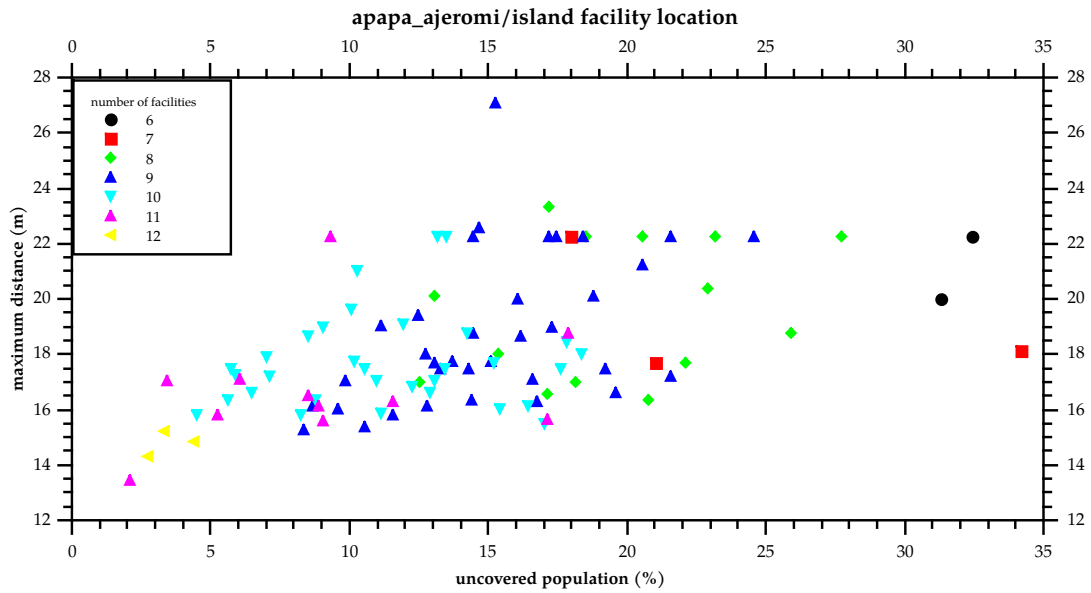


Figure 5.10: Apapa_Ajeromi/Island/Ikeja percentage population uncovered and maximal distance to closest HCF

Solutions with the lower values of p (six and seven) are observed to have the lower population coverage, but not necessarily the higher S values. Suggested solutions were chosen from the left lower locations of the scatter plot for each value of p to ensure that all p -facility solutions were represented.

5.7.1.2 Facility combination alternatives

From Table 5.5:

- The average service distance D desired for healthcare coverage is 12 m which gives 98% maximum coverage, and the minimum value of S is 13.45 m (solution 4). However, this may not be regarded as the best solution since it requires 11 HCFs. Interestingly, seven of the facilities are low-cost, and five are in locations that will require institutional facilities. Although maximum coverage does not indicate a corresponding minimum distance or maximum p value, this solution has the maximum coverage and minimum S value.
- The decision maker may want to avoid establishing a large number of other institutional facilities. The least number of HCFs with no other institutional facilities is two, in the six-facility (solutions 26 and 27), seven-facility (solution 24), and eight-facility (solution 21) solutions. However, for such options, coverage within D is reduced and the difference between D and S is large.
- An equity consideration may choose S to be as close as possible to D .
- All solutions in this area would require other institutional facilities, indicating that the area not only lacks HCFs, but school, religious and market facilities as well.

Table 5.5: Aapa_Ajeromi/Island selected solutions (required $p = 5$; service distance $D = 12$ m)

| solution | low-cost facilities | | standard facilities | | facility | uncov pop | max distance |
|----------|---------------------|------------|---------------------|--------------|----------|-----------|--------------|
| no | with utility | no utility | no utility | with utility | total | % | S (m) |
| 1 | 4 | 1 | 3 | 4 | 12 | 3.33 | 15.23 |
| 2 | 3 | 2 | 3 | 4 | 12 | 2.72 | 14.32 |
| 3 | 4 | 3 | 2 | 3 | 12 | 4.39 | 14.87 |
| 4 | 4 | 3 | 2 | 2 | 11 | 2.11 | 13.45 |
| 5 | 2 | 1 | 3 | 5 | 11 | 3.42 | 17.03 |
| 6 | 4 | 2 | 2 | 3 | 11 | 5.26 | 15.81 |
| 7 | 0 | 0 | 5 | 6 | 11 | 6.05 | 17.09 |
| 8 | 3 | 2 | 1 | 4 | 10 | 4.47 | 15.80 |
| 9 | 2 | 1 | 3 | 4 | 10 | 5.61 | 16.37 |
| 10 | 0 | 0 | 3 | 7 | 10 | 5.70 | 17.49 |
| 11 | 3 | 1 | 2 | 4 | 10 | 5.88 | 17.26 |
| 12 | 3 | 1 | 3 | 3 | 10 | 6.49 | 16.64 |
| 13 | 2 | 3 | 3 | 2 | 10 | 8.77 | 16.37 |
| 14 | 3 | 2 | 2 | 2 | 9 | 8.33 | 15.29 |
| 15 | 3 | 2 | 1 | 3 | 9 | 8.68 | 16.12 |
| 16 | 2 | 2 | 2 | 3 | 9 | 9.56 | 16.03 |
| 17 | 2 | 1 | 3 | 3 | 9 | 9.82 | 17.03 |
| 18 | 3 | 2 | 2 | 2 | 9 | 10.53 | 15.36 |
| 19 | 3 | 0 | 3 | 3 | 9 | 11.58 | 15.80 |
| 20 | 2 | 1 | 3 | 2 | 8 | 12.54 | 17.00 |
| 21 | 2 | 0 | 2 | 4 | 8 | 13.07 | 20.10 |
| 22 | 2 | 1 | 2 | 3 | 8 | 15.35 | 18.03 |
| 23 | 0 | 0 | 4 | 4 | 8 | 17.11 | 16.55 |
| 24 | 1 | 0 | 2 | 4 | 7 | 17.98 | 22.27 |
| 25 | 0 | 0 | 4 | 3 | 7 | 21.05 | 17.71 |
| 26 | 2 | 0 | 2 | 2 | 6 | 31.32 | 20.02 |
| 27 | 0 | 0 | 2 | 4 | 6 | 32.46 | 22.27 |

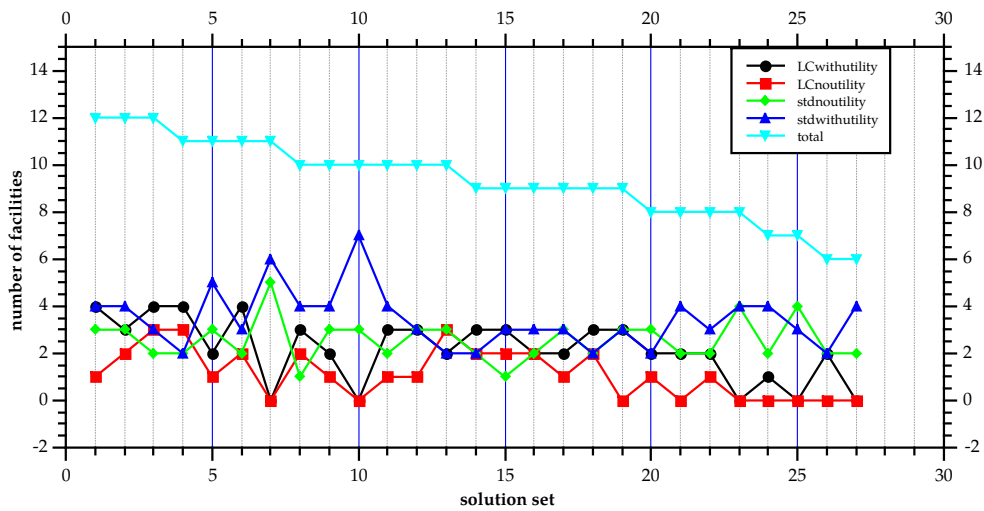
**Figure 5.11:** Set of p -facility in Aapa_Ajeromi/Island

Figure 5.11 and 5.12 show the graphical representations of each p -facility solution. It is shown that:

- All the solutions provide total coverage.
- Each solution in this region requires that at least two HCFs be provided with other institutional facilities, to agree with the WHO (1998) standard.

- A low-cost or standard HCF will require additional cost if other institutional facilities need to be provided within its catchment.

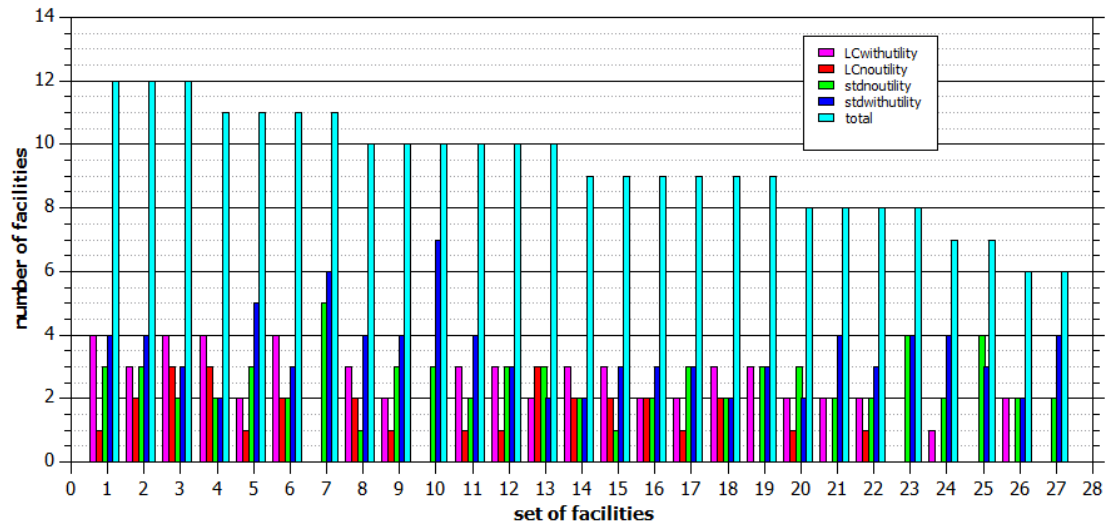


Figure 5.12: p -facility-type in Apapa_Ajeromi/Island

5.7.2 Ikeja

For the Ikeja facility location, the number of HCFs that can give maximal coverage range between four and eight, but four HCFs are required from the facility-population ratio. Table 5.6 gives a summary of the 100 simulation outputs. Six-facility and seven-facility solutions have the higher frequencies.

Table 5.6: Ikeja summary of solutions

| low-cost facilities | | standard facilities | | facility | | uncov pop | | max distance | |
|---------------------|-----|---------------------|-----|----------|------|-----------|---------|--------------|-------|
| with utility | | with utility | | total | | (%) | | S (m) | |
| min | max | min | max | p | freq | min cov | max cov | min | max |
| 2 | 2 | 2 | 2 | 4 | 1 | 40.28 | 40.28 | 17 | 17 |
| 0 | 3 | 2 | 5 | 5 | 16 | 17.1 | 56.21 | 15.57 | 19.92 |
| 0 | 3 | 3 | 6 | 6 | 40 | 0.03 | 37.24 | 13.41 | 18.68 |
| 0 | 3 | 4 | 7 | 7 | 38 | 0.23 | 24.36 | 11.89 | 18.22 |
| 2 | 3 | 5 | 6 | 8 | 5 | 0 | 2.11 | 3.61 | 15.57 |

5.7.2.1 Ikeja paired maximum distance and percentage of uncovered population

Paired maximum distance and percentage of uncovered population is shown in Figure 5.13. These solutions show that more than 99% of the population can be covered with seven or eight HCFs, but four HCFs are required based on population-facility ratio. Although a four-facility solution was found, it only has 60% coverage. The results also show that the maximum covering distance S was less than the service distance

D in some instances. Indicating that the population left uncovered may be closer to an existing HCF or other additional HCF within the catchment of a new facility. The results in this area also show that all HCF locations are in close proximity to other required institutional facilities. The availability of such institutional facilities can be attributed to the fact that Ikeja is an urban area and the capital of Lagos State.

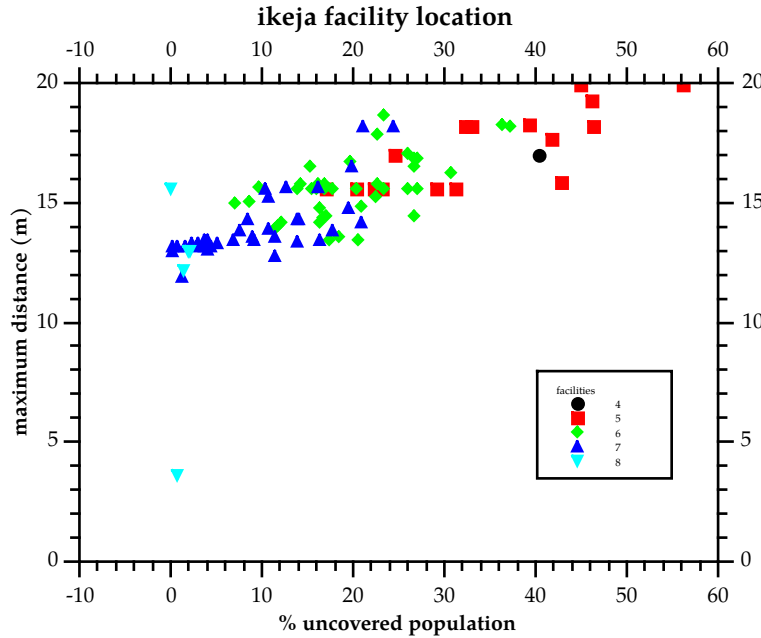


Figure 5.13: Ikeja percentage population uncovered and maximal distance to closest HCF

Table 5.7: Ikeja selected solution (required $p = 4$; service distance $D = 12$ m)

| solution | facilities with utility | | | uncov pop | max distance |
|----------|-------------------------|----------|-------|-----------|--------------|
| no | low-cost | standard | total | % | S (m) |
| 1 | 2 | 6 | 8 | 0.30 | 15.57 |
| 2 | 3 | 5 | 8 | 0.70 | 3.61 |
| 3 | 3 | 5 | 8 | 1.87 | 13.00 |
| 4 | 2 | 6 | 8 | 1.41 | 12.21 |
| 5 | 2 | 5 | 7 | 0.23 | 13.00 |
| 6 | 2 | 5 | 7 | 0.23 | 13.17 |
| 7 | 2 | 5 | 7 | 0.70 | 13.17 |
| 8 | 2 | 5 | 7 | 1.17 | 11.90 |
| 9 | 2 | 5 | 7 | 1.64 | 13.15 |
| 10 | 2 | 4 | 6 | 7.03 | 15.00 |
| 11 | 2 | 4 | 6 | 8.67 | 15.03 |
| 12 | 2 | 4 | 6 | 9.60 | 15.65 |
| 13 | 2 | 4 | 6 | 11.48 | 13.93 |
| 14 | 2 | 3 | 5 | 17.10 | 15.57 |
| 15 | 0 | 0 | 5 | 20.37 | 15.57 |

Four HCFs are required based on facility to population ratio. However, considering other spatial factors, six HCFs and seven HCFs have the higher frequencies.

5.7.2.2 Ikeja facility combination alternatives

Selected solution alternatives are presented in Table 5.7.

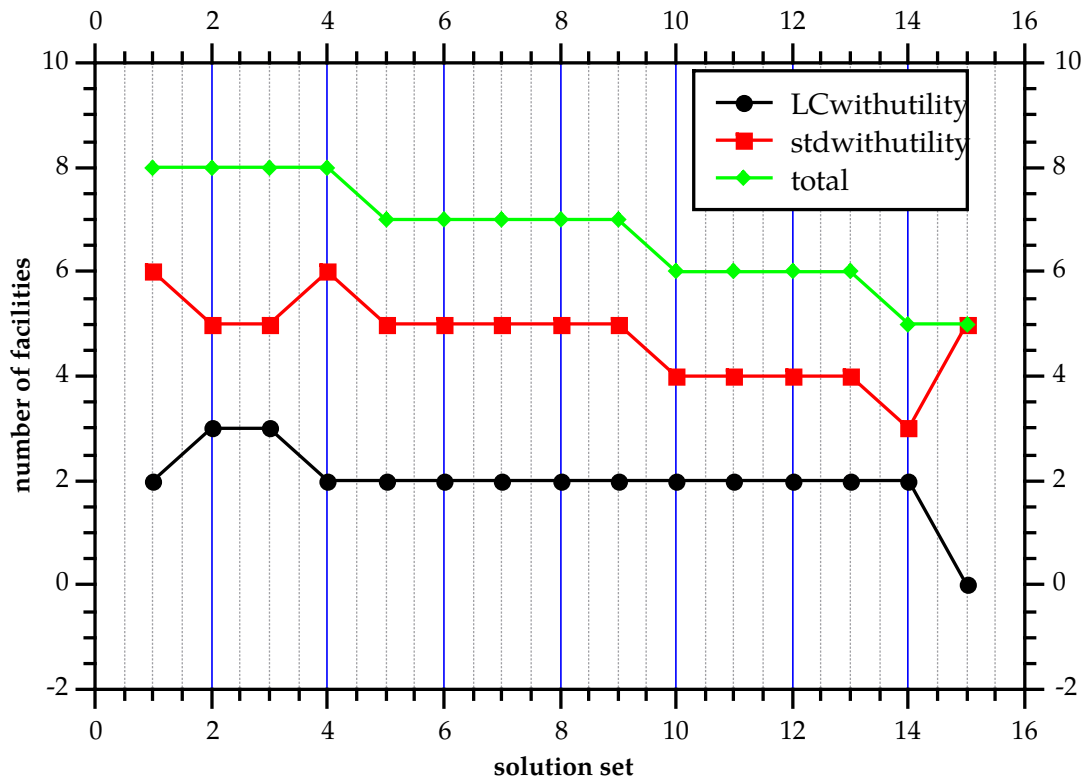


Figure 5.14: p -facility-type in Ikeja

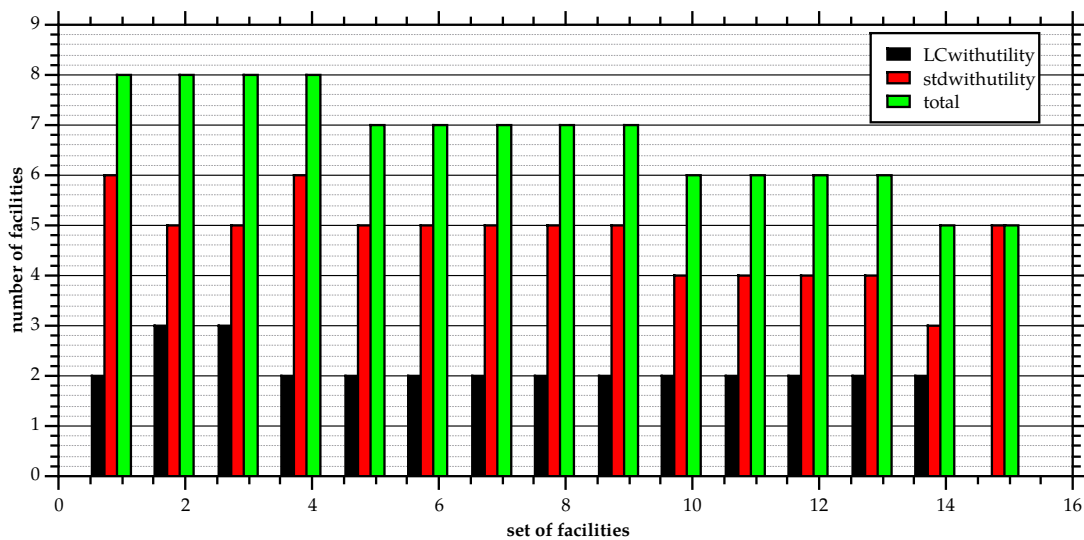


Figure 5.15: Set of solutions for Ikeja

As an alternative to the four-facility solutions, standard four HCFs can be complemented with two low-cost HCFs or 5-facility solution can be adopted comprising of three standard HCFs plus two low-cost HCFs.

5.8 Visualization and Overlay Analysis

For the optimization to have a spatial association, the locations of the new HCFs need to be translated from the ABM coordinate system to real world geographic coordinate system. To do this, the shapefiles of the HCF agents and the link agents connecting patient agents to the closest HCF agents were imported to the GIS. Other spatial features that informed the facility locations were independently overlayed to examine if the simulation is a replication of the real-world.

