International Journal of Mathematical, Engineering and Management Sciences
Vol. 8, No. 1, 43-59, 2023
https://doi.org/10.33889/IJMEMS.2023.8.1.003

Robust Facility Location of Container Clinics: A South African Application

C. Karsten
Department of Industrial and Systems Engineering,
University of Pretoria, Pretoria, Gauteng, South Africa.
CSIR Built Environment,
Pretoria, Gauteng, South Africa.
E-mail: ckarsten@csir.co.za

W. L. Bean
Department of Industrial and Systems Engineering,
University of Pretoria, Pretoria, Gauteng, South Africa.
Corresponding author: wilna.bean@up.ac.za

Q. Van Heerden
CSIR Built Environment,
Pretoria, Gauteng, South Africa.
E-mail: qvheerden@csir.co.za

(Received on July 08, 2022; Accepted on September 02, 2022)

Abstract
There is a lack of dynamic facility location models for developing countries that consider the changes in the problem environment over time, such as patient population and population migration. Therefore, this paper focuses on using optimization and goal programming to locate health care facilities in an uncertain environment using multiple possible future urban development senarios. To achieve this, a robust multi-objective facility location model is developed and used to determine locations for container clinic deployment over multiple years in selected communities in South Africa. A synthetic population and urban growth simulation model are used to estimate population density and distribution from 2018 to 2030 for three development senarios. The results from the urban growth simulation model are then used as input into the facility location model to locate facilities whilst considering the three future development scenarios. Results of the model indicate that the robust model can be used to find locations that provide a relatively good solution to all considered development scenarios, providing key role players with quantitative decision support during network design under uncertainty. An accessibility analysis investigates the impact of the prescribed accessibility percentage on model results and a budget analysis evaluates the impact of a case that includes a budget constraint. From these two analyses it is illustrated that the model is sensitive to changes in parameters and that the model can be used by key stakeholders to combine network design and urban development planning for improved decision making.

Keywords- Facility location, Goal programming, Mobile clinics, Genetic algorithms, Optimization, Multiple objectives.

1. Introduction

Healthcare, especially access to healthcare, is a key metric for countries. In 2017, 9% of South Africa's Gross Domestic Product (GDP) was spent on healthcare. Despite this being 5% higher than recommended by the World Health Organization for a country of its socio-economic status, the country's health outcomes are poor when compared to similar countries. Two of the reasons for this imbalance are inequalities between the public and private healthcare systems and restricted access to healthcare in some communities (Africa Health, 2019).

There are significant differences in households 'proximity to a healthcare facility between rural and urban areas as well as socio-economic groups (Booysen, 2003; Ali et al., 2018). Many people using public healthcare are located far from hospitals or clinics, even within urban areas. According to Yantzi et al. (2001) the distance to hospitals and clinics is a crucial factor when selecting health services and whether to visit health facilities. This was also found by Zimmerman (1997) who highlights the convenience of access as a critical factor when deciding whether or not to visit a healthcare facility. In addition, Ali et al. (2018) confirms this relationship and highlight that there is a strong association between the distance to healthcare facilities and pregnant women's utilization of antenatal care at these facilities. These studies emphasize how the distance to healthcare facilities can often become a barrier to the use of healthcare services and increasing accessibility to healthcare should be a priority.

To make healthcare more accessible, mobile clinics can be used. Mobile clinics are customized vehicles that travel into communities to provide immediate but transient healthcare to people in these communities (Hill et al., 2014). Another alternative for providing accessible healthcare is container clinics, i.e., shipping containers converted into clinics. This is a more permanent alternative but without the extensive cost implications. Investing in container clinics rather than mobile clinics can provide a sense of security since the community knows that it will not disappear overnight and can be accessed regularly.

When a metro decides to invest in container clinics, efficient facility locations may improve utilization of clinics whilst reducing the distribution network cost to these facilities (Afshari and Peng, 2014). However, planning for the longer-term location of facilities whilst considering uncertain demand and population growth can be a challenge and lead to suboptimal solutions. One possible way to address this is to combine network planning and urban growth modelling during planning to assist with improved strategic decision making which can ultimately lead to a more robust network design. Network robustness is a crucial strategic consideration for all organizations (Graham et al., 2015). A robust network's configuration can ensure that the network's performance level or accessibility level stays at the desired level irrespective of changes in the customer base, often due to changes in population growth rates or urbanization.