A visualization analysis of a feasible solution in the Apapa_Ajeromi/Island region is presented in Figure 5.16. The ABM output view is shown in Figure 5.16a and the corresponding GIS map is shown in Figure 5.16b. A visual examination of these two representations confirms that:

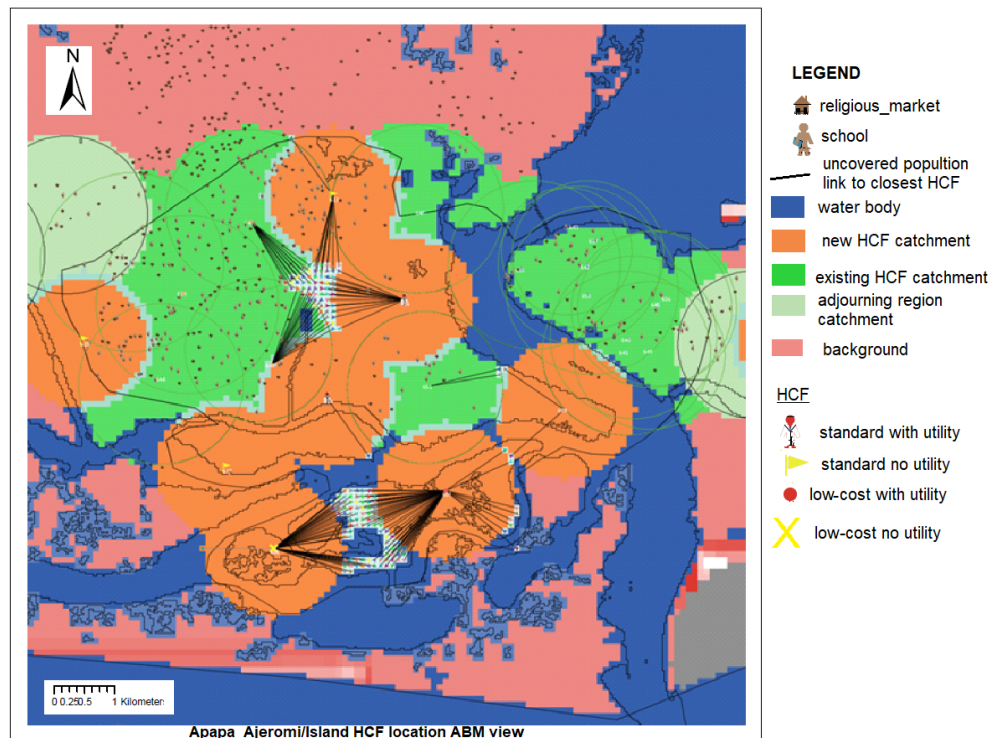
- All the features relate to the same region and the spatial features such as water bodies and schools were accurately simulated in the ABM.
- Uncovered populations were linked to either existing HCF or new HCF location, depending on which is the closest to the population.
- The southern part of the region lacks other amenities in comparison with the northern part.

The GIS map confirms that the objectives and constraints of the CMC-MCLP with LNS algorithm were met. These include:

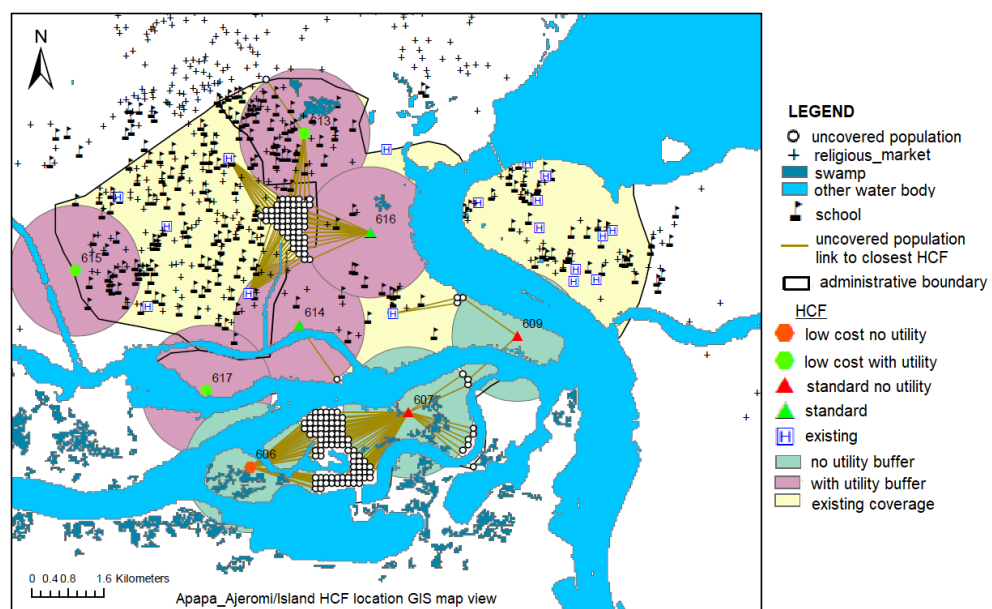
- No new HCF is located on a waterbody.
- All the population left uncovered by the MCLP are linked to the closest HCF.
- No HCF is located outside the boundary of the region.
- New standard HCFs have other amenities within their catchments.
- HCF classified as "no utility" has no other amenities within its catchment.

A further visualization analysis overlayed four different solutions to form the thematic map shown in Figure 5.17 (each solution is differentiated with colors). It is shown that:

- Solutions 1, 3 and 4 have eight new HCFs, while solution 2 has nine.
- Facilities 615, 614, 617 and 618 of solutions 1 to 4 respectively, were observed to be on the same location.
- All four solutions identified the need for HCFs in the southern part, although there are no utilities and the population is low. Hence the consideration for low-cost HCFs.



(a) Apapa_Ajeromi/Island HCF location ABM view



(b) Apapa_Ajeromi/Island HCF location GIS map view

Figure 5.16: ABM and GIS visual analysis of healthcare facility locations

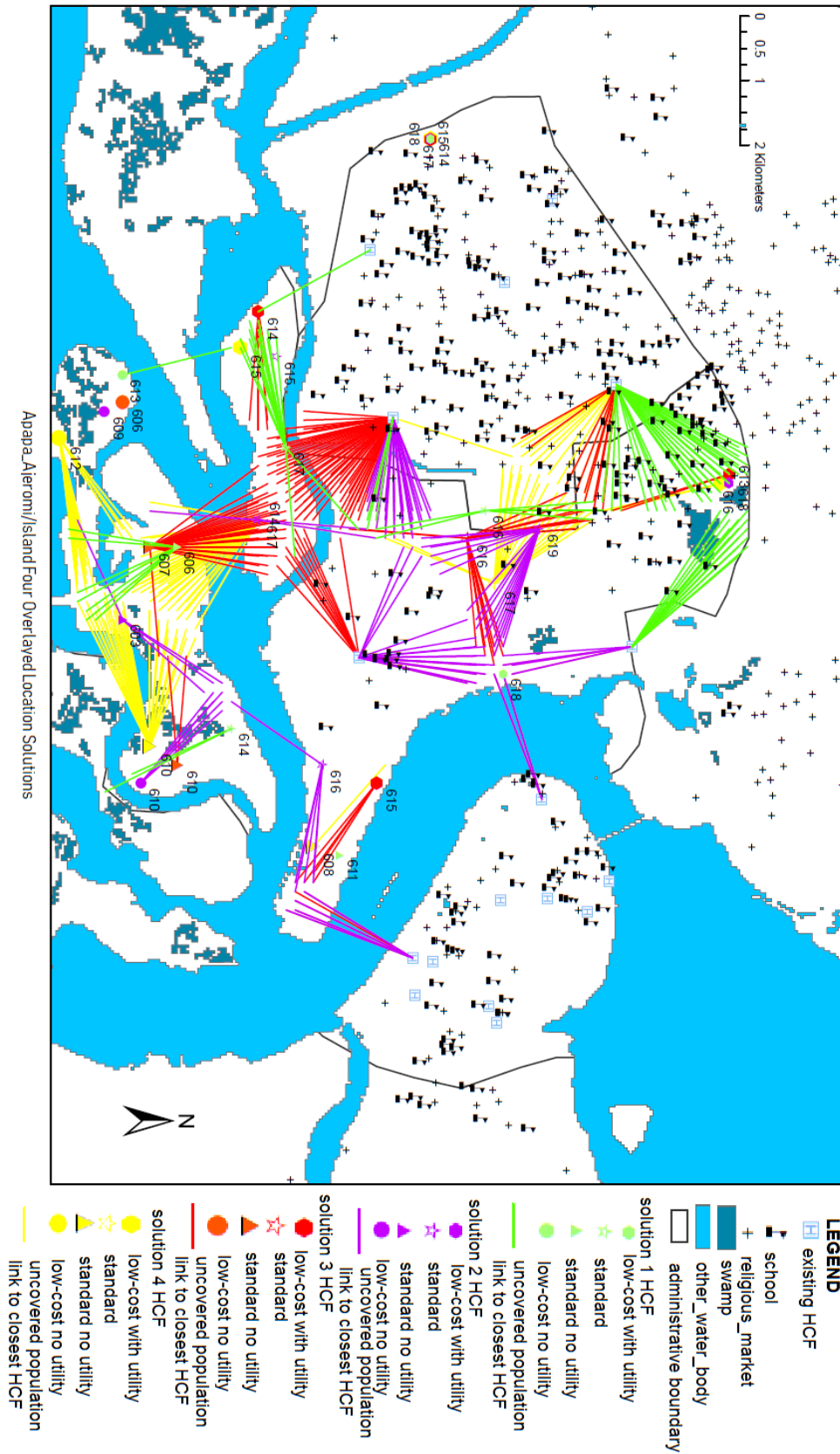


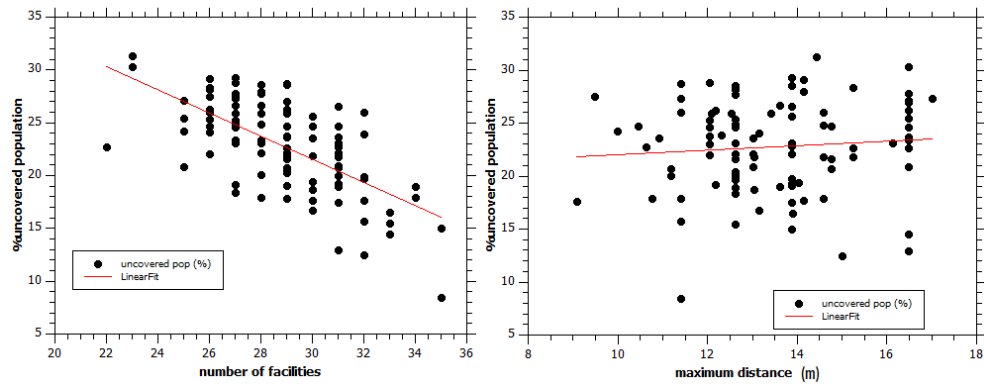
Figure 5.17: Four overlaid feasible solutions to healthcare coverage map

5.9 Non-Spatial Correlation Analysis

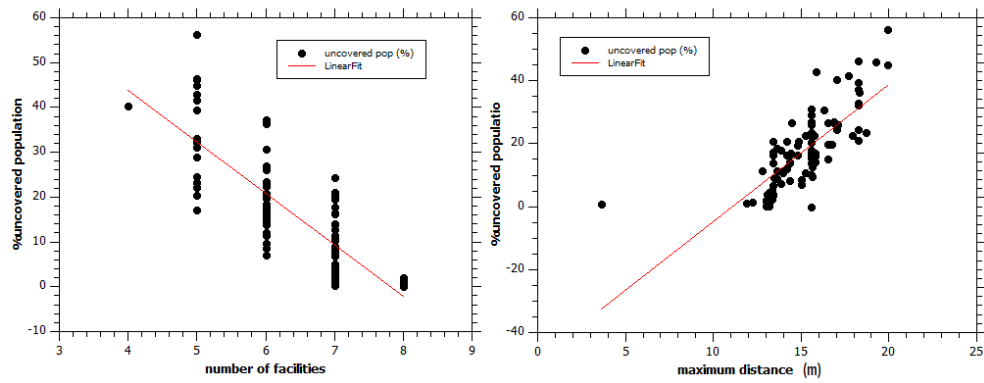
The relationship of the population coverage with the total number of new HCFs and the maximum distance between the farthest uncovered population to its closest HCF was assessed with a non-spatial correlation analysis using 100 model runs. The results show that:

- For every additional unit of HCF, the expected uncovered population decreases. This indicates that on the average, an increased number of HCFs gives a higher percentage of population coverage of healthcare.
- For every additional specified service distance unit, the expected uncovered population (%) decreases.
- Number of HCFs significantly contributed to the coverage ($p < 10^{-5}$).
- Maximum distance on the other hand did not show a significant relationship with the coverage in all regions.

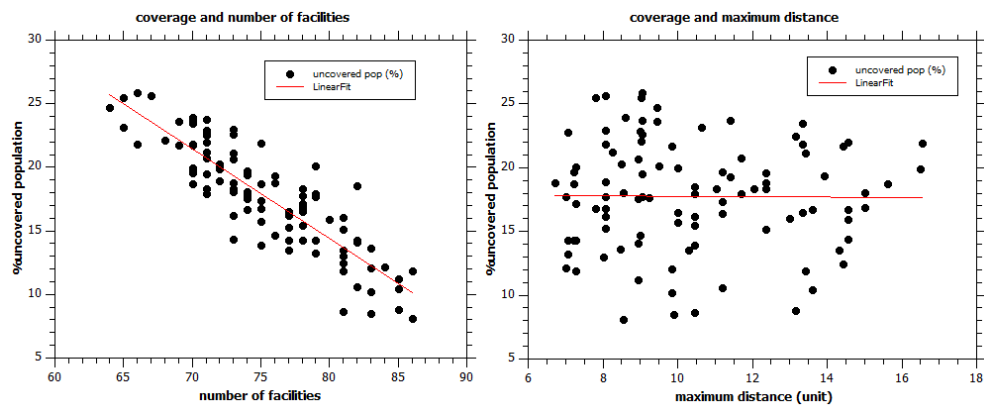
The non-significant relationship between coverage and maximum distance to closest facility can be attributed to the spatial distribution of the existing and newly located HCFs. It is possible that the closest HCF to the population is one of the existing HCFs whose distribution and availability vary in all regions. Figure 5.18 shows the scatter plots and the line fits for North-West zone of Lagos State.



(a) Alimosho region



(b) Ikeja region



(c) Badagri region

Figure 5.18: correlation of uncovered population with number of facilities and maximum service distance

5.10 Spatial Pattern Analyses

Spatial pattern analyses were carried out on the existing HCFs and a set of selected solutions from the Spatial-ABM optimization to assess the adherence of the HCF agents to rules for selecting their locations. These analyses investigate:

1. the spatial distribution of the HCFs
2. the relationship of HCF locations with their neighbouring features

Spatial autocorrelation statistical technique was employed to provide an index for determining the spatial dependence or independence of a feature using the geometric and non-geometric attributes of the feature. Unlike the traditional statistics where independence of observation is anticipated, spatial statistical analyses consider that geographical observations are effected by the geographic space and observations are related to other things. According to Tobler's first law of geography - "everything is related to everything else, but near things are more related than distant things." (Tobler, 1970). The absence of spatial autocorrelation is an indication of a random distribution of the data over the focused area. If random distribution is established, there will be no need for further statistical spatial autocorrelation analyses.

Two types of spatial autocorrelation measures are:

- i **Global autocorrelation:** This aggregates the spatial association with respect to an entire region and does not indicate where the association is most obvious or where there are no associations.
- ii **Local autocorrelation:** This measures the index based on the location of a feature with respect to a specified neighbourhood proximity. If a clustered pattern is identified, the local spatial autocorrelation helps to identify the local features that are strongly responsible for the general spatial pattern in the region. This measure is applied in this analysis considering the differing characteristics of the different region in Lagos State.

Positive autocorrelation indicates that similar data values of features are within the neighbourhood, while **negative autocorrelation** is an indication that the values within the neighbourhood are dissimilar. The spatial statistics that describe such local relationships are called **Local Indicators of Spatial Association (LISA)** among which is the Morans' I proposed by Anselin (1995), and is employed for this thesis using the Open Geoda spatial statistical version 1.12.1.161 software.

The questions that will be answered in the spatial analyses include:

- Do HCFs have a clustered, random or dispersed pattern?
- If there is clustering, what is the significance of the clustering?

- What are some of the contributing factors to the selection of HCF locations?
- Why are there many or less HCFs in some places?
- Is there an association between the features in the region?
- What spatial processes are responsible for the spatial pattern?

The following sections explain the procedure for answering these questions.

5.10.1 Data Preparation

As an initial step for point data, the observations are aggregated within a polygon feature which may be the geographic boundary or a grid of cells drawn to cover the study area. The point feature variables within each polygon feature are then aggregated. For this thesis, the Lagos State boundary map was covered with a grid of polygons so that each polygon grid has variables such as number of HCFs, population size, and area of water body coverage that intersects with the grid. Grid polygons that fell outside the Lagos State boundary and grids with zero values of HCF for the focused variable were removed prior to analysis. The distribution of existing HCFs, the newly located HCFs and a combination of new and existing HCFs were compared.

5.10.2 Exploring HCFs Distribution

The next step is to investigate the location of the datasets for variance and outliers. This was done using the box plot in Geoda. The box plot graphics in Figure 5.19 show that the estimated average number of existing HCFs, new HCFs and the combined sets of HCFs are 2, 5 and 7 respectively (indicated by the green circle); and the median values are 1, 4 and 5 respectively (indicated by the orange line). The datasets for all sets of HCFs reveal slight skewness.

A box map was also created for visual exploration of the datasets and the outliers. The box maps are presented in Figure 5.20. The outliers are also revealed in the maps, including the cells with a high and low number of HCFs. For example, linking the outliers on the box graph to the map shows that the lowest outliers (small number of HCFs within a cell) in the new HCFs data (Figure 5.19b) and the combined new and existing HCFs (Figure 5.19c) are much located on the outskirts and rural areas of the study area. The upper outliers (large number of HCFs within a cell) for the new HCFs (Figure 5.19b) are located towards the south-east and south-west, while the existing HCFs data (Figure 5.20a) has upper outliers around the Lagos metropolis.

For each dataset, most of the values are within the 25% - 75% range. Compared to the newly located HCFs that has most number of HCFs ranging from two to eight, the existing HCFs only has between zero and two values. Upper outliers were revealed in all the datasets, however the existing HCFs have more outliers. These outliers offer valuable information about the distribution of the data. In this case, the number of

HCFs within a grid in each dataset that deviates so much from other values can be investigated further to understand the processes responsible for such high variance in their respective locations.

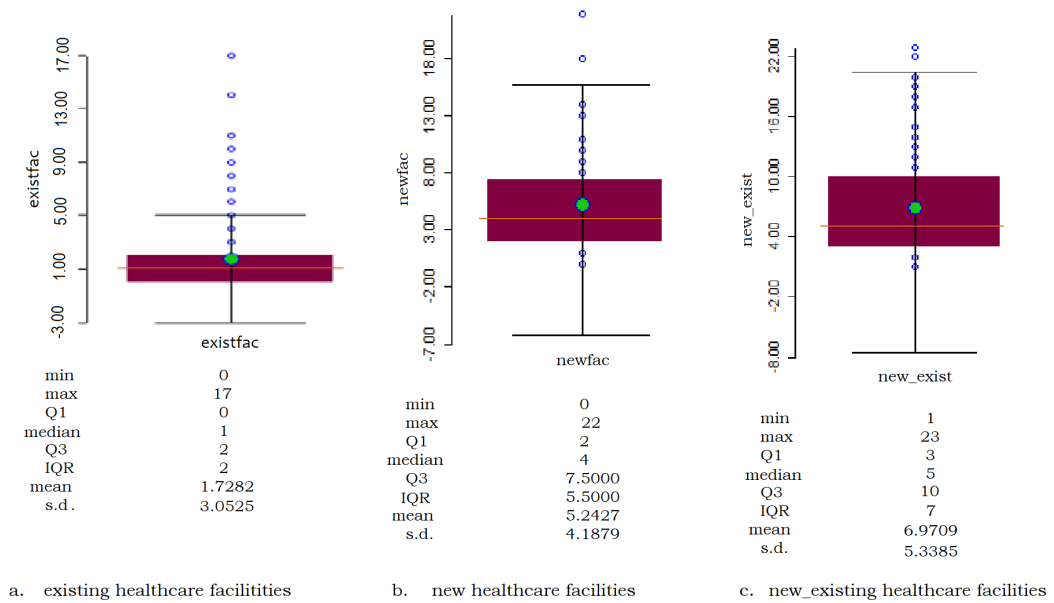


Figure 5.19: Box graphics of healthcare facilities distribution

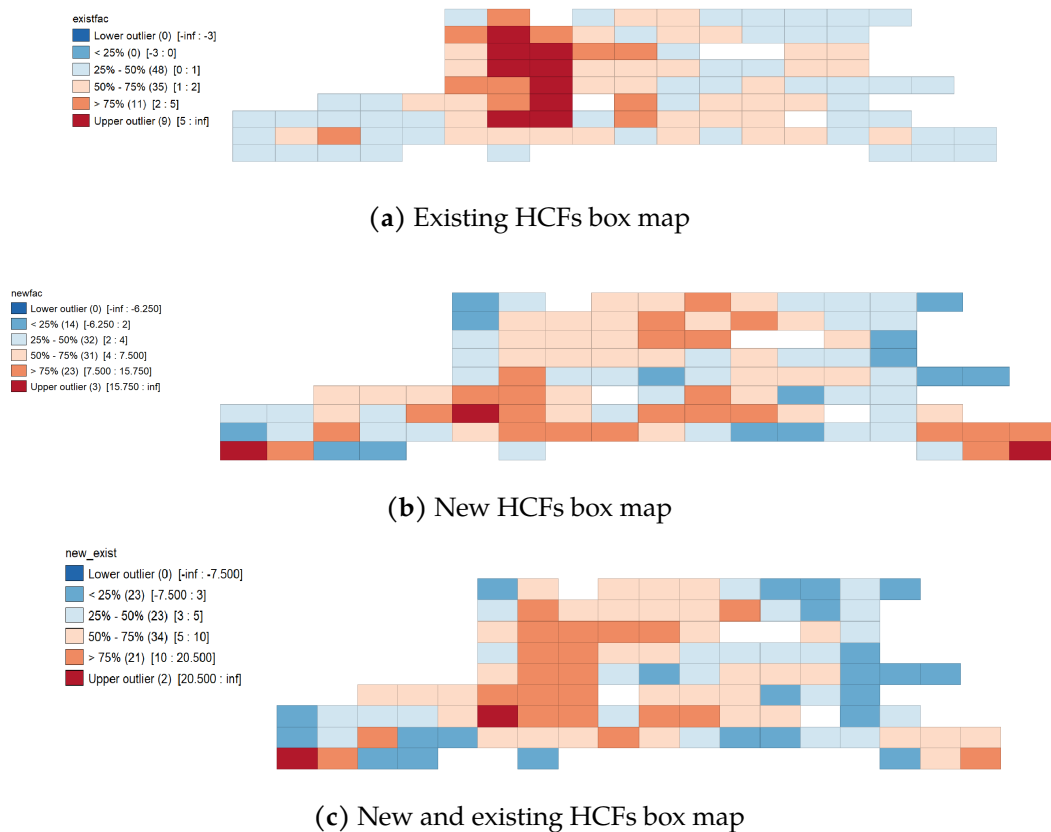


Figure 5.20: Box plot maps for exploratory analysis of aggregated HCFs in grid polygons

A further visual analysis was done by creating 10-category natural breaks maps to classify the number of HCFs in the polygon grids. The maps are provided in Figure 5.21. Figure 5.21a shows that there are few existing HCFs within Lagos State, most of which are within the Lagos metropolis, which consists of nine LGAs: Ikeja - the state capital, Alimosho, Agege, Ifako-Ijaye, Kosofe, Oshodi-Isolo, Mushin, Shomolu and Ojo. The newly located facilities (Figure 5.21b) are more inclined towards the south and the upper north of the state. A combined effect of the HCF distribution is shown in Figure 5.21c which shows higher values and a wider coverage.

The exploratory analyses confirm the existence of clustering and the need for statistical spatial autocorrelation analysis to identify the spatial factors responsible for the clustering and outliers.

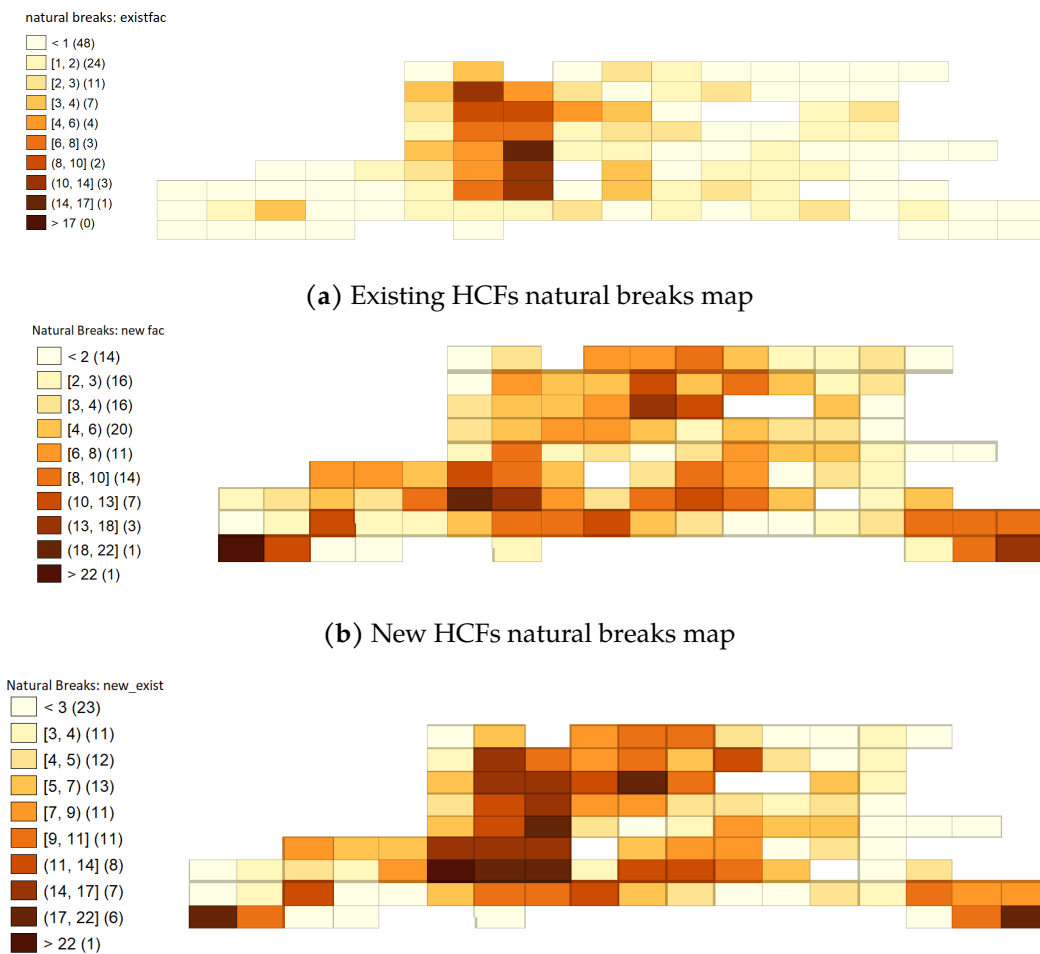


Figure 5.21: Natural breaks of aggregated HCFs in grid polygons

5.11 Spatial Autocorrelation Analysis

This analysis requires that a spatial weight is created. Spatial weights were created with the polygon grids map layer to measure the influence of closer features around each

grid. The analysis helps to know if grids are contiguous and clustering with other grids based on the number of HCFs. A queen contiguity weight that combines both rook-based and bishop-based contiguity was adopted to ensure that all grids have neighbours. As illustrated in Figure 5.22, the rook contiguity shows that cell 5 has cells 2, 4, 6 and 8 as neighbours; in the bishop contiguity, cell 5 has cells 1, 3, 7 and 9 as neighbours; the queen contiguity shows that cell 5 has all neighbours of the rook and bishop contiguity as neighbours. Different spatial weight of four order of contiguity that included lower orders was created for each set of HCFs (new, existing and combination of new and existing), and cells with zero values were removed from the polygon grids layer of each dataset before analysis.

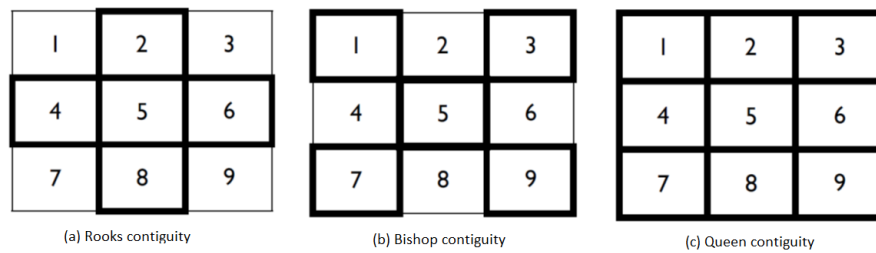


Figure 5.22: Contiguity illustration

The connectivity result for each spatial weight shows that the weight is symmetric and all features have neighbours, except for an isolated grid in the existing set of HCFs. The connectivity histogram of the number of grids for the combined HCFs and the number of neighbours they are connected to is illustrated in Figure 5.23. The minimum number of neighbours is 14 and the maximum is 66 with a total of 103 grids. Mean neighbours is 40.47 and median neighbours is 41. The conceptualization of neighbours was based on these preliminary analyses.

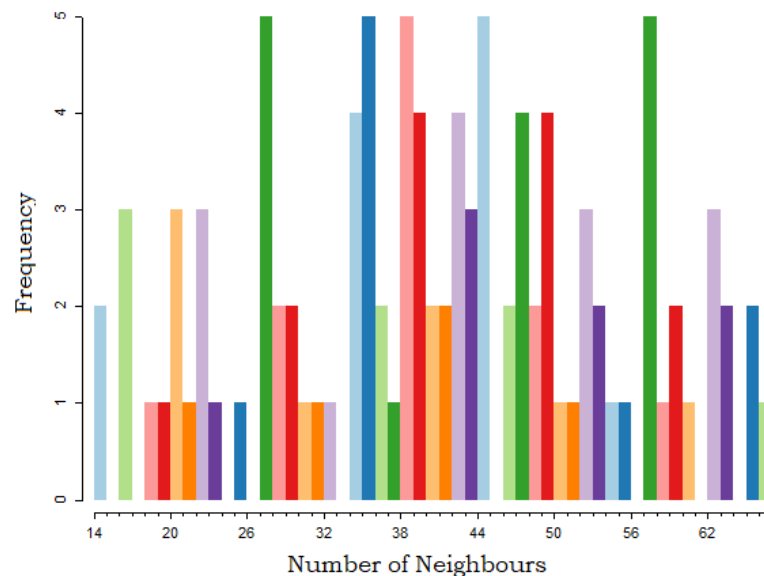


Figure 5.23: Connectivity histogram

For the LISA spatial autocorrelation analyses, three elements are considered: the cluster map, the significance map, and the Moran's index or Moran's I.

1. The Moran's I

The Moran's I is the slope of the regression line of the plot of the spatially lagged variable on the measured variable. The spatially lagged variable is the expected value at each location if the distribution is random. It is the weighted sum or weighted average of neighbouring values for the measured variable.

2. The LISA Significance Map

The LISA significance map shows where there is significant Local Moran statistics results by polygon grid which can be considered to be making significant contribution to the overall or global autocorrelation. Thus, if a significant result is obtained for the whole study area, it is just some grids that actually reveal significant clustering. A legend of significant results and their corresponding p-values with different shades of green is provided with the map. The level of significance obtainable varies according to the number of replications in the randomization procedure. This analysis was performed with 999 permutations.

3. The LISA Cluster Map

The LISA cluster map shows how the attribute that is being analysed clusters. The locations with a significant Local Moran statistics are shown on the map and classified into high-high (HH), low-low (LL), high-low (HL) and low-high (LH) spatial correlation. HH and LL indicate positive spatial autocorrelation (where similar data values or variables are clustered), while HL and LH indicate negative autocorrelation (locations that suggest spatial outliers). The HL values are high values within a low value neighbourhood, while LH values are low data values surrounded by high values.

The spatial clusters shown on the LISA cluster map is the core of the cluster within a neighbourhood where the value, either high or low, is more similar to other values.

5.11.1 Spatial Autocorrelation Analyses Results

Out of the 103 polygon grids that were created covering the study area, one grid has no HCF in the newly created HCFs dataset which is the north western part of Alimosho LGA, however, there are existing HCFs within the grid. For the existing HCFs dataset, 48 grids have no HCF. Results for each dataset are provided as follows:

Existing healthcare facilities

The LISA for existing HCFs is presented in Figure 5.24. Out of the 55 grids covering the existing facilities, 13 are not significant, while the contribution of 16 grids to the clus-

tering is highly significant ($p < 0.001$) (Figure 5.24a). The grids that make significant contribution to the overall autocorrelation are located in the centre and the extreme east of the state.

The LISA cluster map in Figure 5.24b reveals clusters of high number of HCFs and low number of HCFs. The HH clustering is mostly within the Lagos metropolis and adjoining boundaries. The LL clusterings are in parts of Ibeju/Lekki and Epe on the east. Areas with low number of HCFs surrounded by high number (LH clustering) are the outliers located in adjoining cells to the HH clusters: around the outer parts of the metropolis and part of Etiosa region. However, no HL clustering is revealed.

Overall there is positive correlation with Moran's I value 0.065 ($p < 0.05$).

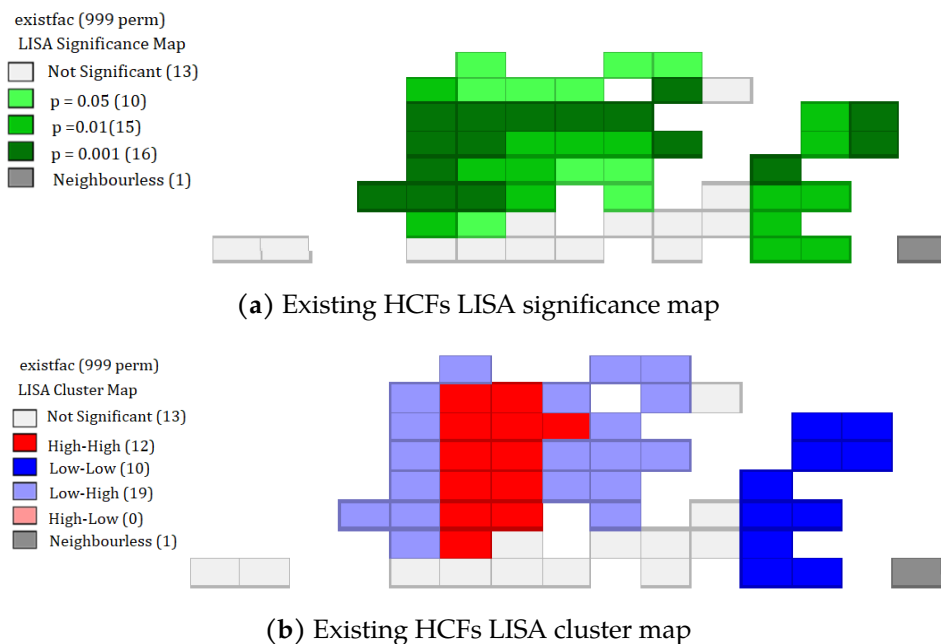


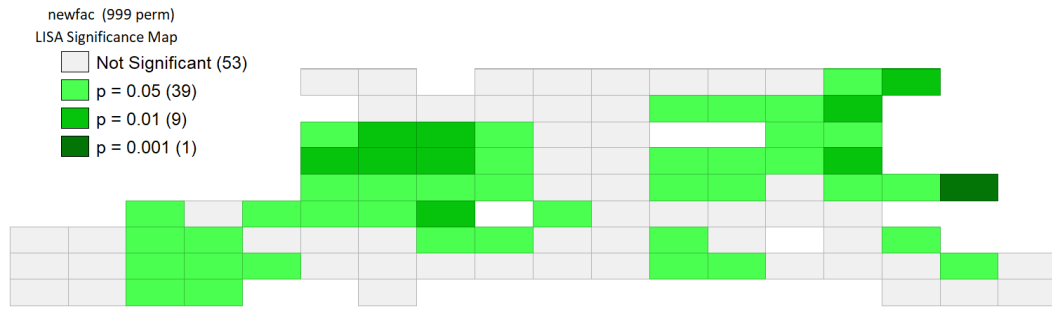
Figure 5.24: LISA for new HCFs

Newly established healthcare facilities

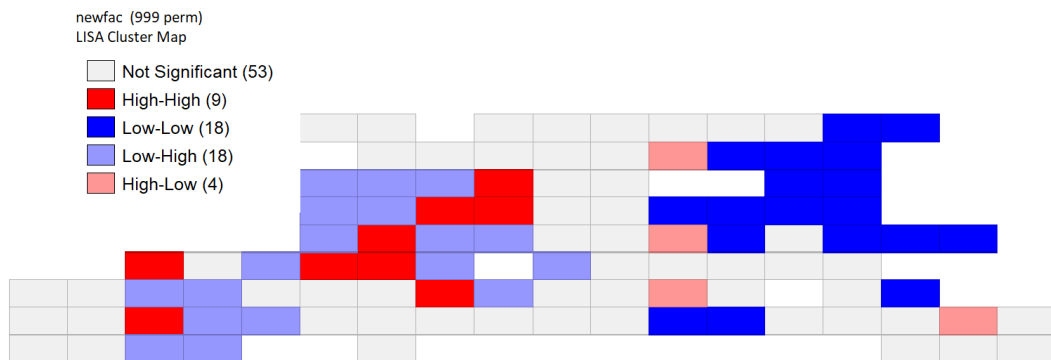
From the newly established HCFs dataset LISA (5.25), 53 out of the 102 grid features are not significant (Figure 5.25a), and only one grid has highly significant ($p < 0.001$) contribution to the clustering (Figure 5.25b). The grids that make significant contribution to the overall autocorrelation span from the west to the east, except the extreme west, the south and some central regions of the state.

From the LISA cluster map, it is revealed that there is statistically significant moderate clustering of number of HCFs. High number of HCFs cluster in parts of the western outskirts and urban central region. Lower number of HCFs cluster in the rural east with relatively lower population. LH outliers are within parts of the Lagos metropolis and the western area of the state, HL outliers bound the LL clusters on the west and south east.

Overall there is positive correlation with Moran's I value 0.011 ($p < 0.05$).



(a) New HCFs LISA significance map



(b) New HCFs LISA cluster map

Figure 5.25: LISA for new HCFs

Existing and new healthcare facilities

From the newly established dataset LISA (5.26), 29 out of the 103 grid features are not significant (Figure 5.26a), and 31 grids highly contribute significantly to the clustering ($p < 0.001$). The grids that make significant contribution to the overall autocorrelation span from the west to the east, except the extreme west, the south and some central regions of the state.

From the LISA cluster map (Figure 5.26b), it is revealed that there is statistically significant moderate clustering of HCFs counts within Lagos State. High numbers of HCFs cluster significantly in parts of the metropolis and other urban regions on the south and central parts of the state. Lower number of HCFs cluster in the rural east. LH outliers adjoin the areas with high clustering, HL are on the extreme south-east and south-west.