A robust container clinic location plan can assist the South African government to improve access to healthcare for South Africans living in communities with little to no access to healthcare facilities. Therefore, there is a need to determine when and where to locate container clinics in lower- to medium-income communities in South Africa, whilst considering various future population growth and metropolitan development scenarios. By combining future urban growth and network design models, a link between the domains of urban planning and facility location can be made and it can be demonstrated how the output from urban growth planning and modelling can be used to facilitate robust network design.

To this end, the aims and objectives of this research is to develop and solve a robust multi-objective facility location optimization model that considers expected future population growth and development scenarios in lower-to-medium income communities in South Africa

To find a robust solution for container clinic locations in South Africa, a facility location problem is modeled for three large metropolitans in Gauteng, South Africa, whilst considering multiple possible future population growth and development scenarios for these metros. To achieve this, literature is scrutinised to determine which household attributes can be used to predict healthcare demand. This information is then used to convert the available and relevant household attributes for the three metros into healthcare demand. The estimated demand is then used as input into multi-objective facility location models to determine where container clinics should be located and when they should be opened, based on different urban growth scenarios. Finally, accessibility and budget analyses are conducted to evaluate the model's usability in various instances.

The rest of the paper is organized as follows. The model and method are discussed in Section 2, with results and additional analyses presented in Sections 3 and 4, respectively. Finally, the conclusions are presented in Section 5 with a future research agenda.

2. Model and Method

2.1 Urban Growth Simulation Modelling

Gauteng is home to just more than a quarter of the South African population (Stats SA, 2020), with the City of Tshwane, City of Johannesburg, and Ekurhuleni the three metros with the highest population density in the province (refer to Figure 1). This research focuses on these three metros since, the proportion of the population considered can be maximized whilst keeping the case study area relatively small for in-depth analysis

In South Africa, municipalities plan for future development based on a master plan that includes an Integrated Development Plan and a Metropolitan Spatial Development Framework. Possible revenues, expenses and development projects are a few of the topics included in these plans. These development projects are either housing or job creation projects and the deployment sequence of these projects affects the growth of municipalities. Although the municipalities have these development plans, not all projects realize for numerous reasons, such as budget constraints and shifts in focus or importance. The realization of deployment strategies ultimately determines the development patterns of the municipality and the city form.

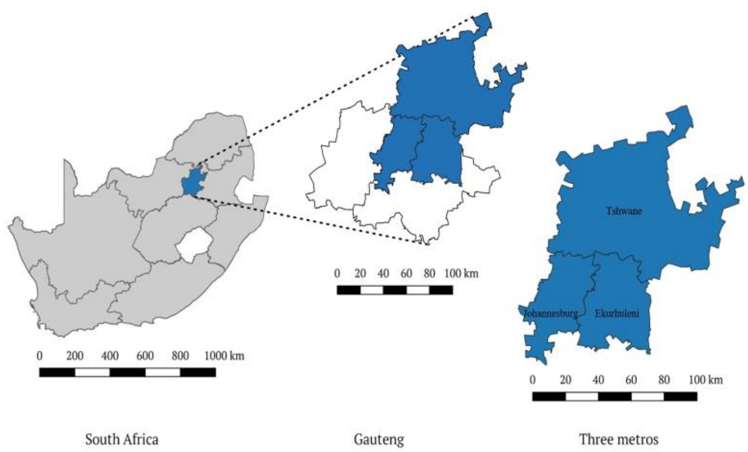


Figure 1. Metros considered in the research.

Urban growth simulation modelling is used to test the likely outcomes of development project deployment strategies in South Africa. These models show possible future city form scenarios based on developments as well as other models such as transportation and location choice models. To this end, the CSIR implemented an urban growth simulation model (UrbanSim) based on synthetic population data for these three metros in Gauteng (Waldeck et al., 2020). This model was used to predict urban growth and development in the three metros from 2019 to 2030. The following three scenarios were developed and tested with various development project alternatives and schedules in UrbanSim, these are also the three scenarios considered in this research:

- **Trend scenario:** This scenario is a continuation of the current development trend and increases the number of households as well as the number of jobs over time.

- **Economic spike scenario:** This scenario is based on the trend scenario with one exception, the development focus is on creating employment opportunities.

- **Relocation scenario:** This scenario is also based on the trend scenario, but more housing development projects are implemented to move households closer to job opportunities.

Demand for healthcare is closely related to the health-seeking behaviour of individuals (Sarma, 2009). Availability and affordability are two of the critical decision influencers related to seeking healthcare, especially in developing countries. According to O'donnell (2007), the effect of affordability on healthcare demand is more pronounced in developing countries, because a large portion of the population has a lower income and no medical aid. Various studies found a strong positive correlation between household income and the probability of seeking healthcare (Mwabu et al., 1993; Abera Abaerei et al., 2017; Ali et al., 2018; Paul and Chouhan, 2020). Even when free healthcare is provided, the monetary and time cost of travel to health facilities is sometimes seen as a healthcare cost and can be a reason to refrain from seeking healthcare, especially when the facility is too far away (Nteta et al., 2010; McLaren et al., 2014). This is confirmed by various studies that found a negative correlation between seeking medical care and the distance to the healthcare facility (Mwabu et al., 1993; Hotchkiss, 1998; Booysen, 2003; Ali et al., 2018).

Another factor identified is the age of the patient, with various studies confirming a correlation between age and the likelihood of seeking medical care (Nteta et al., 2010; Masiye and Kaonga, 2016). This is confirmed by Abera Abaerei et al. (2017) who found that a one-year increase in age increases the odds of seeking medical care by 2%.

Some other attributes that also contribute to healthcare utilization were found in the literature, i.e. the number of children in a household (Nteta et al., 2010), the perceived quality of healthcare (Wellay et al., 2018), gender (Abera Abaerei et al., 2017), employment status (Wabiri et al., 2016), etc. However, factors with the strongest correlation to the likelihood of seeking healthcare are distance to the facility and affordability. Other factors such as age, gender, education, and employment status also have an impact, but this impact is generally less significant and will therefore not be considered in this research.

For this research, two datasets are used. The first is a set of all the existing public clinics and hospitals in the three metros under investigation. This data is required to calculate the initial accessibility measures. The existing public healthcare facilities are used as the base facilities and all the container clinics opened are added to this set. The second dataset required is the synthetic household distribution and attribute data. The household data is necessary for the primary healthcare demand calculation in the model. Based on investigated literature, the household income, the number of children, distance to nearest healthcare facility and household size are used to determine healthcare demand. To convert the data into healthcare demand, the following attributes and probabilities are used:

- **Number of children:** From the literature, it was found that children are more likely to visit a health care facility. Therefore, in this study the probability of visiting a health care facility when ill is set to 0.85 for households with one or more children, whereas this probability is reduced to 0.75 for households without children.

- **Household income:** Each income class is allocated a probability of visiting a health facility when ill as indicated in Table 1. These probabilities are in line with findings from the literature.

- **Distance to the nearest facility:** For all households, the probability of visiting a clinic within a 5 km radius is set to 1. This probability decreases exponentially with the increase in distance. Following the distance decay function used by Mitropoulos et al. (2013), the distance decay function used for this research is given by $y = 0.95^x$, where x is the distance to the nearest clinic, and y is the probability of visiting a clinic when ill.

Table 1. Probability of individual visiting a clinic when ill based on annual household income.

Income class	Household income per year (ZAR)	Probability to visit health facility when ill
1	0 - 9 600	0.6
2	96 601 - 42 000	0.7
3	42 000 - 108 000	0.75

To determine the healthcare demand per household, these probabilities are multiplied by the household size. This demand per household is aggregated to a total demand per zone used in the location model.

2.2 Facility Location Modelling

To identify vacant land, the combined area of all three metros is divided into 28 461, mostly homogeneous, square zones of approximately 1km^2 in size, with each zone consisting of several parcels of varying size (illustrated in Figure 2).