Overall there is positive correlation with Moran's I value 0.101 ($p < 0.05$).

combined new and existing HCFs (999 perm)

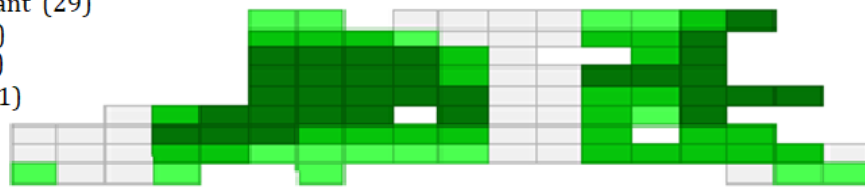
LISA Significance Map

□ Not Significant (29)

■ $p = 0.05$ (16)

■ $p = 0.01$ (27)

■ $p = 0.001$ (31)



(a) New and existing HCFs LISA significance map

combined new and existing HCFs (999 perm)

LISA Cluster Map

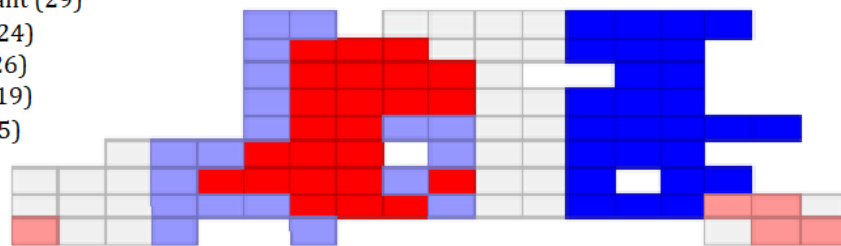
□ Not Significant (29)

■ High-High (24)

■ Low-Low (26)

■ Low-High (19)

■ High-Low (5)



(b) New and existing HCFs LISA cluster map

Figure 5.26: LISA for new and existing HCFs

5.11.1.1 Assessing neighbouring features to HCF

Considering the population map presented in Figure 5.27, it is evident that high population is the contributing factor to the high number of existing HCFs. The regions with low population are however, isolated. While population size is an important factor to be considered in healthcare facility location, the populations' spatial distribution and the road network also impact on access as revealed in the accessibility analyses in Section 5.4. Although large health service area as is the case in the urban areas and the metropolis may indicate higher population within the catchment compared to regions with lower accessibility, hence the need for more facilities. Another geographic characteristic of the location of the HH clustering is that the high population settlement is enclosed in small administrative boundaries. On the other hand the areas with the very low data values have dispersed settlement due to large coverage of water bodies and large boundary extent, particularly in the eastern part of the state. These spatial factors are likely contributors to the low positive clustering. Existing HCFs are largely located in regions with little or no water bodies.

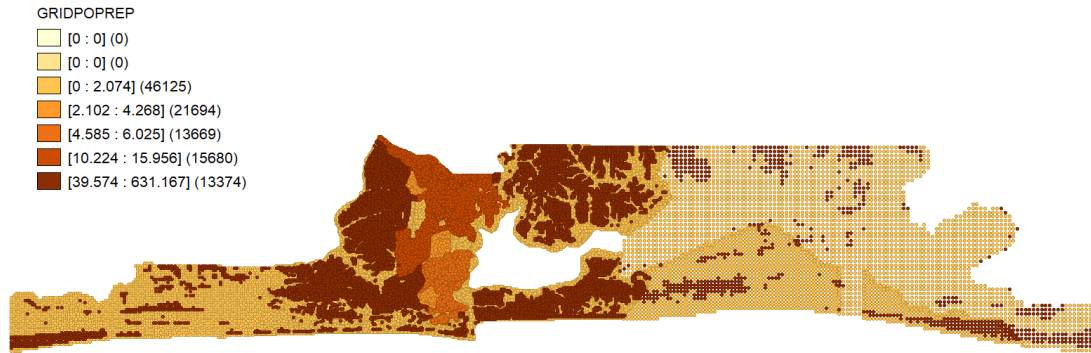


Figure 5.27: Population map

The location-allocation for this research only considered locating new HCFs in areas that are outside the catchments of the existing HCFs. It is also clear that the significant HH clusterings are within the highly populated regions. Not only because of the high population size, but low accessibility as well that is responsible for small catchment area encompassing high population number. The lower facility numbers that seem to cluster in the sparsely populated regions are surrounded with high values. This spatial pattern may be attributed to the population size and distribution, unstructured road networks and water coverage.

The relationship of the HCFs are compared with water and population distribution in the following section.

5.11.1.2 Correlation of HCFs with water and population distribution

The correlation plot matrix in Figure 5.28 and Figure 5.29 compare relationship of the existing and newly located HCFs with water and population distribution. Histograms, scatter plots, and the slope of the regression line of the scatter plot were generated. Observations from these comparisons are highlighted below:

- i The histogram on the diagonal shows the shape of the distribution of each variable in the study area. That is the distribution of HCFs, population size, and area covered by water feature. The histograms reveal that most of the grid features have low percentage of water coverage and population values. However, there are some extremely high values in a few areas. The existing HCFs also have low values in most parts of the state and extremely high values in a few places, unlike the newly established HCFs that have high values in most regions.
- ii The correlation matrix in Figure 5.28 indicates weak positive correlation (0.064) between the exiting HCFs and the new HCFs, while there is a strongly significant and positive relationship (0.891, $p < 0.01$) between the existing HCFs and the population. The relationship with water is negative (-0.114) and non significant. The newly established HCFs have weak positive relationship (0.111) with population

and water (0.068). This can be attributed to the fact that the new HCFs did not only consider the population size, but other geographic factors such as accessibility and distribution of the population.

- iii Interestingly, water coverage and population size have a non significant relationship. As revealed in the scatter plots: where there are high number of population, there is low water coverage, and vice versa. However, this is not to say that people do not dwell in places with water coverage to have excluded such areas from being considered for healthcare coverage, but they are dispersed within the few dry habitable spaces in such regions. Unfortunately, as revealed in the scatter plots, where there is high water coverage, very few HCFs exist. This gap has been filled by the DSAFL model presented in this thesis, where high water coverage does not indicate low number of new HCFs. It is expected that dispersed population will require higher number of facilities. However, cost can be reduced if the size of the population is given consideration in budget allocation, as indicated by the low-cost HCFs suggested in this thesis.
- iv Figure 5.29 shows the overall effect of this intervention. The new and existing facilities combined now have a significant positive relationship with a population correlation value of (0.596, $p < 0.01$). As indicated in the coverage analysis, new HCFs are only considered where there is no coverage of existing healthcare.

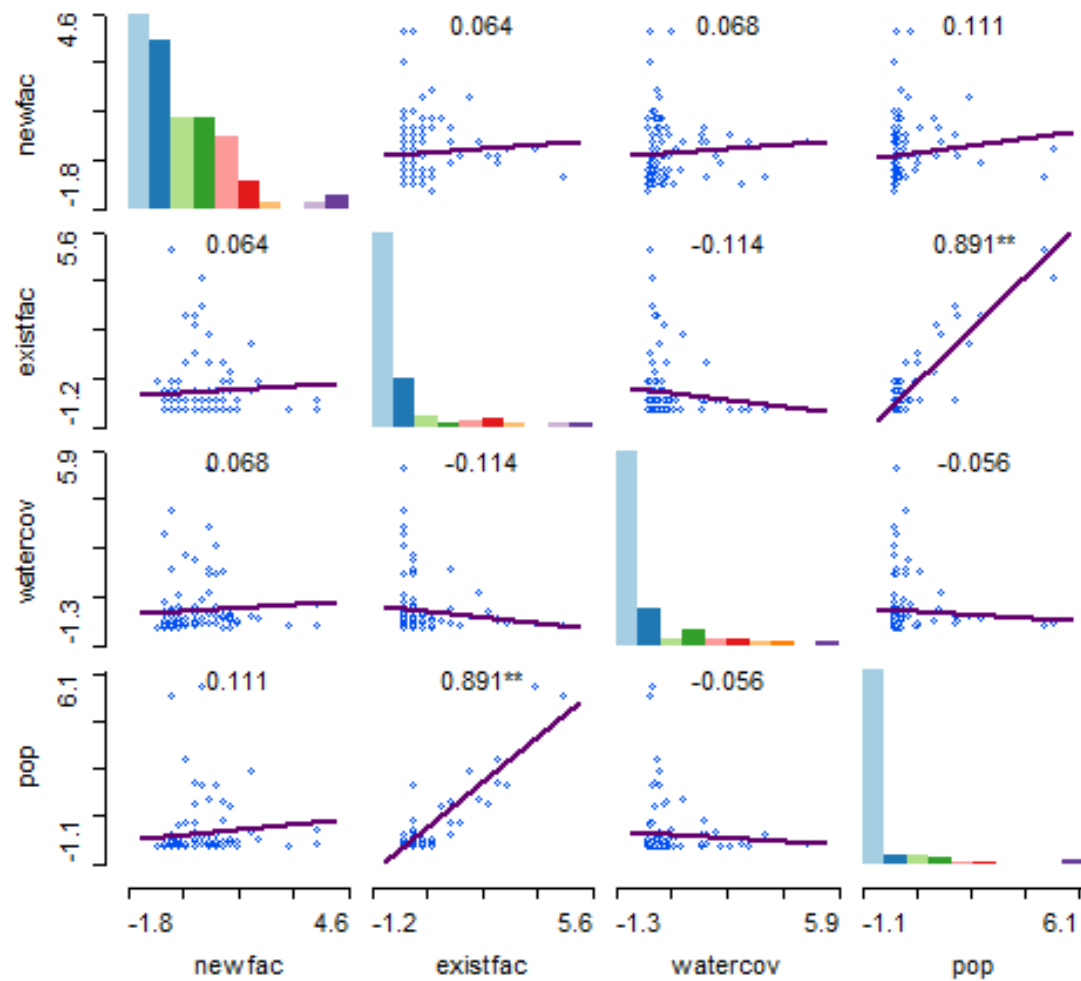


Figure 5.28: Correlation of existing and new healthcare facilities vs water bodies and population

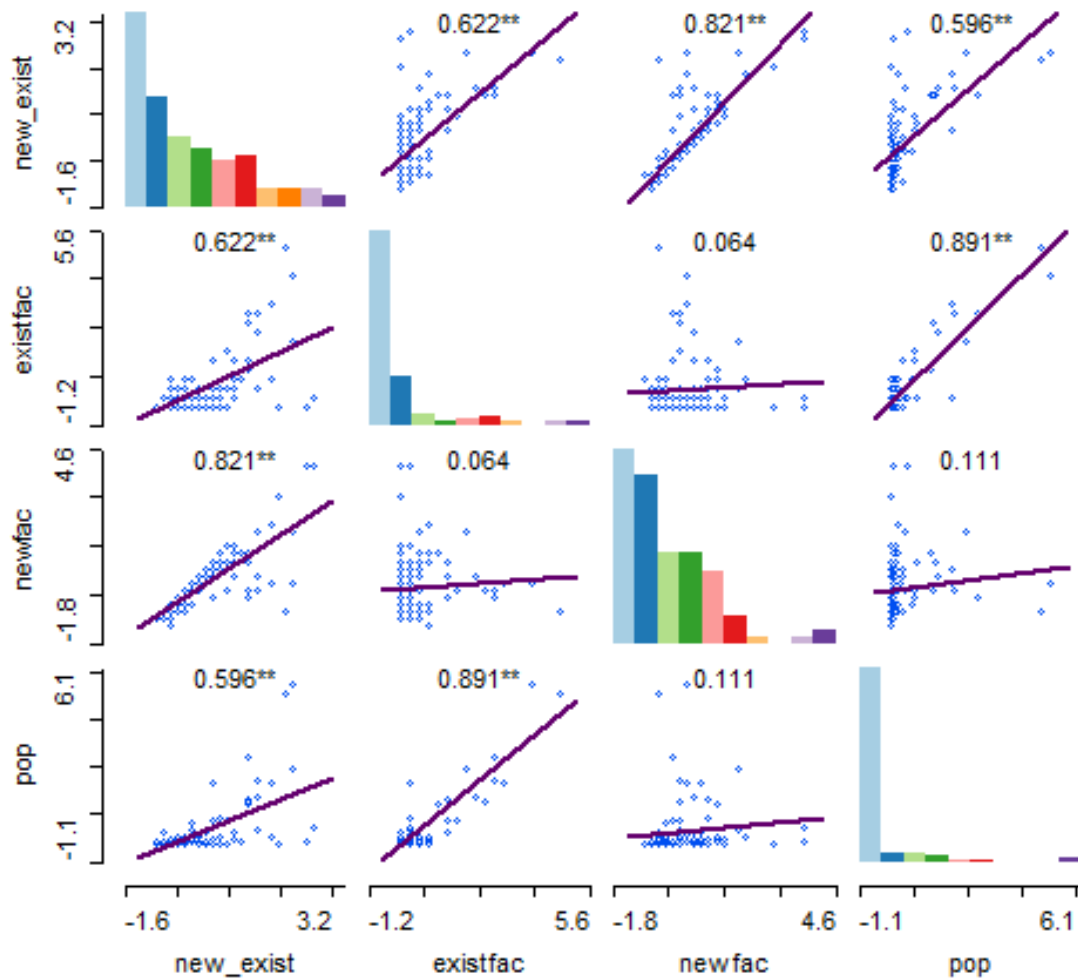


Figure 5.29: Correlation matrix of total healthcare facilities vs population

5.12 Overall Pattern Distribution of Optimized HCF Locations With Nearest Neighbour Analysis

Using the ESRI ArcGIS Pro 2.3.0, an overall pattern distribution of the newly established HCFs was performed with the nearest neighbour analysis that examines the distance between each new HCF location and the closest new HCF to it, in comparison with an expected complete spatial randomness pattern. Considering that the optimized HCFs were only located within the area uncovered by the existing HCFs, the pattern analysis was within the uncovered area.

The null hypothesis for average nearest neighbour states that HCFs are randomly distributed and are located independently of one another.

Alternative hypothesis: HCFs are not randomly distributed and are not located independently of one another.

The result revealed a dispersed pattern within this area. With a z-score of 6.6405 ($p < 0.01$), the newly established HCFs have a significant dispersed pattern (Figure 5.30).

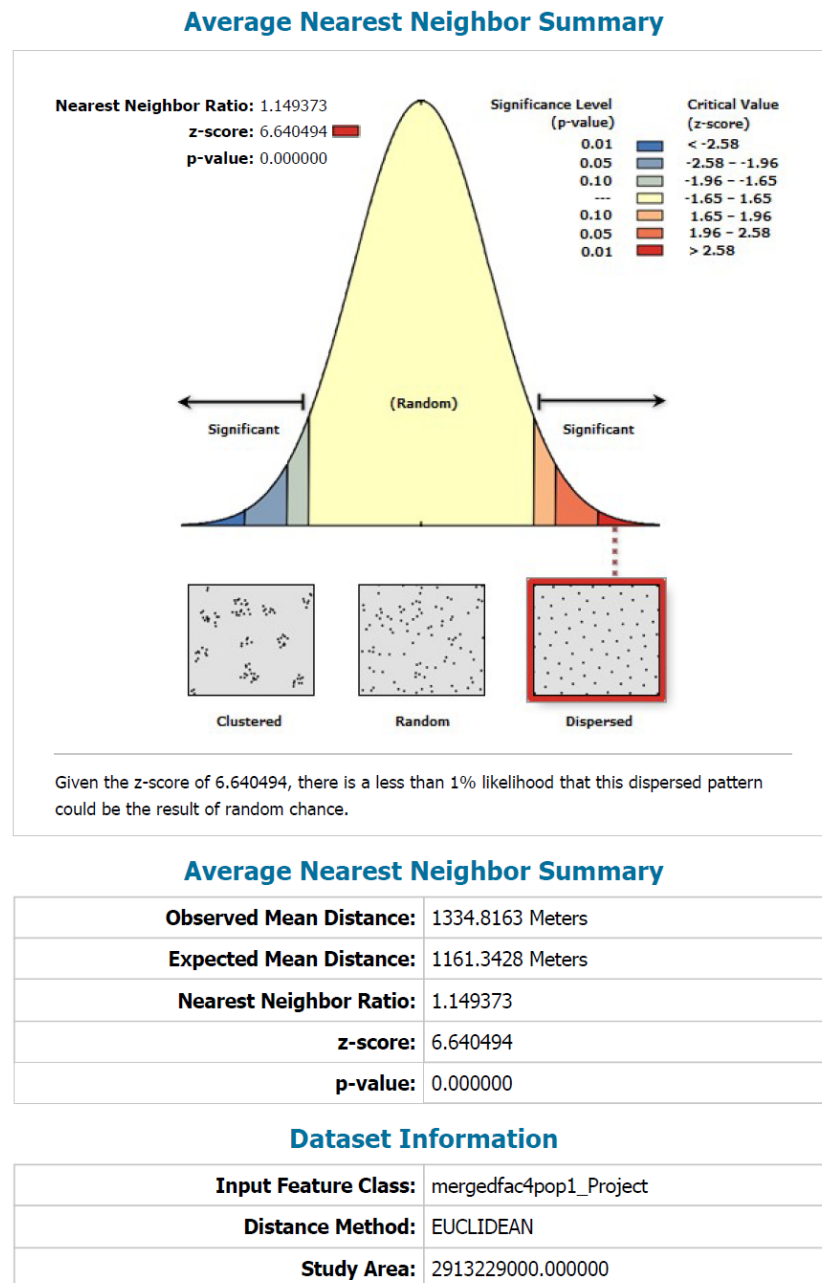


Figure 5.30: Average nearest neighbour analysis of new HCFs located within the

We therefore reject the null hypothesis of complete spatial randomness of optimized HCF locations and accept the alternative hypothesis that the optimized HCF locations have spatial relationship. The location of the HCFs are not a result of chance.

5.13 Results of Located Healthcare Facilities

The total number of facilities for the sample solution is 642 comprising of different types based on the objective functions and constraints in the algorithm. Figure 5.31 shows that Epe/Lekki region requires a large number of facilities. This region has a very large geographic extent and water coverage. Although the population in other parts of the

state may be low compared to the metropolis, theoretically, clustered populations require less number of facilities than dispersed population. Poor road network, and large expanse of land covered mostly by water body that impacts on accessibility implies that greater number of facilities will be required. However, with the population size factor, establishment costs can be greatly reduced as innovatively provided in this thesis.

It is also interesting to know that this isolated region also lacks other institutional facilities such as schools and markets, while HCFs in many of the urban regions such as Ikeja, Alimosho and Ifako-Agege have such facilities in their catchments. Overall HCF types with no such amenities are of greater number as revealed in Figure 5.32 and the pie chart in Figure 5.33.