Parcels are classified according to their underlying land use and a parcel could either be built-up (having one or more buildings present) or vacant (having no buildings). The built-up parcels can further be classified as commercial or residential, depending on the building's use. Any built-up parcels cannot be developed, whereas vacant parcels can be developed and, therefore, only these parcels are considered when calculating the vacant area in the zone. Based on the typical container footprint covering approximately 28m^2, vacant land of at least 35m^2 is deemed sufficient to locate a typical container clinic (Cooke et al., 2010). To determine the vacant area in a zone, the areas of all vacant parcels in the zone are aggregated. All zones with enough vacant space for container clinics are included in the set of candidate locations for the facility location model. The locations of these zones within the three metros are indicated in blue in Figure 2. A pharmaceutical distribution centre, illustrated with a red dot, is used in this study. For modelling purposes, all deliveries to the container clinics are made from this distribution centre.

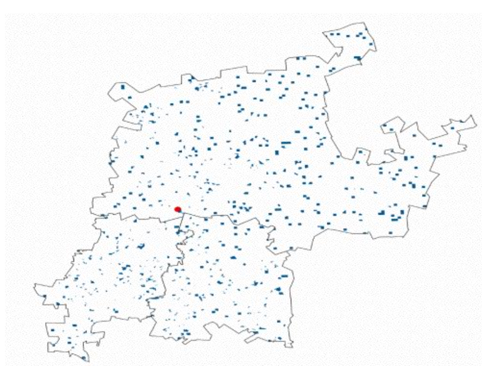


Figure 2. Candidate location zones.

This paper focuses on discrete facility location modelling since these models are most often used in the healthcare context. In addition, these models have fixed sets of locations where facilities can be placed and assume that individual demand points can be grouped into selected discrete demand points. This enables the modeller to represent a geographical area with fewer demand points making it a suitable modelling approach. Discrete location modelling is divided into, amongst others, covering-based and median-based problems (Meskarian et al., 2017).

When considering the types of problems investigated in the literature related to healthcare, the majority focus on median-based location problems. Mestre et al. (2015) applied a p-median model that locates facilities whilst constraining the allowable distance to a health care facility in an uncertain demand environment. Beheshtifar and Alimohammadi (2014) defined a p-median problem with 4 466 demand points and 100 candidate sites to find the optimal locations for clinics whilst minimizing transportation cost and land costs. Kim and Kim (2013) determined locations for public health care facilities within a given budget, whilst maximizing the number of patients served in both private and public facilities. Das et al. (2020) used a p-facility location model to place new facilities amongst existing facilities whilst minimizing the total transportation cost.

The problem considered in this paper is similar to these p-median problems, intending to minimize cost and improve accessibility. In the majority of these p-median healthcare-related problems, either transportation cost or building cost are minimized. In this paper, the model aims to locate container clinics whilst minimizing building cost, patient travel distance and total distribution distance, simultaneously. To achieve these two models are used: 1) a facility location model for each development scenario, and 2) a robust facility location model combining all three scenarios. The first model determines the best network configuration, i.e., when and where to locate container clinics from a set of available locations, for each urban development scenario in isolation, whereas the robust facility location model determines a good configuration for all three scenarios combined. In essence, this model finds a configuration as close as possible to the individual model results of the three scenarios.

During modelling it is assumed that all deliveries of supplies to clinics are made from the distribution center indicated in Figure 2. In addition, since 100% accessibility is not necessarily feasible, a 90% accessibility level is selected as an accessibility target that must be reached within the first five years to address the lack of accessibility as soon as possible. No operating costs are considered in the model. The model only considers the cost of opening the clinics. By not including operating costs, clinics will be built as soon as possible as only the initial construction cost is minimized. The building cost is however increased by 3.5% each year to incorporate building inflation. In addition, it is assumed that all candidate locations have the same building cost and straight-line distances multiplied by a crow-fly factor of 1.265 are used as the distances between facilities and zones (Barthelemy, 2011). Using this distance rather than actual road distances implies that some "accessible" facilities may not fill within 5 km travel distance due to geographical aspects, such as mountains or rivers, not being part of the calculation. However, comprehensive road networks are not necessarily available in all areas and straight-line distances with a crow-fly factor circumvent this problem whilst reducing the problem size and complexity.

To formulate the model, we define the following sets: **E** the set of existing clinics, **L** the set of candidate locations, **S** the set of scenarios, **Y** the number of years included in planning, and **R** the set of household zones. For the individual scenarios, two objective functions are minimized. The first is the total building cost calculated in Equation (1), where b_y represents the building cost per year y and w_{ly} a binary variable indicating if the clinic is opened at location l in year y (value 1) or not (value of 0). The second is the total distance travelled by the households to the nearest clinics and the total distance from the distribution centre to all the open clinics combined. The second objective is shown in equation (2), where m_{ri} is the distance (in km) between a household in zone r and a clinic i, n_i is the distance (in km) from the distribution centre to clinic, and x_{riy} is a binary that will be 1 if household in zone r is served by clinic i in year y.

$$min\, z_1 = \sum_{y\in \boldsymbol{Y}} \sum_{l\in \boldsymbol{L}} b_y w_{ly}. \quad (1)$$

$$min\, z_2 = \sum_{y\in \boldsymbol{Y}} \sum_{r\in \boldsymbol{R}} \sum_{i\in \boldsymbol{E},\boldsymbol{L}} x_{riy} m_{ri} + \sum_{y\in \boldsymbol{Y}} \sum_{l\in \boldsymbol{L}} \sum_{i\in \boldsymbol{E},\boldsymbol{L}} (12-y) w_{ly} n_i \quad (2)$$

To ensure that least one clinic is built and that a clinic cannot be built more than once, equations (3) and (4) are included.

$$\sum_{y\in \boldsymbol{Y}} \sum_{l\in \boldsymbol{L}} w_{ly} \geq 1. \quad (3)$$

$$\sum_{y\in \boldsymbol{Y}} w_{ly} \leq 1\, \forall l \in \boldsymbol{L}. \quad (4)$$

Equation (5) ensures that there is a clinic within 5 km of 90% of the households investigated in this study, and to ensure that all households are serviced by at least one clinic, equation (6) is included.

$$P\big(min\big(m_{ri} x_{riy}\big) \leq 5\big) \geq 0.9 \forall r \in \boldsymbol{R}, y \in \boldsymbol{Y}, i \in \boldsymbol{E}, \boldsymbol{L} \quad (5)$$

$$\sum_{y\in \boldsymbol{E},\boldsymbol{L}} x_{riy} \geq 1 \forall r \in \boldsymbol{R}, y \in \boldsymbol{Y} \quad (6)$$

Equation (7) ensures that a clinic can only serve patients if it has been opened that year or in a previous year. The probability of a household member going to a clinic when ill based on the distance to the nearest facility is calculated in equation (8).

$$\sum_{y\in \boldsymbol{Y}} x_{riy} \leq 13 \sum_{y\in \boldsymbol{Y}} w_{iy} - \sum_{y\in \boldsymbol{Y}} \boldsymbol{y} w_{iy}\, \forall r \in \boldsymbol{R}, i, l \in \boldsymbol{L} \quad (7)$$

$$p_{2r} = 0.95^{m_{ri}} \forall r \in \boldsymbol{R}, i \in \boldsymbol{E}, \boldsymbol{L} \quad (8)$$

The calculation of the healthcare demand based on the available household attributes and the distance to the facility is given in equation (9). Finally, binary conditions are imposed in equations (10) and (11).

$$d_{ry} = o_{ry} p_{ry} p_{2r} \forall r \in \boldsymbol{R}, \boldsymbol{y} \in \boldsymbol{Y} \quad (9)$$

$$x_{riy} \in \{0;1\} \forall r \in \boldsymbol{R}, i \in \boldsymbol{E}, \boldsymbol{L} \quad (10)$$

$$x_{riy} \in \{0;1\} \forall r \in \boldsymbol{R}, i \in \boldsymbol{E}, \boldsymbol{L} \quad (11)$$