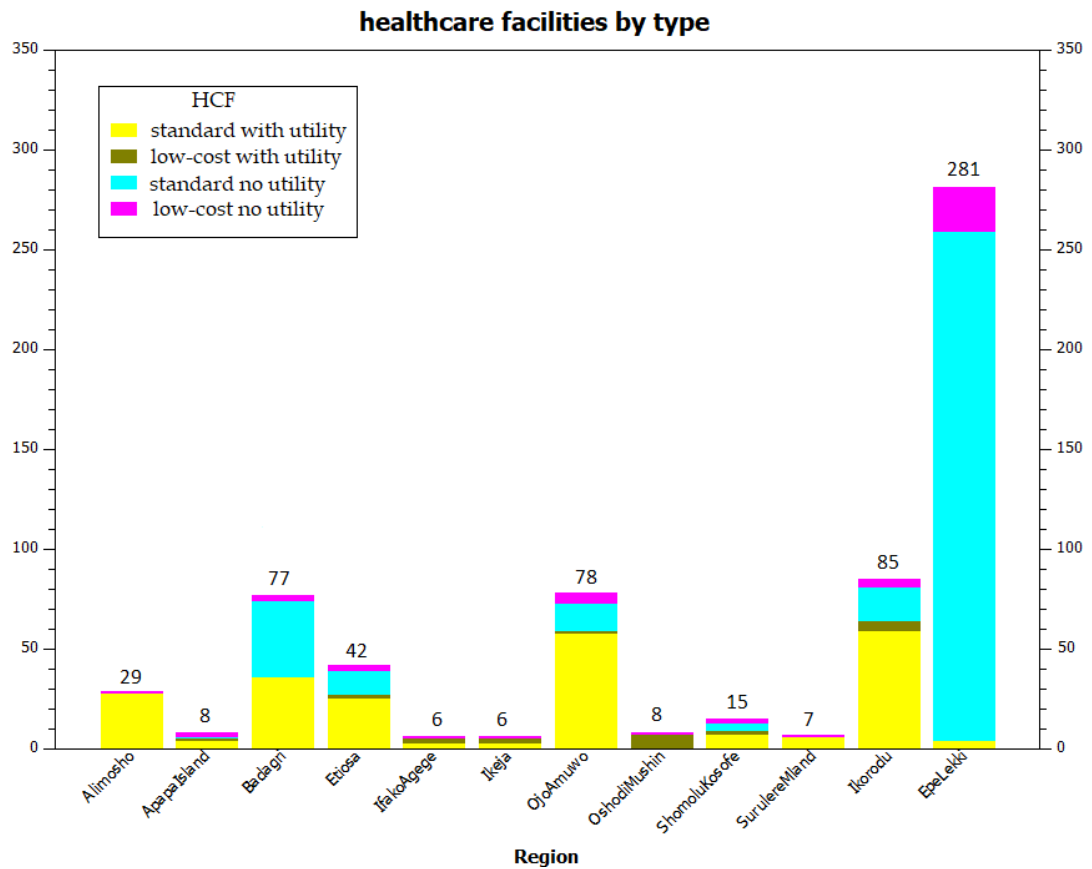


Figure 5.31: region healthcare facilities by type

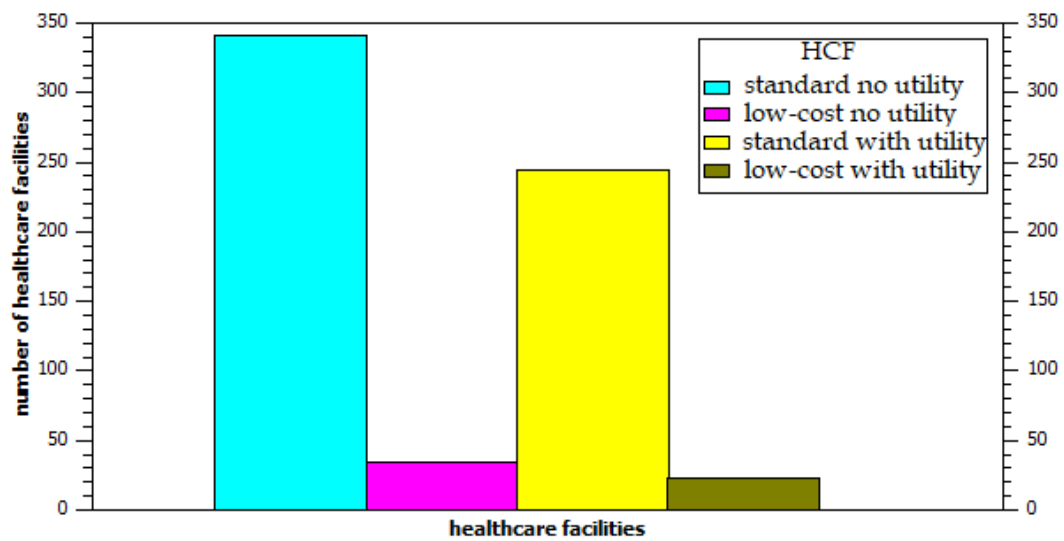


Figure 5.32: total located healthcare facilities by type

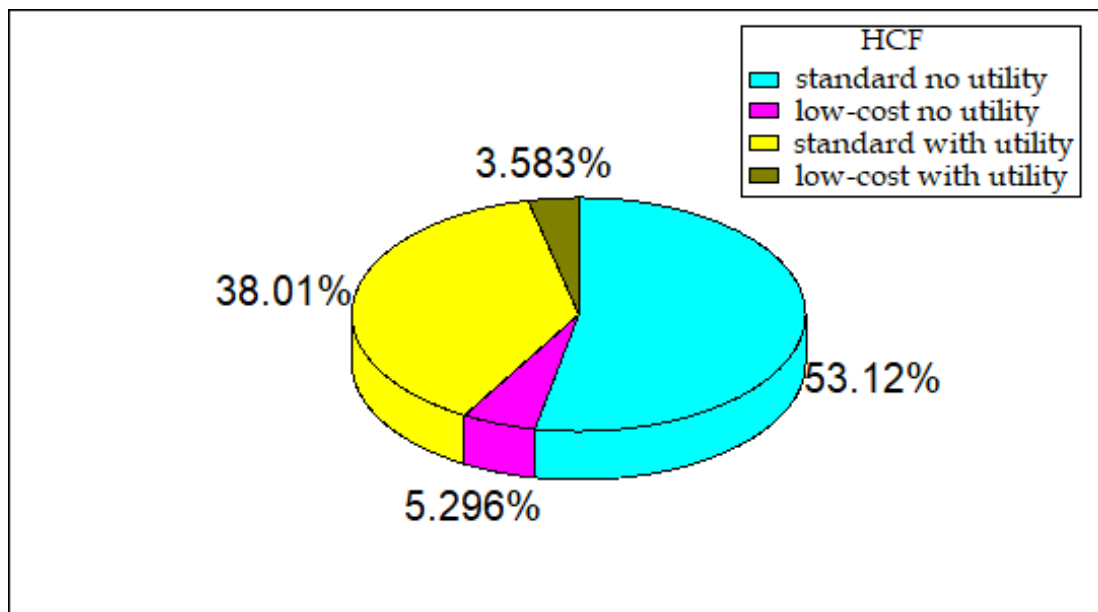


Figure 5.33: type of healthcare facilities by percentage

This is one of the great advantages of the flexibility and adaptability of the Spatial-ABM LNS algorithm for this thesis, which has not only revealed the shortage of HCFs, but other features that meet the health needs of the citizens.

Also, the results have shown the benefit of disaggregated analysis. All regions within Lagos state do not have the same geographic characteristics in terms of accessibility, coverage, road infrastructure, population distribution and availability of amenities that impact on health. For example, Oshodi/Mushin region only requires low-cost facilities and Alimosho requires mostly standard HCFs - all having other utilities in their catchments. However some regions need to be considered for other utilities.

The low-cost facilities provide the flexibility of allocating less resources for their establishment for future upgrades as population increases.

Other supporting consideration for the decision maker is the improvement on road infrastructure that can increase accessibility in some area, thus reducing the number of HCFs that need to be established. the flexibility of the model can also be utilised to increase the carrying capacity of the HCFs as part of budget-cutting strategies, depending on the prevailing circumstances.

These analyses and results have shown that the covering model developed in this thesis is a powerful decision support tool that has utilised Agent-Based modelling and Geospatial techniques at ensuring complete healthcare coverage for global or local health burdens. As demonstrated and revealed in the results, coverage of facilities in general is achievable. The approach also has the potential of revealing the phenomenon that are not included in the initial concept.

The map of the new HCFs is provided in Figure 5.34. In the figure:

LCwithUT = low-cost with utility facility

LCnoUT = low-cost no utility facility

STwithUT = standard with utility facility

STnoUT = standard no utility facility.

Figure 5.35 presents both existing and newly located HCFs; overlayed with water coverage.

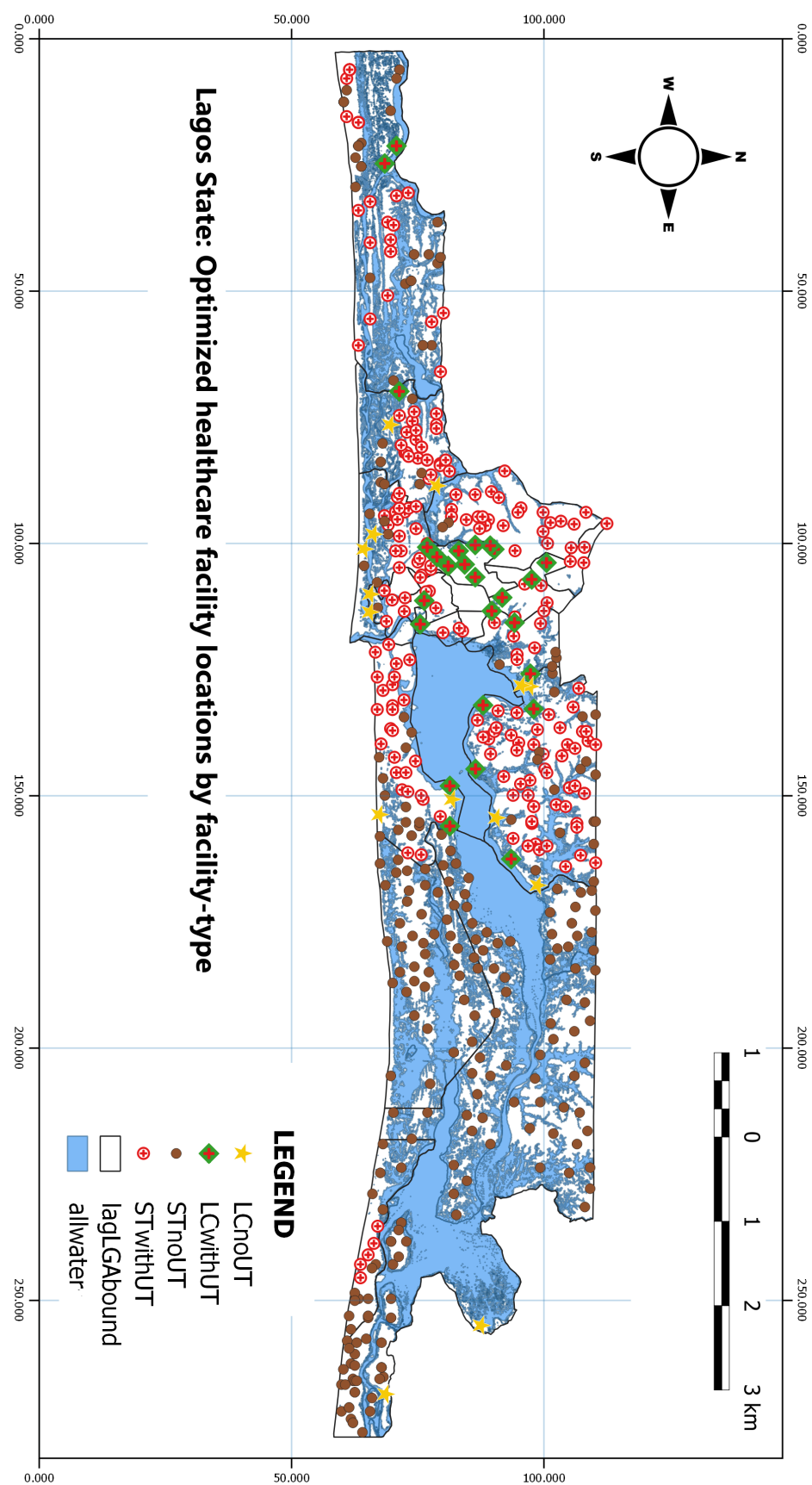


Figure 5.34: Lagos State: optimized healthcare facility locations by facility-type

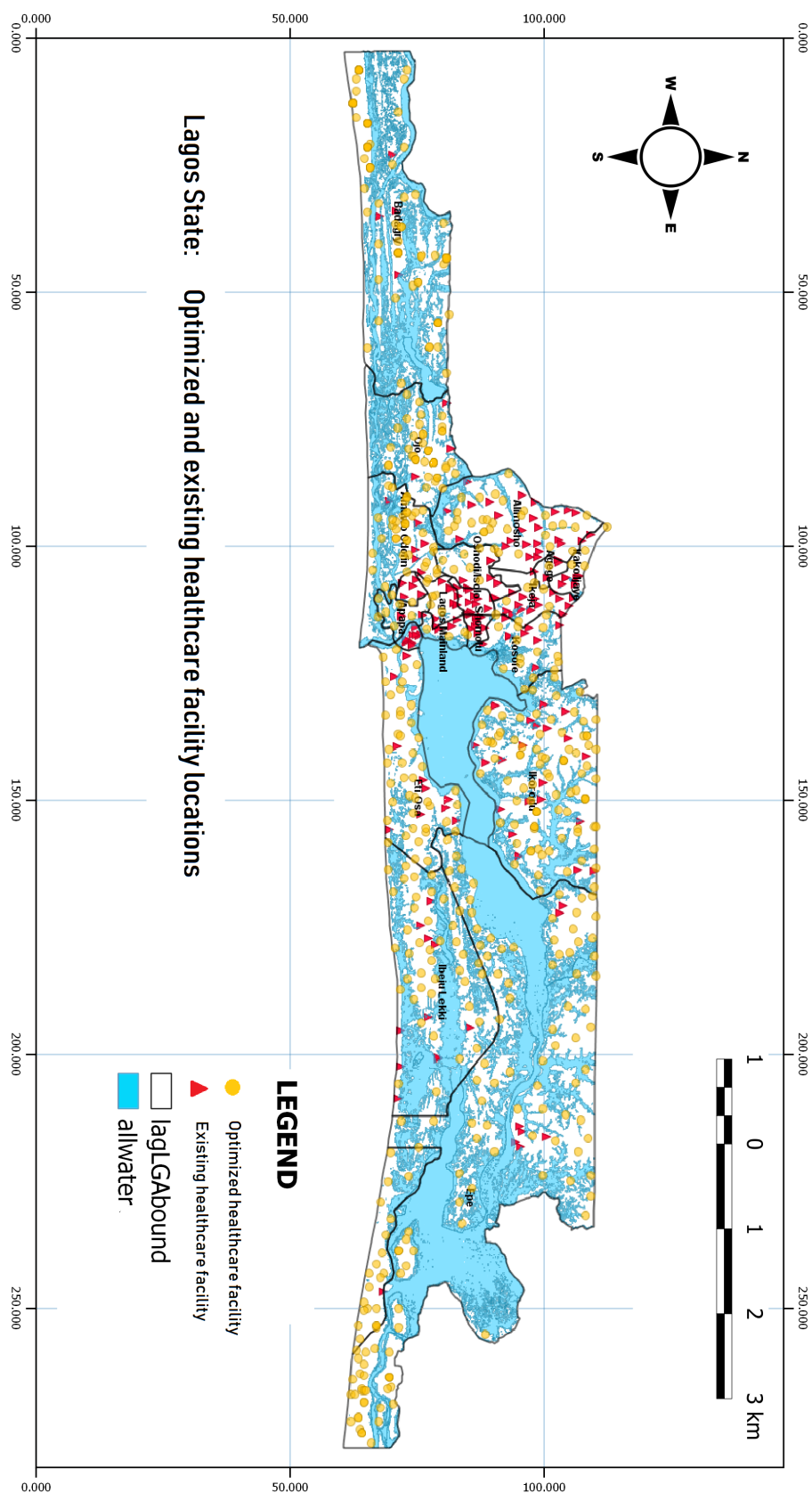


Figure 5.35: Lagos State: optimized and existing healthcare facility locations

5.14 Summary

In this chapter, the proposed models for this thesis were implemented on Lagos State, Nigeria - a case study with varying geographical and socio-economic characteristics. The geographic factors that influence HCF location were identified. The 30 minutes network proximity analysis reveals that accessibility to HCFs varies within regions in the state, with a consequent variation in HCFs' service area values. This depends on the road structure. There are clear indications that the 177 existing HCFs in Lagos state are highly insufficient and do not provide adequate coverage. Their locations are highly favourable to the densely populated urban and central regions, while the rural and outskirts are isolated as revealed in the spatial pattern analyses. The isolated areas were totally provided HCF coverage with the proposed DSABFL model that uses an establishment cost measure to differentiate low-cost HCFs from costly HCFs. Spatial pattern analysis of a feasible solution with 642 selected HCFs revealed a statistical significant dispersed pattern of the newly established HCFs that reflects the pattern distribution of the population, and other features such as water. This thesis based on a novel decision support DSABFL model, has identified the isolated regions that not only require HCFs, but other institutional facilities such as schools. These regions form the higher percentage of the areas secluded from health service coverage. It is apparent that certain regions can basically be provided with low-cost HCFs. This novel intervention can improve care for general well being in the study area.

Chapter 6

Conclusions and Future Work

6.1 Introduction

This chapter presents the summary and the conclusions drawn from this research. The contributions are discussed, including the limitations and considerations for future works.

6.2 Discussion

The main objective of this research to develop a novel capacitated covering-based optimization model for public healthcare facility locations using adaptable Spatial Agent-Based modelling methodology, through the integration of different spatial, economic and social characteristics has been successfully achieved. Mainly, the research focus is to facilitate sustainable healthcare policies for universal coverage of healthcare and ease of access to health service. This subsequently serves as part of strategies for reducing health burden in Nigeria. The research has provided a cost-effective location-allocation measure that incorporates the population size in the type of facilities to be established, while ensuring complete health coverage for all citizens within a realistic travel time threshold. From the questions posed in Chapter 1, the achievements and results of this thesis confirms thus:

1. Spatial Agent-Based modelling analysis can support decision-making in health-care coverage.
2. Geographic Information System (GIS) and Agent-Based Model (ABM) can be integrated for geospatial accessibility and facility location analysis for health interventions.

The requirement for solving the central healthcare facility location problem is considerably data based, involving real-world geographic, environmental and socio-economic complexity and heterogeneity. Typically, these problems are solved using

sophisticated mathematical or GIS models, or a combination of the two where the GIS provides the required spatial support. The mathematical model would require additional equation for each feature or complexity added to the problem, while the GIS analysis is a static and rational representation of this complexity. Hence the need for the new integrated approach offered by this thesis, based on the Maximum Covering Location Problem (MCLP) that seeks to cover as much demand as possible with fixed number of facilities, p , within a predefined service distance, D , as a cost-effective consideration for limited resources. An alternative approach to ensure that the demands outside the coverage of D do not travel inconsiderable distance compared to D is to introduce a distance value, S , greater than D as a constraint in the objective function, beyond which no demand is expected to travel to reach its closest facility. To realise these MCLP with mandatory closeness constraints and objectives, the decision maker is faced with several other challenges identified in Chapter 2. These include knowing the values of p , D and S ; and where to locate the p -facilities - without compromising full coverage and ease of access for resource availability. This research has provided a better cost-effective decision support that is flexible, adaptable and can ensure total health coverage. The capacity of the facility is integrated to propose a lower establishment cost for facilities that are under-served. The value of D is the average travel distance in the region of focus, while multiple values of p and S , together with location configurations, are endogenously suggested by the model. The systematic approach and findings of this thesis are provided as follows:

- A loose coupling method was adopted for integrating GIS and ABM. Before the choice of this integration approach, this work tried using close coupling of Agent Analyst extension in the ESRI ArcGIS 10.0 software and discovered that the extension is not compatible with subsequent versions of ArcGIS. The choice of loose coupling is considered to be the best approach such that the spatial data can be integrated without restriction of a system into the other system's programming environment. Therefore, this synergy is completely independent of any GIS toolkit, either proprietary or open source.
- A methodological framework was developed in Chapter 3 to identify the required spatial data and agents, and the data acquisition processes. The algorithms for proximity and location-allocation analyses were also described. GIS data acquisition and management capability was used to provide the data for the algorithms in ABM. Considering the limited time, technical and financial resources for this research – which is also usually the case in real-life decision-making processes, a budget-friendly and simple road data acquisition process was adopted for the travel-time analysis. Road networks were simply traced out on Google Earth Satellite imagery and converted to vector polyline dataset, considering that the open source ABM tool used does not require the level of accuracy and sophistication that the GIS needs for road connectivity. This was after confirming considerable gaps in the available data used by the on-line routing and service area tools and the image data. To represent patients' destinations, dwellings were

randomly simulated in the ABM.

- The coverage of the existing public healthcare services was assessed with a proximity analysis. As an improvement to the buffer analysis in resource constrained circumstances, the travel-time model allowed autonomous and mobile agents to travel along road network using the shortest path, within a user-defined time threshold and travel-speed. The polygon joining the end of the agents' journey formed the catchment of the existing healthcare facilities (HCFs). A straight-line measure was also included for the decision maker to simultaneously compare the two measures.
- The Adaptable Spatial Agent-Based Facility Location (ASABFL) technique proposed in this thesis has provided novel solutions to the covering problem by converting the continuous space into discrete points as potential facility locations, using meta-heuristics method. A Large Neighbourhood Search (LNS) algorithm was developed for which agents obeyed spatial rules such as: distance for determining their closeness to other agents, containment that examines agents' location within certain boundaries, adjacency where polygon features share boundary, connectivity for connecting demand locations to their closest HCFs, and intersection to check the occupancy of two agents on the same location. The agents take independent decisions while obeying these rules intelligently. Agents are either passive by not moving in the environment, or active by moving and changing their locations - following a suitability analysis of examining their locations within the neighbourhood, and avoiding forbidding locations such as water bodies or committed zones.

The LNS meta-heuristic algorithm is a six-phased process – initialization for determining initial feasible solution with greedy random location of facilities; construction for relocating facilities after evaluation of conflicts of interest based on intersection and containment spatial rules; destruction for eliminating agents based on adjacency, containment, distance and intersection rules; repair for regenerating more agents based on capacity rules; sorting for classifying selected HCFs based on size of population served and proximity to other amenities within their catchments, and determining uncovered population; and lastly improvement and stopping for linking uncovered population to the closest facility, and returning the values of p and S .

- Chapter 4 describes the model verification and validation processes. The robustness and consistencies of the models were confirmed via multiple runs and parameter variation. With 30 minutes travel limit at 48 m/min, real-life geographic locations of sets of origin (HCF) and destination (agent's travel limit) were used as inputs for walking travel distance and time requests to the Google Maps Distance Matrix Application Programming Interface (API). Evaluation of the results with Paired T-Test for Equivalence using the two-one-sided test (TOST) confirmed the equivalence of the Spatial-ABM travel-time model and Google Maps

results. Interestingly, the regions with no updated data on Google Maps either returned no values or unreliable results. The unrealistic Google Maps travel result was further confirmed with visual comparison of the paths taken in the two measures. Comparison of the Spatial-ABM and GIS buffer catchment values gave no notable difference.

The covering-based location-allocation model has produced 28.6% increase in population coverage against the existing optimal 5-facility results on the 55-node dataset for MCLP with mandatory distance constraints. Validation with optimal MCLP complete coverage within certain values of D indicates that equivalent and better results are obtainable with the proposed model.

Unlike the optimal solutions of other works that were compared, the ASABFL LNS meta-heuristic algorithm also provided better alternative values for p , D and S , including the configuration and locations of facilities for the decision maker by introducing a capacity concept into the model. Facilities serving less than a population threshold can be established with less resources.

- Chapter 5 describes the implementation of the developed method to Lagos State, Nigeria for healthcare coverage. It was revealed that a greater percentage of the people in Lagos State are uncovered by healthcare within the 30 minutes maximum travel threshold guideline of the World Health Organization (WHO). New HCFs were successfully located in the uncovered areas within the 20 Local Government Areas (LGAs), with each having different geographic and population settlement characteristics. These differing spatial characteristics are better analysed with disaggregated or individual level measure that yields more realistic results, and gives more insights into spatial relationship and agents' behaviour.

The pattern distribution and geostatistical spatial autocorrelation analysis with Morans' I revealed that the spatial distribution of existing HCFs is mainly within the highly populated urban metropolis. With the application of this model that has considered both capacity of HCFs and the spatial distribution of the population across the geographic space, complete healthcare coverage has been provided within the state. Various feasible alternative solutions are provided for policy decisions.

6.3 Conclusion

As part of strategy for the global healthcare coverage, a novel methodology that can integrate GIS and ABM for analysing public healthcare coverage and optimize HCF locations has been developed. A facility-level decision-making and management has been provided. Based on this synergy, a clear understanding of the relationship between travel time or distance and healthcare coverage was revealed: a well-structured road network gives a wider spatial coverage. The implementation and analyses revealed gap in health coverage in the study area.

With a knowledge of the existing coverage gap, potential sites for new HCFs were optimized for the uncovered population. This considered the spatial distribution of the population, proximity to other amenities and service distance. Lack of other amenities were revealed in the analyses. Population allocation to HCFs determined if the facility was serving population within, below or above the threshold, based on facility density ratio. Population was found to have an impact on facility establishment. Less population requires less resources.

Different alternative solutions for HCF configurations were produced to support decision on either to provide other amenities or select a configuration that requires no additional amenities, depending on the spatial characteristics of the focused region.

6.3.1 Contributions

The research has produced new methods that can be used in covering-based health-care facility location problems with effective combination of objectives of maximizing coverage, and minimizing the maximum distance between demand and health service locations, as identified in Chapter 2. The ABM adaptability and flexibility offer added advantage over the traditional modelling techniques. The significant contributions of this thesis are given in this section.

•Considerable increase in health service coverage and access for efficient healthcare system:

This thesis has produced better results than the optimal results from linear programming method at improving health service coverage and access to healthcare. Although, the meta-heuristic does not provide an optimal solution, feasible solution is returned at the end of each model run. While the study did not make a decision on the best configuration and number of HCFs, the LNS meta-heuristics algorithm provides alternative solutions that provide the decision maker a wide range of practicable and reliable choices that can improve healthcare coverage, save costs of establishment, and yield an efficient and successful healthcare system. This is better than a single solution that may be difficult to implement. The algorithm also recovers from initial input parameters without series of iterations.

•Simplification of complex decision making process is efficiently realised:

In this work, it is established that the geographic distribution and size of the population play key roles in the number and cost of HCFs that can be established. In other words, more facilities may not necessarily mean more expenditure as this depends on demographic and spatial factors that need to be integrated into the model. This research provided a novel method of integrating these complex spatial factors in the model to endogenously determine the number of HCFs and the distance that the uncovered

population in the MCLP will travel to their closest HCF.

The ABM simplifies the decision making process which the mathematical or GIS models cannot provide. This capability has supported the decision maker with better suggestions than can be produced with predetermined values. For example, the difference between D and S can be minimised and be as close as possible. This is a novel contribution, considering the complex and data driven decision processes that will be required to meet these objectives. This approach has contributed immensely in removing uncertainties faced in determining the number of HCFs and coverage distance for effective decision support and equity in public healthcare provisions.

●**Cost-effectiveness and data concerns:**

This thesis has contributed in the following cost-effective measures:

On-line road data requires local inputs and update prior to healthcare coverage analysis. Many travel time, service area or catchment analysis are based on the on-line road datasets, which may be driven by resource availability. However, this thesis has established the non-reliability of such datasets in some regions. They require to be verified and updated before being integrated into health coverage analysis, considering the great importance of health on the entire populations. A valuable contribution has been made by revealing that certain regions in Lagos State such as the rural communities can be secluded in health coverage plan if analysis is based on the traditional data availability assumption.

Improved alternative to circular buffer technique in case of scarce resources was provided with a network-based travel-time model developed with open source ABM toolkit that produced not only the travel time and distance, but revealed the travel pattern of patients along road network. The travel limit point feature can be imported to GIS where the service or catchment area is generated by connecting the point features with polygon features. Without the high technicality in connectivity graph as required in GIS, road data can be updated at reduced cost.

Number of facilities is not necessarily indicative of high establishment costs. The choice of the number of facilities has generally been considered a cost-effective measure because of the little marginal difference in population coverage with each additional facility. For example if six facilities will cover 90% population and seven facilities cover 90.5%, the decision maker may consider establishing six facilities. However, with the capacity integration in this thesis, low-cost HCF type that serves the population that is below the threshold value within their catchments is an additional valuable consideration for cost-effectiveness. An economic evaluation of more HCFs, including the low-cost HCFs, against fewer standard HCFs may result in less expenditure in terms of establishment, which includes factors such as health workers deployment. Interestingly, the same number of HCFs can suggest different establishment costs.

A reasonable consideration requires that health service is provided close to everyone notwithstanding their geographic locations. Fixing HCFs may be isolating a significant proportion of the population or greatly reducing accessibility to healthcare. This approach is providing an innovative adaptable and flexible support, and long-term impact for public HCFs provision. Future prediction of upgrading the low-cost HCFs when population increases, or expanding service area of the HCF when proximity is improved with better road infrastructure then becomes viable. This low-cost HCF type budget accommodating support is also valuable to the bottom-up approach for establishing community clinics.

●Generating general data and theory:

One novel information generated by this thesis is that a facility location model is capable of generating data that can create other valuable and distinct theories. The question of which areas are in need of institutional facilities such as schools, religious or commercial centres have been effectively formed and answered. This also represents a new and potential enlightenment for the development of theories. For example, why there are no HCFs in certain areas or why agents do not follow certain road paths.

The adaptable spatial-ABM as opposed to the GIS stand-alone static analysis has been able to identify regions that are without road infrastructure and other institutional facilities in the study area. These findings are important because they have impact on the health of the citizens. The policy makers can utilize a single analysis to plan for such facilities instead of allocating another budget for the research. For example, improved road infrastructure improves accessibility and consequently less HCFs may be required. As a further contribution to information and knowledge, the categorization of HCFs has included locations that do not satisfy the criteria of closeness to other institutional facilities - indicating their need for HCFs. Lack of such amenities should not impede the right of the people to good health. A consideration may therefore be to improve the infrastructural needs of such populations.

Another interesting inference from the outputs of the covering model is that the series of p and D provided shows the adaptability of the ABM to capture heterogeneous mixing and agent interactions, which enables it to give a better overall view of the coverage parameters.

●Experience with the case study:

Guided by the WHO specifications, the healthcare coverage for Lagos State was improved by optimizing new HCFs for the secluded populations based on 30 minutes travel time accessibility. Prior to optimizing the HCF locations, an analysis with the existing HCFs revealed that the greater percentage of people living in Lagos State are not within health service. Spatial accessibility differs greatly in all regions. Regions with poor road networks, especially the rural communities have lower spatial accessi-

bility. This low coverage explains one of the reasons for high mortality and morbidity rates due to health burden in Nigeria, considering that Lagos State is the most populated.

The distribution of the existing HCFs as revealed with the spatial pattern analysis is intensely inclined towards the urban areas where the population is mostly clustered. More importantly, the spatial autocorrelation analysis with Morans' I revealed statistical significant clustering of high values of the HCFs in the neighbourhood of the Lagos metropolis and urban settlements, and low values in the rural and outskirts of the state. The HCFs are located where there is little or no water coverage as revealed in their negative correlation of -0.114 against water covers. However, their association with population size is a strong positive correlation (+0.821). This poor pattern of HCF distribution has potential adverse health consequences on the people.

In this research geographical factors such as spatial distribution of the population, existing HCFs, other institutional facilities and water bodies which cover a larger proportion of the state were integrated into the analyses. More reliable results are produced when these spatial features are incorporated into the objectives. It was revealed as expected, that there is slight positive relationship of the newly established HCFs with water coverage and population with correlation of +0.068 and +0.111 respectively.

This thesis has therefore improved on the existing HCFs layout by providing additional HCFs to be accessible to the uncovered population. The selected HCFs were in two major categories: standard for HCFs that meet the capacity standard; and low-cost for HCFs that fall short of the capacity threshold. Each major category also contains subset of HCFs that do not pass the test of having other institutional facilities within their catchments. This unique classification resulted in a clear revelation that most areas with no health coverage are indeed lacking such institutional facilities. 58.416% of the selected HCFs are in this category.

●Global Impact

Lack of adequate health service coverage is a global health problem responsible for most global deaths and the depletion of the economy of the world. The cost-effective approach of spatial optimization modelling in resource-constrained circumstances presented in this thesis has clearly contributed efficient and significant global decision supports: providing knowledge and information on the existing healthcare coverage in the region with one of the highest morbidity and mortality rates in the world; effective improvement of healthcare provisions in this region; and realisation of the **WHO Global Technical Strategy for Malaria 2016-2030** at reducing death rates of malaria by at least 40% from 2015 levels in malaria endemic countries (WHO, 2015; 2018a), of which Nigeria is one.

The resource-management support has produced multi-facility-types and budget-friendly solutions that can serve as decision support for the deployment of health ser-

vice workers and community intervention. Establishing low-cost clinics within communities with small population size contributes to achieving a **Universal Health Coverage strategy**: to provide community-level health service.

With this intervention the global health burden will reduce, with a consequent effect on poverty, morbidity and mortality reduction; increased well-being; and the **Sustainable Development Goals (SDGs)** achievement. From the outcome of this research, it is apparent that this approach is effective and helpful to those in the front-line of health service provisions and those using optimization problems. Not only in the under-developed countries, but other general applications.

The Spatial-ABM approach is a simple, interactive tool for analysing proximity, coverage, accessibility and availability of not only healthcare facilities, but other public facility locations. Its performance against existing standard measures confirms its applicability with or without capacity constraints, and that it is transferable. Simultaneous analysis of network and straight-line analysis, including varying service distance in a model run is also possible. The location-allocation model can also adapt to different spatial population representation, either as nodal or continuous, as demonstrated on the 55-node and the case study datasets.

Also, the findings of this adaptable multi-criteria Spatial-ABM facility location models contribute to the existing literature by proffering better alternative solutions to facility locations that can be compared with complex mathematical models. This bottom-up approach can allow individual region to plan within their region rather than at national or state level.

6.4 Recommendation

Although, thankfully to insight to many health issues and technological advancement in the world, mortality and morbidity are reducing with consequent increase in life expectancy, there are still more than half of the global population that are not within reasonable reach of healthcare. Millions of people are still dying prematurely as a result of preventable health issues and lack of financial capability for healthcare. According to the World Bank and WHO (The World Bank, 2017), approximately half of the world's population do not have access to basic health services, 800 million people suffer financial hardship due to lack of access to healthcare by spending at least 10% of their income on healthcare, leading to extreme poverty for an estimated 100 million people, mortality and morbidity. Despite this scale of misfortune and loss, policy makers in countries that are hard pressed, especially with harsh economic and political conditions, may question why they should devote scarce resources and limited political term in office to analysing complex spatial healthcare coverage and establishing more HCFs. Many of these countries are still struggling with various resource, technological, political and economic imbalance that hinder the provision of adequate healthcare.

As part of the measures to overcome health coverage challenges in many developed countries, the tele-healthcare is being introduced. Healthcare is delivered remotely or over a distance through telephone or other electronic devices. Although it is said to improve coverage, preventive care, and management of health conditions, however, it is discovered to be costly and does not provide patient satisfaction. Studies have shown that most patients would prefer this to be complementary to face-to-face consultation rather than a substitute. Issues have also been raised about confidentiality, relationship with healthcare professionals, disruptions to normal workflows, internet and telephone network coverage and reliability (Black et al., 2011).

The proposed facility coverage tool in this thesis as opposed to existing proprietary tools requires little financial investment – the software is free, data acquisition is less costly, less technical resources are required, usage requires no internet, model can be run on external disks. In addition, spatial data acquisition is cost-saving, not labour or technical intensive and time consuming. The outcome is a complete healthcare coverage based on an effective management of budget, health workers, HCF establishment expenditures and equipment distribution. Patients can travel to HCFs within reasonable distance/time, with a consequent reduction in travel costs. The adaptation to standard facility/population ratio has the potential to reduce service waiting time and increase attention of health workers towards patients. The population well-being is also improved. The tool also factors in the information on availability of other institutional facilities which could have been carried out separately with additional resources. This also is a cost-effective means of improving the availability of such amenities within the community.

When dealing with healthcare for the population, it is not just about doing the right thing, but it is about doing things right with a conscious effort to give people what is rightly theirs in terms of health. Not taking the right decision to improve physical coverage of healthcare through facility location will incur more economic loss to individuals, government, private and public businesses, labour and productivity. Improving healthcare leads to significant increases in workforce, revenue, and profits. Quality healthcare can be provided, and millions of lives can be saved and standard of living of the world's population improved.

For policy makers, this tool is a platform and base line for realisation of other aspects of the SDGs and governance such as education and road infrastructure. As less resources are expended on siting HCFs (the population size within the catchment of a HCF will determine the resource allocation due to such facility), more revenue is generated from the healthy workforce that can also boost the economy. Within the limited office tenure of governance, a wider impact will therefore be made. Compared to existing facility location models, a larger percentage of coverage is obtained with the same number of facilities. Such intervention will address inequality in HCF distribution and access to healthcare. The cost-effective methodology serves as a decision-support tool for locating HCFs and distinguishes HCFs based on the present carrying capacity in comparison to the expected carrying capacity for appropriate resource allocation,

which includes health workers.

The proposed tool is recommended to be integrated in health policies and strategies for UHC to be executed within the phases of UHC realization. It offers cost-effective support in health coverage for all populations in the world, take over 100 million people out of poverty level, reduce global mortality and morbidity. In addition, there is the potential impact of reducing health burden, and over-stretching the existing workers and facilities, through effective distribution. These give a very large return on investment. With its implementation on Lagos, Nigeria, there is a clear indication that the tool meets the UHC goal. With increased pressure on the existing HCFs and a projected population size of twice the present population by 2050, Nigeria stands to prevent vicious poverty, adverse health burden, increased mortality and morbidity, economic downturn and impeded development with the adoption of this HCF optimization support in health policies.

Given the justification for universal health service coverage, this thesis recommends the new cost effective adaptable multi-criteria agent-based HCF location model and resource management decision support tool for HCF establishment for improved global well-being and poverty reduction.

6.5 Limitations and Future Works

While it can be concluded that this thesis has met its aims and objectives, some limitations are recognised and highlighted in this section:

The population data integrated into the analyses does not accurately represent the population in the study area. As pointed out in the research there is no concise demographic information available and the values used were estimates from Center for International Earth Science Information Network - CIESIN - Columbia University (2016) who also noted the unreliability of the data. However the research has demonstrated how population data can be integrated into the model both as vector or raster data.

This research cannot claim comprehensive information on other institutional facilities which were geocoded from on-line mapping tools. Also, the thesis has only integrated some uninhabitable regions. It is understood that some other land-use/land-cover features such as forest reserves can be integrated as forbidden zones. Other WHO considerations in the selection of site such as pollution, and flood risk will also need to be integrated in future work.

The results of the thesis have been presented based on walking scenario. Although the travel-speed could be varied to reflect a vehicular or other journey mode, however other factors such as waiting and turning at junctions can be incorporated in future work. Further improvement on the network travel-time model will create polygon features joining the agents' locations in the ABM. The estimated coverage for the network analysis was produced with the minimum bounding circle to represent the catchment,

however catchment may be obtained with other boundary geometries such as the convex hull: the output of course will vary. Nonetheless, the approach adopted in this thesis reveals that proximity is unequal from the centre of a facility in all directions.

The spatial accessibility in the covering model adopted the floating catchment method with straight-line using the average radius of the bounding circle in the region of focus. Future work will consider a network catchment definition as an extension of this model.

The thesis assumes that the existing HCFs already provide service coverage within their catchments based on travel-time analysis. However, this may not be the case if the population within the existing catchments is factored in. Future work will include this factor and consider if an existing HCF is over-served or under-served for possible suggestion of adding or relocating some HCFs.

6.6 Summary

The target for Universal Health Coverage has been on focus since the adoption of the Sustainable Development Goals in 2015, yet it is a basic fact that approximately half of the world's population are not within coverage of health services. Many countries, especially the low-and-middle-income nations lack the appropriate tools that can provide information on the population that is outside health service and the possible locations for new healthcare facilities. Hence, they face the difficulties of implementing technologies that they cannot sustain, finance and accomplish. Therefore, this research has addressed these challenges by providing alternative measure that can support healthcare providers and decision makers at improving health service coverage using geosimulation with GIS and ABM.

In summary, this dissertation has developed an Adaptable Spatial Agent-Based Facility Location (ASABFL) model by assessing the coverage of existing healthcare facilities with an adaptable and interactive travel-time model that measures the proximity of the population to the HCFs. The performance of the models were compared with other models and datasets. With this process, the travel-time was found to be comparable with web-mapping tools and has also provided better dataset and information. The location-allocation covering model endogenously provides alternative sets of solutions that the decision maker can consider in the choice of number of facilities to locate, expenditure and population coverage. With its application to existing datasets, the model produced comparable and better results. The analysis on the implementation to Lagos State shows that the spatial distribution of demands has a large impact on facility location. Geospatial analysis shows significance relationship of neighbouring geographical features in location decisions. Finally, the thesis demonstrated usefulness of the model in simultaneously supporting decisions in providing other amenities. The model's successful application to other dataset established its transferability.

Appendix A

Research Papers and Award

The research papers and award received in the course of this research are given below.

Journal article

Olowofoyeku, O., Jethro, S., Lipika, D. and Eric, G. (2019). Healthcare facility coverage for malaria and sickle cell disease treatment. *International Journal of Health, Wellness, and Society*. (In press).

Conference

Olowofoyeku O.O. (2018). Vector-based geosimulation travel time model for health-care facilities catchment. ACM-W UK INSPIRE conference 20 April - Poster and lightning presentation.

Submitted journal article

Olukemi O. Olowofoyeku, Jethro Shell, Lipika Deka, Francisco Chiclana and Eric Goodyer. (2019). Disaggregated Spatial-agent based coverage modeling in resource constrained systems: A decision support for healthcare in Ikorodu Local Government Area, Nigeria. *Global Health Research and Policy*. (Under review)

Award received

Best poster award - Poster and lightning presentation. 3rd ACM-W UK inspire conference 2018.

Bibliography

Abar, S., Theodoropoulos, G. K., Lemarinier, P. and OHare, G. M. (2017). Agent based modelling and simulation tools: a review of the state-of-art software, *Computer Science Review* **24**: 13–33.

Abasiubong, F., Bassey, E. A., Udobang, J. A., Akinbami, O. S., Udoh, S. B. and Idung, A. U. (2012). Self-medication: potential risks and hazards among pregnant women in Uyo, Nigeria, *Pan African Medical Journal* **13**(1).

Abbas, I. I., Auta, S. Z., Naiya, R. M. et al. (2012). Health care facilities mapping and database creation using GIS in Chikun Local Government, Kaduna State, Nigeria, *Global Journal of Human-Social Science Research* **12**(10-B).

Abdou, M., Hamill, L. and Gilbert, N. (2012). Designing and building an agent-based model, *Agent-based Models of Geographical Systems*, Springer, pp. 141–165.