For the robust model, two objective functions were minimized with a sum of regrets model. The difference between the good values of the scenarios and the objective value of the current configuration is minimized in equation (12), where g_1 is the total cost of the good configuration for a scenario and v_1 is the cost of a configuration for a scenario. The variation between the total distance travelled for the scenarios given and the current configuration is minimized in equation (13), where g_{2f} is the total distance travelled for the good configuration and v_{2f} is the total distance travelled for a configuration per scenario f.

$$min\, z_3 = \sum_{f\in \boldsymbol{S}} |g_{1f} - v_{1f}| \quad (12)$$

$$min\, z_{34} = \sum_{f\in \boldsymbol{S}} \big|g_{2f} - v_{2f}\big| + \sigma(v_{21}, v_{22}, v_{23}) \quad (13)$$

Model Solution

A variety of tools and techniques are available to solve facility location problems. Exact methods or approximate methods can be used. If the network is relatively small, exact methods can be used to find optimal solutions. These methods are sure to find an optimal solution to the problem. If the problem gets too large, exact methods are no longer feasible and approximate methods have to be used to obtain a reasonably good solution in a reasonable time (Talbi, 2009). Facility location problems are considered NP-hard problems with the NP-hardness increasing with the size of the network. The aggregation level influences the size and the NP-hardness of the problem (Cebecauer and Buzna, 2017). Since robust multi-facility location problems are often classified as NP-hard problems and the study area in this research comprises thousands of zones, Genetic Algorithms (GAs) are used to solve these models. GAs is often used to solve facility location problems and have proven to provide good solutions to these problems in a reasonable time (Arostegui et al., 2006; Shariff et al., 2012; Beheshtifar and Alimohammadi, 2014).

The mathematical model is solved using GAs. These models are solved in Python using the GA included in the DEAP package (Fortin et al., 2012). Once good configurations for the scenarios and the robust configuration are determined, the model is verified and validated. The verification of the model is to confirm that the model does what it is supposed to do: find good scenario configurations and robust configurations. The validation determined if the results from the model are valid. The robust configuration is deemed valid for this study if the total difference between the robust configuration and scenario configuration is no more than 25% for each of the scenarios. This robustness level can be adjusted based on the similarity required in the solutions. The smaller the robustness level, the closer the solution has to be to the original solutions.

3. Results and Findings

3.1 Individual Senarios

For each scenario, the good locations for the container clinics and the year they should be opened are identified using the individual scenario model described in Section 3.2. Opening these facilities improves the overall accessibility of lower-income households to primary healthcare. This improvement makes primary healthcare accessible to more households in the considered income classes without having to incur additional transportation costs since the clinics are within walking distance from their houses.

For the **trend scenario**, most of the clinics are opened in the first year to respond to the immediate demand and accessibility target that has to be reached within the first five years. The rest of the clinics are opened as the demand increased over the years. There is a significant increase in accessibility from about 60% to 90% in the first two years. The opening of the initial clinics also led to a large decrease in total household distance travelled as the clinics are now much closer to the households. There is also an increase in the total distribution distance from the distribution centre to all the open clinics as there are 283 more clinics that must be served. From year four onwards, the accessibility fluctuates between 90% and 91%. This accessibility is influenced by the new households and the new clinics opened in that year. The total distance travelled by the households slowly increases as the number of households increases per year based on the scenario population growth and relocation rates.

In **the economic spike scenario,** the majority of clinics are opened in the first year, however, clinics were also opened in all the other years. In this scenario, fewer clinics were opened from year two onwards than in the trend scenario. This decrease in clinics opened is because in the economic spike scenario there is a decrease in the lower-income households and, therefore, a decrease in demand for primary healthcare. There is once again a drastic change in the accessibility and the total distance travelled by households to their nearest clinic in the first two years. After 5 years, the 90% accessibility target is reached and the clinics opened from then onwards maintain this accessibility level. The total distance travelled by households to clinics drops drastically in the first year, after that the total distance travelled by the households to clinics slowly increases each year as the number of households increases due to population growth and household relocation.