Abdulraheem, I., Olapipo, A. and Amodu, M. (2012). Primary health care services in Nigeria: Critical issues and strategies for enhancing the use by the rural communities, *J Public Health Epidemiol* **4**(1): 5–13.

Aday, L. A. and Andersen, R. (1974). A framework for the study of access to medical care, *Health Services Research* **9**(3): 208.

Afolabi, O. J., Hassan, A. M. and Age, L. O. (2017). Behavioral pattern of commercial public transport passengers in Lagos metropolis, *Journal of Sustainable Development of Transport and Logistics* **2**(1): 40–50.

Afshari, H., McLeod, R. D., ElMekkawy, T. and Peng, Q. (2014). Distribution-service network design: An agent-based approach, *Procedia CIRP* **17**: 651–656.

Afshari, H. and Peng, Q. (2014). Challenges and solutions for location of healthcare facilities, *Ind Eng Manag* **3**(1): 12.

Agency for Healthcare Research and Quality (AHRQ) (2014). Gap analysis. how to improve hospital quality and safety, accessed 17 January 2016.

URL: <http://www.ahrq.gov/sites/default/files/wysiwyg/professionals/systems/hospital/qitoolkit/combined>

Ahmadi-Javid, A., Seyedi, P. and Syam, S. S. (2017). A survey of healthcare facility location, *Computers & Operations Research* **79**: 223–263.

- Ahuja, R. K., Ergun, Ö., Orlin, J. B. and Punnen, A. P. (2002). A survey of very large-scale neighborhood search techniques, *Discrete Applied Mathematics* **123**(1-3): 75–102.
- Aikens, C. H. (1985). Facility location models for distribution planning, *European Journal of Operational Research* **22**(3): 263–279.
- Alegana, V. A., Wright, J. A., Pentrina, U., Noor, A. M., Snow, R. W. and Atkinson, P. M. (2012). Spatial modelling of healthcare utilisation for treatment of fever in Namibia, *Int J Health Geogr* **11**(6): 10–1186.
- Ali, A. N. and Onokala, P. C. (2008). Availability of healthcare facilities: implications for rural development in Enugu State, Nigeria, *Nigerian Journal of Geography and the Environment* **1**(1): 89–100.
- Ali, S. (2012). *Multidisciplinary Computational Intelligence Techniques: Applications in Business, Engineering, and Medicine: Applications in Business, Engineering, and Medicine*, IGI Global.
- Anselin, L. (1995). Local indicators of spatial association–LISA, *Geographical Analysis* **27**(2): 93–115.
- Arah, O. A., Klazinga, N. S., Delnoij, D. M., Asbroek, A. t. and Custers, T. (2003). Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement, *International Journal for Quality in Health care* **15**(5): 377–398.
- Arcury, T. A., Gesler, W. M., Preisser, J. S., Sherman, J., Spencer, J. and Perin, J. (2005). The effects of geography and spatial behavior on health care utilization among the residents of a rural region, *Health Services Research* **40**(1): 135–156.
- Aregbeshola, B. S. and Khan, S. M. (2017). Primary health care in Nigeria: 24 years after Olikoye Ransome-Kutis leadership, *Frontiers in Public Health* **5**: 48.
- Aringhieri, R., Carello, G. and Morale, D. (2007). Ambulance location through optimization and simulation: the case of Milano urban area, *XXXVIII Annual Conference of the Italian Operations Research Society Optimization and Decision Sciences*, pp. 1–29.
- Awoyemi, T., Obayelu, O. and Opaluwa, H. (2011). Effect of distance on utilization of health care services in rural Kogi State, Nigeria, *Journal of Human Ecology* **35**(1): 1–9.
- Ayoade, M. (2016). The location of public maternal health care facilities and maternal mortality in ibadan metropolis, Nigeria, *International Journal of Social Sciences* **5**(4): 1–18.
- Baeten, R., Spasova, S., Vanhercke, B. and Coster, S. (2018). Inequalities in access to healthcare, *A study of National Policies, European Social Policy Network (ESPN), Brussels: European Commission* .
- Balcik, B. and Beamon, B. M. (2008). Facility location in humanitarian relief, *International Journal of Logistics* **11**(2): 101–121.

- Banik, A., Bhattacharya, B. B. and Das, S. (2014). Minimum enclosing circle of a set of fixed points and a mobile point, *Computational Geometry* **47**(9): 891–898.
- Barbati, M., Bruno, G. and Genovese, A. (2011). An agent-based framework for solving an equity location problem, *Agent and Multi-Agent Systems: Technologies and Applications*, Springer, pp. 486–494.
- Barbati, M., Bruno, G. and Genovese, A. (2012). Applications of agent-based models for optimization problems: A literature review, *Expert Systems with Applications* **39**(5): 6020–6028.
- Basu, R., Jana, A. and Bardhan, R. (2018). A health care facility allocation model for expanding cities in developing nations: strategizing urban health policy implementation, *Applied Spatial Analysis and Policy* **11**(1): 21–36.
- Batty, M. (2012). A generic framework for computational spatial modelling, *Agent-based Models of Geographical Systems*, Springer, pp. 19–50.
- Beheshtifar, S. and Alimoahmadi, A. (2015). A multiobjective optimization approach for location-allocation of clinics, *International Transactions in Operational Research* **22**(2): 313–328.
- Berman, O., Drezner, Z. and Krass, D. (2010). Generalized coverage: New developments in covering location models, *Computers & Operations Research* **37**(10): 1675–1687.
- Black, A. D., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., McKinstry, B., Procter, R., Majeed, A. and Sheikh, A. (2011). The impact of ehealth on the quality and safety of health care: a systematic overview, *PLoS Medicine* **8**(1): e1000387.
- Bolouri, S., Vafaeinejad, A., Alesheikh, A. A. and Aghamohammadi, H. (2018). The ordered capacitated multi-objective location-allocation problem for fire stations using spatial optimization, *ISPRS International Journal of Geo-Information* **7**(2): 44.
- Boonmee, C., Arimura, M. and Asada, T. (2017). Facility location optimization model for emergency humanitarian logistics, *International Journal of Disaster Risk Reduction* **24**: 485–498.
- Borshchev, A. and Filippov, A. (2004). From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools, *Proceedings of the 22nd International Conference of the System Dynamics Society*, Vol. 22.
- Boulos, M. K., Roudsari, A. V. and Carson, E. R. (2001). Health geomatics: an enabling suite of technologies in health and healthcare, *Journal of Biomedical Informatics* **34**(3): 195–219.
- Brabyn, L. and Skelly, C. (2002). Modeling population access to New Zealand public hospitals, *International Journal of Health Geographics* **1**(1): 3.

- Breheny, M. J. (1978). The measurement of spatial opportunity in strategic planning, *Regional Studies* **12**(4): 463–479.
- Brinkhoff, T., Kriegel, H.-P. and Schneider, R. (1993). Comparison of approximations of complex objects used for approximation-based query processing in spatial database systems, *Data Engineering, 1993. Proceedings. Ninth International Conference on*, IEEE, pp. 40–49.
- Brown, D. G., Riolo, R., Robinson, D. T., North, M. and Rand, W. (2005). Spatial process and data models: Toward integration of agent-based models and GIS, *Journal of Geographical Systems* **7**(1): 25–47.
- Bruglieri, M., Pezzella, F. and Pisacane, O. (2018). An adaptive large neighborhood search for relocating vehicles in electric carsharing services, *Discrete Applied Mathematics* .
- Burns, L. D. and Golob, T. F. (1976). The role of accessibility in basic transportation choice behavior, *Transportation* **5**(2): 175–198.
- Caccetta, L., Dzator, M. et al. (2005). Heuristic methods for locating emergency facilities, *Proceedings of the 16th International Congress on Modelling and Simulation, Melbourne, Australia, Citeseer*, pp. 1744–1750.
- Carlson, T., York, S. and Primomo, J. (2011). The utilization of geographic information systems to create a site selection strategy to disseminate an older adult fall prevention program, *The Social Science Journal* **48**(1): 159–174.
- Castle, C. J. and Crooks, A. T. (2006). Principles and concepts of agent-based modelling for developing geospatial simulations.
- Center for International Earth Science Information Network - CIESIN - Columbia University (2016). Gridded Population of the World, Version 4 (GPWv4): Population Density, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Accessed 03 September 2017.
URL: <http://dx.doi.org/10.7927/H4HX19NJ>
- Chanda, E., Mukonka, V. M., Mthembu, D., Kamuliwo, M., Coetzer, S. and Shinondo, C. J. (2012). Using a geographical-information-system-based decision support to enhance malaria vector control in Zambia, *Journal of Tropical Medicine* **2012**.
- Chauhan, A. and Singh, A. (2016). A hybrid multi-criteria decision making method approach for selecting a sustainable location of healthcare waste disposal facility, *Journal of Cleaner Production* **139**: 1001–1010.
- Chen, Y., Li, X., Liu, X. and Liu, Y. (2010). An agent-based model for optimal land allocation (AgentLA) with a contiguity constraint, *International Journal of Geographical Information Science* **24**(8): 1269–1288.
- Christaller, W. (1966). *Central places in southern Germany*, Prentice Hall.

- Church, R. L. (1984). Symposium on location problems: In memory of Leon Cooper: The planar maximal covering location problem, *Journal of Regional Science* **24**(2): 185–201.
- Church, R. L. (1999). Location modelling and GIS, *Geographical information systems* **1**: 293–303.
- Church, R. L. (2002). Geographical information systems and location science, *Computers & Operations Research* **29**(6): 541–562.
- Church, R. L. and ReVelle, C. S. (1976). Theoretical and computational links between the p-median, location set-covering, and the maximal covering location problem, *Geographical Analysis* **8**(4): 406–415.
- Church, R. L. and Roberts, K. L. (1983). Generalized coverage models and public facility location, *Papers in Regional Science* **53**(1): 117–135.
- Church, R. and ReVelle, C. (1974). The maximal covering location problem, *Papers of the Regional Science Association*, Vol. 32, Springer, pp. 101–118.
- Ciari, F., Axhausen, K. W., Axhausen, K. W. and Axhausen, K. W. (2011). *Modeling location decisions of retailers with an agent-based approach*, Eidgenössische Technische Hochschule Zürich, IVT, Institut für Verkehrsplanung und Transportsysteme.
- Cohen, R. and Katzir, L. (2008). The generalized maximum coverage problem, *Information Processing Letters* **108**(1): 15–22.
- Cowen, D. et al. (1990). Unit 01-What is GIS?, in M. F. Goodchild and K. K. Kemp (eds), *NCGIA Core Curriculum in GIS*, National Center for Geographic Information and Analysis, UC Santa Barbara, chapter 01.
- Cremers, A. L., Alege, A., Nelissen, H. E., Okwor, T. J., Osibogun, A., Gerrets, R. and Vant Hoog, A. H. (2019). Patients and healthcare providers perceptions and practices regarding hypertension, pharmacy-based care, and mHealth in Lagos, Nigeria: a mixed methods study, *Journal of Hypertension* **37**(2): 389.
- Cromley, E. K. and McLafferty, S. L. (2011). *GIS and public health*, Guilford Press.
- Crooks, A., Croitoru, A., Lu, X., Wise, S., Irvine, J. M. and Stefanidis, A. (2015). Walk this way: Improving pedestrian agent-based models through scene activity analysis, *ISPRS International Journal of Geo-Information* **4**(3): 1627–1656.
- Crooks, A. T. (2006). Exploring cities using agent-based models and GIS, *Centre for Advanced Spatial Analysis Working Paper Series, Paper 109*, ISSN 1467-1298, Centre for Advanced Spatial Analysis, UCL (University College London), accessed 12 January 2016.
URL: <http://discovery.ucl.ac.uk/3341/>
- Crooks, A. T. and Heppenstall, A. J. (2012). Introduction to agent-based modelling, *Agent-based Models of Geographical Systems*, Springer, pp. 85–105.

- Crooks, V. A. and Andrews, G. J. (2009). Community, equity, access: core geographic concepts in primary health care, *Primary Health Care Research & Development* **10**(03): 270–273.
- Daskin, M. S. and Dean, L. K. (2005). Location of health care facilities, *Operations Research and Health care*, Springer, pp. 43–76.
- Daskin, M. S. and Owen, S. H. (1999). Two new location covering problems: The partial p-center problem and the partial set covering problem, *Geographical Analysis* **31**(3): 217–235.
- Daskin, M. S. and Stern, E. H. (1981). A hierarchical objective set covering model for emergency medical service vehicle deployment, *Transportation Science* **15**(2): 137–152.
- De Smith, M. J., Goodchild, M. F. and Longley, P. (2018). *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*, 6th edn, Online. Accessed January 18, 2019.
URL: <https://www.spatialanalysisonline.com/HTML/index.html>
- Dekle, J., Lavieri, M. S., Martin, E., Emir-Farinas, H. and Francis, R. L. (2005). A Florida county locates disaster recovery centers, *Interfaces* **35**(2): 133–139.
- Delamater, P. L., Messina, J. P., Shortridge, A. M. and Grady, S. C. (2012). Measuring geographic access to health care: raster and network-based methods, *International Journal of Health Geographics* **11**(1): 1.
- Demir, E., Bektaş, T. and Laporte, G. (2012). An adaptive large neighborhood search heuristic for the pollution-routing problem, *European Journal of Operational Research* **223**(2): 346–359.
- Djurhuus, S., Hansen, H. S., Aadahl, M. and Glümer, C. (2015). Building a multimodal network and determining individual accessibility by public transportation, *Environment and Planning B: Planning and Design* p. 0265813515602594.
- Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *Cartographica: The International Journal for Geographic Information and Geovisualization* **10**(2): 112–122.
- Dudko, Y., Robey, D. E., Kruger, E. and Tennant, M. (2018). Selecting a location for a primary healthcare facility: combining a mathematical approach with a geographic information system to rank areas of relative need, *Australian Journal of Primary Health* **24**(2): 130–134.
- Dzator, M. and Dzator, J. (2018). The p-median problem and health facilities: Cost saving and improvement in healthcare delivery through facility location, *Data and Decision Sciences in Action*, Springer, pp. 369–380.

- Ejiagha, I. R., Ojiako, J. C. and Eze, C. G. (2012). Accessibility analysis of healthcare delivery system within Enugu urban area using geographic information system.
- Ekpenyong, M. S., Bond, C. and Matheson, D. (2019). Challenges of maternal and prenatal care in Nigeria.
- Fadahunsi, J., Kufoniyi, O. and Babatimehin, O. (2017). Spatial analysis of distribution patterns of healthcare facilities in Osun State, Nigeria, *Universal Journal of Public Health* 5(7).
- Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A. and Roventini, A. (2017). Validation of agent-based models in economics and finance, *Technical report*, LEM Working Paper Series.
- Farahani, R. Z., Asgari, N., Heidari, N., Hosseini, M. and Goh, M. (2012). Covering problems in facility location: A review, *Computers & Industrial Engineering* 62(1): 368–407.
- Farahani, R. Z., Hekmatfar, M., Fahimnia, B. and Kazemzadeh, N. (2014). Hierarchical facility location problem: Models, classifications, techniques, and applications, *Computers & Industrial Engineering* 68: 104–117.
- Filatova, T., Verburg, P. H., Parker, D. C. and Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: Challenges and prospects, *Environmental Modelling & Software* 45: 1–7.
- Fisher, R. and Lassa, J. (2017). Interactive, open source, travel time scenario modelling: tools to facilitate participation in health service access analysis, *International Journal of Health Geographics* 16(1): 13.
- Freiria, S., Ribeiro, B. and Tavares, A. O. (2015). Understanding road network dynamics: link-based topological patterns, *Journal of Transport Geography* 46: 55–66.
- Friedrich, C. J. and Weber, A. (1929). *Alfred Weber's theory of the location of industries*, University of Chicago Press.
- Galvão, R. D. (2004). Uncapacitated facility location problems: contributions, *Pesquisa Operacional* 24(1): 7–38.
- Galvão, R. D. and ReVelle, C. (1996). A lagrangean heuristic for the maximal covering location problem, *European Journal of Operational Research* 88(1): 114–123.
- Gething, P. W., Noor, A. M., Zurovac, D., Atkinson, P. M., Hay, S. I., Nixon, M. S. and Snow, R. W. (2004). Empirical modelling of government health service use by children with fevers in Kenya, *Acta Tropica* 91(3): 227–237.
- Gilbert, G. N. (2008). *Agent-based models*, number 153, Sage.
- Gilliland, J. A., Shah, T. I., Clark, A., Sibbald, S. and Seabrook, J. A. (2019). A geospatial approach to understanding inequalities in accessibility to primary care among vulnerable populations, *PloS one* 14(1): e0210113.

- Gu, W., Wang, X. and McGregor, S. E. (2010). Optimization of preventive health care facility locations, *International Journal of Health Geographics* **9**(1): 1.
- Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R. and Hudson, M. (2002). What does 'access to health care' mean?, *Journal of Health Services Research & Policy* **7**(3): 186–188.
- Gulzar, L. (1999). Access to health care, *Image: the Journal of Nursing Scholarship* **31**(1): 13–19.
- Hacar, M., Kılıç, B. and Şahbaz, K. (2018). Analyzing openstreetmap road data and characterizing the behavior of contributors in Ankara, Turkey, *ISPRS International Journal of Geo-Information* **7**(10): 400.
- Hafez, R. (2018). Nigeria health financing system assessment, World Bank Group, Washington, D.C. Health, Nutrition and Population (HNP) Discussion Paper.
URL: <http://documents.worldbank.org/curated/en/782821529683086336/Nigeria-health-financing-system-assessment>
- Haghani, A. (1996). Capacitated maximum covering location models: Formulations and solution procedures, *Journal of Advanced Transportation* **30**(3): 101–136.
- Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph, *Operations Research* **12**(3): 450–459.
- Harper, P., Shahani, A., Gallagher, J. and Bowie, C. (2005). Planning health services with explicit geographical considerations: a stochastic location–allocation approach, *Omega* **33**(2): 141–152.
- Hartmann, T. and Zerjav, V. (2014). Optimizing the location of out-care centers in urban space using agent-based modeling, *Construction Research Congress 2014@ sConstruction in a Global Network*, ASCE, pp. 2375–2384.
- Higgs, G., Langford, M., Jarvis, P., Page, N., Richards, J. and Fry, R. (2019). Using geographic information systems to investigate variations in accessibility to extended hours primary healthcare provision, *Health & Social Care in the Community*.
- Huisman, O. and De By, R. (2009). Principles of geographic information systems, *ITC Educational Textbook Series* **1**: 17.
- Humann, J. and Madni, A. M. (2014). Integrated agent-based modeling and optimization in complex systems analysis, *Procedia Computer Science* **28**: 818–827.
- Idiegbeyan-Ose, J., Nkiko, C., Idahosa, M. and Nwokocha, N. (2016). Digital divide: issues and strategies for intervention in Nigerian libraries, *Journal of Cases on Information Technology (JCIT)* **18**(3): 29–39.
- Ingram, D. R. (1971). The concept of accessibility: a search for an operational form, *Regional Studies* **5**(2): 101–107.