Most of the clinics in the **relocation scenario** are also built in the first year as there is an immediate need for clinics and it is the cheapest year to build the needed clinics. The rest are built as the demand creates a need. The clinics built in the first five years are located more towards the periphery of the metros due to the strong spatial expansion focus to urban growth in this scenario. The total cost per year is the highest in year 2 as most of the clinics are opened in that year. The total distance travelled by households increased over the years, even though accessibility remained relatively constant at around 90%. This steady increase over the years is linked to population and urban growth. The total distance travelled from the distribution centre to the open clinics increases each year. This increase is expected as the total number of clinics serviced each year increases. In this scenario, the total distance travelled by households is noticeably higher than in the other two scenarios. This noticeable difference can be attributed to the fact that there are many more households in this scenario than in the other two.

Each scenario has a specific configuration of when and where to open clinics. When comparing the configurations, almost 80% of the clinics are in completely different locations. Therefore, a good solution for one scenario is not necessarily a good solution for another. If decision makers are confident about which specific scenario will play out, they can optimize the most likely scenario and base their decisions on those results. Using the individual scenario solutions works well for each scenario in isolation and the most cost-effective configuration can be determined while reaching the desired level of accessibility. However, if, as in many real-world situations, multiple scenarios could realize, the robust model provides decision makers with evidence-based decision support that considers a combination of different scenarios and not just one in isolation.

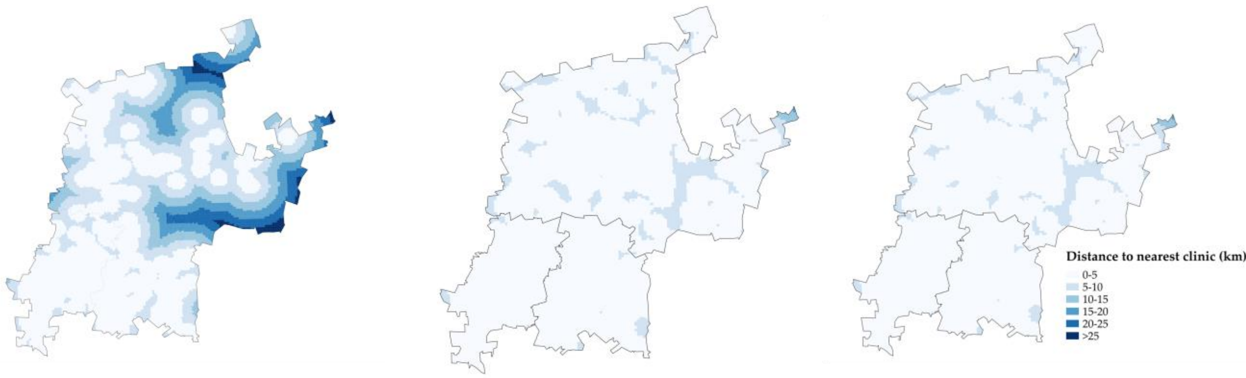
3.2 Robust Solution

The robust model uses goal programming to find an acceptable solution for all three scenarios whilst staying as close as possible to the individual scenario solutions. The objectives of the individual scenarios are set as the goals in the goal programming model. The robust model places facilities over the years whilst minimizing the cost and the total travel distance to come as close to the individual scenarios as possible. It, therefore, seeks the best compromise between the three scenarios.

The number of clinics opened for the robust scenario is more than in the trend and economic spike scenarios, as there is a greater demand in the relocation scenario. Once again, most of the clinics are opened in the first year to meet the immediate need in all the scenarios and no penalty is included in the form of operating costs. A summary of the accessibility improvement over time is provided in Figure 3. Fewer clinics are opened on the outskirts that cater for the relocation scenario demand, however, enough clinics are still in operation that can cater for most of the demand, should it arise. Since the majority of clinics are opened in the first year, the most significant improvement is in the first year. For all the scenarios, the percentage of households within 5km of a clinic given the robust configuration is above 90%. Therefore, the robust configuration adheres to the accessibility constraint in all the scenarios. No households are further than 10km from the nearest clinic in any of the scenarios given the robust configuration and can be reached by foot if necessary.