- Isma'il, M. and Musa, I. J. (2014). Analysis of accessibility to healthcare facilities in Giwa and Tofa Local Government Areas of Nigeria: GIS approach, *Journal of Scientific Research & Reports* **3**(22).
- Ismaila, A. and Usul, N. (2013). A GIS-based spatial analysis of health care facilities in Yola, Nigeria, *GEO Processing: The Fifth International Conference on Advanced Geographic Information Systems, Applications, and Services*. Turkey.
- Jacobs, B., Ir, P., Bigdeli, M., Annear, P. L. and Van Damme, W. (2012). Addressing access barriers to health services: an analytical framework for selecting appropriate interventions in low-income Asian countries, *Health Policy and Planning* **27**(4): 288–300.
- Jamtsho, S., Corner, R. and Dewan, A. (2015). Spatio-temporal analysis of spatial accessibility to primary health care in Bhutan, *ISPRS International Journal of Geo-Information* **4**(3): 1584–1604.
- Jia, H., Ordóñez, F. and Dessouky, M. (2007). A modeling framework for facility location of medical services for large-scale emergencies, *IIE Transactions* **39**(1): 41–55.
- Jordan, H., Roderick, P., Martin, D. and Barnett, S. (2004). Distance, rurality and the need for care: access to health services in South West England, *International Journal of Health Geographics* **3**(1): 21.
- Keeton, V., Soleimanpour, S. and Brindis, C. D. (2012). School-based health centers in an era of health care reform: Building on history, *Current Problems in Pediatric and Adolescent Health care* **42**(6): 132–156.
- Kim, D.-G. and Kim, Y.-D. (2013). A lagrangian heuristic algorithm for a public health-care facility location problem, *Annals of Operations Research* **206**(1): 221–240.
- Kokash, N. (2005). An introduction to heuristic algorithms, *Department of Informatics and Telecommunications* pp. 1–8.
- Kress, D. H., Su, Y. and Wang, H. (2016). Assessment of primary health care system performance in Nigeria: Using the primary health care performance indicator conceptual framework, *Health Systems & Reform* **2**(4): 302–318.
- Kroll, T., Jones, G. C., Kehn, M. and Neri, M. T. (2006). Barriers and strategies affecting the utilisation of primary preventive services for people with physical disabilities: a qualitative inquiry, *Health & Social care in the Community* **14**(4): 284–293.
- Kuponiyi, O. T., Amoran, O. E. and Kuponiyi, O. T. (2016). School health services and its practice among public and private primary schools in western Nigeria, *BMC Research Notes* **9**(1): 203.
- Lagos State Bureau of Statistics (2017). Digest of Statistics 2017, Lagos State Government, accessed 03 September 2018.
URL: lagostate.gov.ng

- Lagos State Ministry of Health (2015a). Infrastructure and equipment, accessed 15 January 2016.
URL: <http://www.lsmoh.com/infrastructure.php>
- Lagos State Ministry of Health (2015b). State health facility monitoring and accreditation agency (HEFAMAA), accessed 30 September 2015.
URL: <http://www.lagosstateministryofhealth.com/agency>
- LaMondia, J. J., Blackmar, C. E. and Bhat, C. R. (2010). Comparing transit accessibility measures: A case study of access to healthcare facilities.
- Lei, T. L. and Church, R. L. (2014). Vector assignment ordered median problem: A unified median problem, *International Regional Science Review* **37**(2): 194–224.
- Lei, T. L., Church, R. L. and Lei, Z. (2016). A unified approach for location-allocation analysis: integrating GIS, distributed computing and spatial optimization, *International Journal of Geographical Information Science* **30**(3): 515–534.
- Leknes, H., Aartun, E. S., Andersson, H., Christiansen, M. and Granberg, T. A. (2017). Strategic ambulance location for heterogeneous regions, *European Journal of Operational Research* **260**(1): 122–133.
- Levesque, J.-F., Harris, M. F. and Russell, G. (2013). Patient-centred access to health care: conceptualising access at the interface of health systems and populations, *Int J Equity Health* **12**(1): 18.
- Li, Q., Zhang, T., Wang, H. and Zeng, Z. (2011). Dynamic accessibility mapping using floating car data: A network-constrained density estimation approach, *Journal of Transport Geography* **19**(3): 379–393.
- Liu, H., Silva, E. A. and Wang, Q. (2016). Incorporating GIS data into an agent-based model to support planning policy making for the development of creative industries, *Journal of Geographical Systems* **18**(3): 205–228.
- Liu, S. and Zhu, X. (2004). Accessibility analyst: an integrated GIS tool for accessibility analysis in urban transportation planning, *Environment and Planning B: Planning and Design* **31**(1): 105–124.
- Love, D. and Lindquist, P. (1995). The geographical accessibility of hospitals to the aged: a geographic information systems analysis within Illinois., *Health Services Research* **29**(6): 629.
- Lu, C.-C. (2013). Robust weighted vertex p-center model considering uncertain data: An application to emergency management, *European Journal of Operational Research* **230**(1): 113–121.
- Luo, W. (2004). Using a GIS-based floating catchment method to assess areas with shortage of physicians, *Health & Place* **10**(1): 1–11.

- Luo, W. and Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians, *Health & Place* **15**(4): 1100–1107.
- Luo, W. and Wang, F. (2003). Spatial accessibility to primary care and physician shortage area designation: a case study in Illinois with GIS approaches, *Geographic Information Systems and Health Applications*, IGI Global, pp. 261–279.
- Ma, L., Luo, N., Wan, T., Hu, C. and Peng, M. (2018). An improved healthcare accessibility measure considering the temporal dimension and population demand of different ages, *International Journal of Environmental Research and Public Health* **15**(11): 2421.
- Macal, C. M. (2005). Model verification and validation, *Workshop on: Threat Anticipation: Social Science Methods and Models*.
- Macal, C. M. and North, M. J. (2006). Introduction to agent-based modeling and simulation, *MCS LANS Informal Seminar*.
- Macal, C. M. and North, M. J. (2010). Tutorial on agent-based modelling and simulation, *Journal of simulation* **4**(3): 151–162.
- MacKinney, C., Coburn, A., Lundblad, J. P., McBride, T. D., Mueller, K. J. and Watson, S. D. (2014). Access to rural health care - a literature review and new synthesis, *Rural Policy Research Institute*.
- Maguire, D. J. (1991). An overview and definition of GIS, *Geographical Information Systems: Principles and Applications* **1**: 9–20.
- Mahabir, R., Stefanidis, A., Croitoru, A., Crooks, A. T. and Agouris, P. (2017). Authoritative and volunteered geographical information in a developing country: A comparative case study of road datasets in Nairobi, Kenya, *ISPRS International Journal of Geo-Information* **6**(1): 24.
- Makinde, O. A., Sule, A., Ayankogbe, O. and Boone, D. (2018). Distribution of health facilities in Nigeria: implications and options for universal health coverage, *The International Journal of Health Planning and Management* **33**(4): e1179–e1192.
- Malczewski, J. (1999). *GIS and multicriteria decision analysis*, John Wiley & Sons.
- Marceau, D. J. and Benenson, I. (2011). *Advanced Geo-Simulation Models*, Bentham Science Publishers.
- McColl, M. A. (2005). Disability studies at the population level: Issues of health service utilization, *American Journal of Occupational Therapy* **59**(5): 516–526.
- McLafferty, S. L. (2003). GIS and health care, *Annual Review of Public Health* **24**(1): 25–42.

- Mehretu, A. (1985). A spatial framework for redressing disparities in rural service delivery systems, *Tijdschrift voor Economische en Sociale Geografie* **76**(5): 363–373.
- Mesgari, M. and Masoomi, Z. (2008). GIS applications in public health as a decision making support system and its limitation in Iran, *World appl Sci J (Supple 1)* **3**: 73–77.
- Meskarian, R., Penn, M. L., Williams, S. and Monks, T. (2017). A facility location model for analysis of current and future demand for sexual health services, *PloS one* **12**(8): e0183942.
- Mestre, A. M., Oliveira, M. D. and Barbosa-Póvoa, A. P. (2015). Location–allocation approaches for hospital network planning under uncertainty, *European Journal of Operational Research* **240**(3): 791–806.
- Mosadeghrad, A. M. (2012). A conceptual framework for quality of care, *Materia Socio-medica* **24**(4): 251.
- Moulay, D., Aziz-Alaoui, M. and Cadivel, M. (2011). The chikungunya disease: modeling, vector and transmission global dynamics, *Mathematical Biosciences* **229**(1): 50–63.
- Moulin, B., Chaker, W., Perron, J., Pelletier, P., Hogan, J. and Gbei, E. (2003). MAGS Project: Multi-agent geosimulation and crowd simulation, *International Conference on Spatial Information Theory*, Springer, pp. 151–168.
- Mulrooney, T., Beratan, K., McGinn, C. and Branch, B. (2017). A comparison of raster-based travel time surfaces against vector-based network calculations as applied in the study of rural food deserts, *Applied Geography* **78**: 12–21.
- Murad, A. (2018). Using GIS for determining variations in health access in Jeddah city, Saudi Arabia, *ISPRS International Journal of Geo-Information* **7**(7): 254.
- Murawski, L. and Church, R. L. (2009). Improving accessibility to rural health services: The maximal covering network improvement problem, *Socio-Economic Planning Sciences* **43**(2): 102–110.
- Nelson, A. (2000). Accessibility, transport and travel time information, *CIAT Hillside Project Report*.
- Nigera Communications Week (2018). NCC says 40m Nigerians out of internet coverage.
URL: <https://nigeriacommunicationsweek.com.ng/ncc-says-40m-nigerians-out-of-internet-coverage/>
- Noor, A. M., Amin, A. A., Gething, P. W., Atkinson, P. M., Hay, S. I. and Snow, R. W. (2006). Modelling distances travelled to government health services in Kenya, *Tropical Medicine & International Health* **11**(2): 188–196.
- Nwakeze, N. M. and Kandala, N.-B. (2011). The spatial distribution of health establishments in Nigeria, *African Population Studies* **25**(2).

- Ogege, L. (2017). *A GIS approach for the analysis of current traffic situation, and transport system in Lagos, Nigeria.*, Master's thesis, Aalto University Finland.
URL: <http://urn.fi/URN:NBN:fi:aalto-201712188210>
- Oladigbolu, R., Oche, M., Kaoje, A., Gana, G. and Makusidi, M. (2018). Out-of-pocket spending and access to healthcare services in Sokoto, Nigeria, *International Journal of TROPICAL DISEASE & Health* pp. 1–13.
- Olatubi, M. I., Oyediran, O. O., Adubi, I. O. and Ogidan, O. C. (2018). Health care expenditure in Nigeria and national productivity: A review, *South Asian Journal of Social Studies and Economics* pp. 1–7.
- Oliver, A. and Mossialos, E. (2005). European health systems reforms: Looking backward to see forward?, *Journal of Health Politics, Policy and Law* **30**(1-2): 7–28.
- Olowofoyeku, O., Jethro, S., Lipika, D. and Eric, G. (2019). Healthcare facility coverage for malaria and sickle cell disease treatment, *International Journal of Health, Wellness, and Society* . (In press).
- Olowofoyeku, O. O. (2018). Vector-based geosimulation travel time model for health-care facilities catchment, [Poster], ACM-W UK 3rd INSPIRE conference, 20 April.
- Oshodi, L. (2010). Housing, population and development in Lagos, Nigeria, [Blog] International Development, Urban Infrastructure and Governance. Accessed July 30 2017.
URL: <https://oshlookman.wordpress.com/2010/11/24/housing-situation-in-lagos-nigeria>
- O'Sullivan, D. (2008). Geographical information science: agent-based models, *Progress in Human Geography* .
- Penchansky, R. and Thomas, J. W. (1981). The concept of access: definition and relationship to consumer satisfaction, *Medical Care* **19**(2): 127–140.
- Persson, F., Erlandsson, D., Larsson, A. and Johansson, M. (2013). Solving location problems with simulation modelling: a case from the construction industry, *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*, IEEE Press, pp. 3363–3374.
- Peters, D. H., Garg, A., Bloom, G., Walker, D. G., Brieger, W. R. and Hafizur Rahman, M. (2008). Poverty and access to health care in developing countries, *Annals of the New York Academy of Sciences* **1136**(1): 161–171.
- Pisinger, D. and Ropke, S. (2010). Large neighborhood search, *Handbook of Metaheuristics*, Springer, pp. 399–419.
- Pourabdollah, A., Morley, J., Feldman, S. and Jackson, M. (2013). Towards an authoritative OpenStreetMap: conflating OSM and OS OpenData national maps road network, *ISPRS International Journal of Geo-Information* **2**(3): 704–728.

- Rahman, S.-u. and Smith, D. K. (2000). Use of location-allocation models in health service development planning in developing nations, *European Journal of Operational Research* **123**(3): 437–452.
- Ray, N. and Ebener, S. (2008). AccessMod 3.0: computing geographic coverage and accessibility to health care services using anisotropic movement of patients, *International Journal of Health Geographics* **7**(1): 63.
- ReVelle, C., Murray, A. T. and Serra, D. (2007). Location models for ceding market share and shrinking services, *Omega* **35**(5): 533–540.
- ReVelle, C. S. and Swain, R. W. (1970). Central facilities location, *Geographical Analysis* **2**(1): 30–42.
- Rodrigue, J.-P., Comtois, C. and Slack, B. (2013). *The geography of transport systems*, Routledge.
- Safar, M. and Shahabi, C. (1999). 2D topological and direction relations in the world of minimum bounding circles, *Database Engineering and Applications, 1999. IDEAS'99. International Symposium Proceedings*, IEEE, pp. 239–247.
- Salari, M. (2014). An iterated local search for the budget constrained generalized maximal covering location problem, *Journal of Mathematical Modelling and Algorithms in Operations Research* **13**(3): 301–313.
URL: <https://doi.org/10.1007/s10852-013-9233-9>
- Salau, T. (2015). Public transportation in metropolitan Lagos, Nigeria: analysis of public transport users socioeconomic characteristics, *Urban, Planning and Transport Research* **3**(1): 132–139.
- Schuurman, N., Fiedler, R. S., Grzybowski, S. C. and Grund, D. (2006). Defining rational hospital catchments for non-urban areas based on travel-time, *International Journal of Health Geographics* **5**(1): 43.
- Shannon, G. W., Spurlock, C. W., Gladin, S. T. and Skinner, J. L. (1975). A method for evaluating the geographic accessibility of health services*, *The Professional Geographer* **27**(1): 30–36.
- Shariff, S. R., Moin, N. H. and Omar, M. (2012). Location allocation modeling for healthcare facility planning in Malaysia, *Computers & Industrial Engineering* **62**(4): 1000–1010.
- Shaw, P. (1998). Using constraint programming and local search methods to solve vehicle routing problems, *International Conference on Principles and Practice of Constraint Programming*, Springer, pp. 417–431.
- Siebers, P.-O. and Aickelin, U. (2008). Introduction to multi-agent simulation, *Encyclopedia of Decision Making and Decision Support Technologies*, IGI Global, pp. 554–564.

- Snyder, L. V. (2006). Facility location under uncertainty: a review, *IIE Transactions* **38**(7): 547–564.
- Stummer, C., Doerner, K., Focke, A. and Heidenberger, K. (2004). Determining location and size of medical departments in a hospital network: A multiobjective decision support approach, *Health care Management Science* **7**(1): 63–71.
- Su, M.-D., Lin, M.-C., Hsieh, H.-I., Tsai, B.-W. and Lin, C.-H. (2010). Multi-layer multi-class dasymetric mapping to estimate population distribution, *Science of the Total Environment* **408**(20): 4807–4816.
- Swain, R. (1971). *A decomposition algorithm for a class of facility location algorithms*, PhD thesis, Cornell University, Ithaca, N.Y.
- The World Bank (2010). *Improving Primary Health Care Delivery in Nigeria*, The World Bank.
URL: <https://elibrary.worldbank.org/doi/abs/10.1596/978-0-8213-8311-7>
- The World Bank (2015). Nigeria | data, accessed 05 November 2015.
URL: <http://data.worldbank.org/country/nigeria>
- The World Bank (2017). World Bank and WHO: half the world lacks access to essential health services, 100 million still pushed into extreme poverty because of health expenses, accessed 15 August 2019.
URL: <https://www.worldbank.org/en/news/press-release/2017/12/13/world-bank-who-half-world-lacks-access-to-essential-health-services-100-million-still-pushed-into-extreme-poverty-because-of-health-expenses>
- The World Bank (2018). Countryprofile | world development indicators, accessed 15 August 2019.
URL: <https://databank.worldbank.org/views/reports/reportwidget.aspx?ReportName=CountryProfileId=b450fd57tbar=ydd=yinf=nzm=ncountry=NGA>
- Thiede, M., Akweongo, P., McIntyre, D., McIntyre, D. and Mooney, G. (2007). Exploring the dimensions of access, *The Economics of Health Equity* pp. 103–123.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region, *Economic Geography* **46**(sup1): 234–240.
- Toregas, C., Swain, R., ReVelle, C. and Bergman, L. (1971). The location of emergency service facilities, *Operations Research* **19**(6): 1363–1373.
- United Nations Children’s Fund (UNICEF) (2019). Unicef data: Monitoring the situation of children and women, accessed 15 August 2019.
URL: <https://data.unicef.org/country/nga/>
- United States Agency for International Development (2016). Nigeria - accessibility to EmONC facilities in the State of Cross River, Accessed 2 December 2018.
URL: https://pdf.usaid.gov/pdf_docs/pa00mdwb.pdf

UN Millennium Project | *About the MDGs*. (n.d.).

Uno, T., Katagiri, K. and Kato, K. (2011). A multi-dimensionalization of competitive facility location problems, *Int J Innov Comput Inf Control* **7**: 2593–2601.

Vadrevu, L. and Kanjilal, B. (2016). Measuring spatial equity and access to maternal health services using enhanced two step floating catchment area method (E2SFCA)—a case study of the Indian Sundarbans, *International Journal for Equity in Health* **15**(1): 87.

Varga, C. A. and Veale, D. (1997). Isihlambezo: Utilization patterns and potential health effects of pregnancy-related traditional herbal medicine, *Social Science & Medicine* **44**(7): 911–924.

Vo, A., Plachkinova, M. and Bhaskar, R. (2015). Assessing healthcare accessibility algorithms: A comprehensive investigation of two-step floating catchment methodologies family.

Wang, L., Shi, H. and Gan, L. (2018). Healthcare facility location-allocation optimization for Chinas developing cities utilizing a multi-objective decision support approach, *Sustainability* **10**(12): 4580.

Wang, S. and Ma, S. (2018). Efficient methods for a bi-objective nursing home location and allocation problem: A case study, *Applied Soft Computing* **65**: 280–291.

Wilensky, U. (1999). *NetLogo*, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

URL: <http://ccl.northwestern.edu/netlogo/>

Wojuade, C. A. (2017). Public Transport Pricing in Nigeria, *Advances in Social Sciences Research Journal* **4**(17).

World Health Organization (1978). Declaration of Alma Ata, *International Conference on Primary Health Care Alma-Ata, USSR, 6-12 September, 1978*.

World Health Organization (1998). *District health facilities: Guidelines for development and operations*, WHO regional publications. Western Pacific series ; no.22, Manila : WHO Regional Office for the Western Pacific.

World Health Organization (2006). Constitution of the world health organization.

World Health Organization (2011). The Abuja declaration:ten years on.

World Health Organization (2013). Service availability and readiness assessment (SARA): an annual monitoring system for service delivery: reference manual.

World Health Organization (2015). *Global technical strategy for malaria 2016-2030*, World Health Organization.

- World Health Organization (2017). WHO Malaria Policy Advisory Committee (MPAC) meeting.
URL: <http://apps.who.int/iris/bitstream/handle/10665/259396/WHO-HTM-GMP-2017.21-eng.pdf;jsessionid=38725EA36A92B5F5025EC7C288EC897A?sequence=1>
- World Health Organization (2018a). Update on the E-2020 initiative of 21 malaria-eliminating countries: report and country briefs, (2018 (WHO/CD-S/GMP/2018.13)).
- World Health Organization (2018b). WHO called to return to the Declaration of Alma-Ata, accessed 29 January 2018.
URL: http://www.who.int/social_determinants/tools/multimedia/alma_ata/en/
- Ye, F., Zhao, Q., Xi, M. and Dessouky, M. (2015). Chinese national emergency warehouse location research based on VNS algorithm, *Electronic Notes in Discrete Mathematics* **47**: 61–68.
- Ye, H. and Kim, H. (2016). Locating healthcare facilities using a network-based covering location problem, *GeoJournal* **81**(6): 875–890.
- Yin, C., He, Q., Liu, Y., Chen, W. and Gao, Y. (2018). Inequality of public health and its role in spatial accessibility to medical facilities in China, *Applied Geography* **92**: 50–62.
- Yin, P. and Mu, L. (2012). Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles, *Applied Geography* **34**: 247–254.
- Yuan, M. (1996). Temporal GIS and spatio-temporal modeling, *Proceedings of Third International Conference Workshop on Integrating GIS and Environment Modeling, Santa Fe, NM*.
- Zarandi, M. F., Davari, S. and Sisakht, S. H. (2011). The large scale maximal covering location problem, *Scientia Iranica* **18**(6): 1564–1570.
- Zhang, W., Cao, K., Liu, S. and Huang, B. (2016). A multi-objective optimization approach for health-care facility location-allocation problems in highly developed cities such as Hong Kong, *Computers, Environment and Urban Systems* **59**: 220–230.
- Zhou, Y. (2016). Path finding model using the A-star algorithm in NetLogo, Accessed 16 November 2016.
URL: <http://geospatialcss.blogspot.co.uk/2016/01/path-finding-model-using-a-star.html>
- Zinszer, K., Charland, K., Kigozi, R., Dorsey, G., Kamya, M. R. and Buckeridge, D. L. (2014). Determining health-care facility catchment areas in Uganda using data on malaria-related visits, *Bulletin of the World Health Organization* **92**: 178–186.