Configuration Overlap

When comparing the robust solution to the individual scenario solutions, there is a 36% configuration overlap, where clinics are opened in the same location and year, for the trend and economic spike scenarios. For the relocation scenario, almost 40% of the clinics have the same configuration. The similarities between the robust configuration and the scenario configurations are much higher than the similarities between the scenario configurations. Therefore, the robust configuration is a good compromise between the different scenarios. The majority of the overlapping is in the City of Tshwane which is the metro with the lowest accessibility to clinics. The locations of the overlapping zones show that the model caters for this lack of accessibility. The clinics open in the same place every time to cater to this lack of initial accessibility. These overlapping zones should be noted as they are the zones that will improve accessibility. Decision makers can use these overlaps to priorities the clinics to be built when there are time, budget, or other constraints.

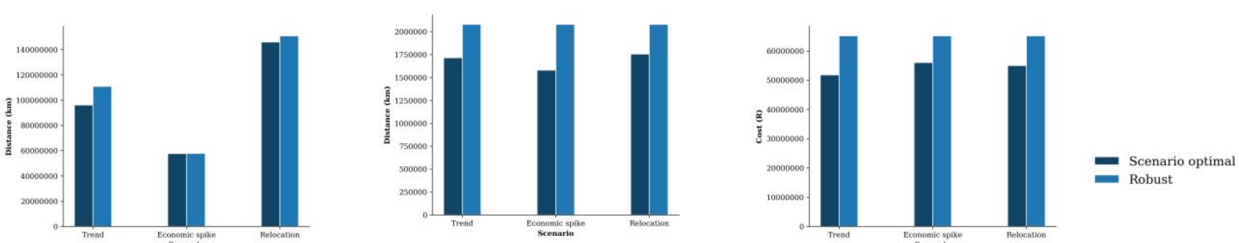


(a) 2018 (b) 2024 (c) 2030

Figure 3. Accessibility improvement with the robust model.

Objective Function Comparison

A comparison of the scenario solutions and the robust solution is provided in Figure 4. The robust values are all the same or higher than the goals. This is expected since the robust configuration is not the optimal configuration for any of the scenarios, but it is a better overall solution. These differences are the cost of having a robust solution, however, the cost of not having a robust solution can be much higher if the individual scenario does not realize.



(a) Travel distance to clinics (b) Travel distance from DC (c) Total cost to open clinics

Figure 4. Comparison of the scenario solutions and the robust solutions per scenario.

The total household distance travelled for the economic spike scenario is the same as for its scenario solution. In the economic spike scenario, the population is denser, creating significant demand points close to city centres where all scenarios have significant demand points. For the other two scenarios, the total distance travelled by households is slightly higher with the combined model than for the individual solutions. This can be attributed to the fact that more households are located towards the peripheries of the municipalities and the population is less dense in these scenarios. If a clinic is not located within 5km of these peripheral demand points, the nearest clinic can be much further away than for households closer to the city centre.

The total distribution distance for the three scenarios is relatively close to each another. The small change can be due to the fact that the distribution centre is located more or less in the centre between the three municipalities. The total distance travelled from the distribution centre to the open clinics for the robust configuration is slightly higher than the scenario solutions. The economic spike scenario has the biggest difference between the scenario solutions and robust total distribution distance since there are not many clinics placed on the outskirts of the municipalities in this scenario keeping the distribution distance low. As more clinics are opened in the combined robust solution than in the individual scenario solutions to cater for the uncertain demand. The total distribution distance for the other two scenarios is also higher than with their individual configurations.

The total costs for the individual scenarios are very similar. This is because the majority of the clinics are opened in the first year, therefore, inflation had a minimal effect on the total cost. The number of clinics opened is the main contributing factor to the total cost. The robust combination model total cost is higher than in any of the scenarios. This cost difference can be attributed to the observation that more clinics are opened in the robust scenario than in the individual scenarios to deal with uncertainty. The difference in the total cost between the scenarios and the robust configuration is due to the years in which the clinics are opened, and the number of clinics opened.

The difference between the robust configuration and the scenario configuration is within a 25% range for all the considered variables. Thus, the robust configuration performs well or at an acceptable level in all scenarios. In this case, if all three variables are considered, the solution is a good solution to combat the uncertainty of how the municipalities will develop in the future. City planners can look specifically at the total household travel distance and the total cost to determine the feasibility of the robust solution, whilst pharmaceutical or logistics companies can do their strategic planning using the total distribution distance.

These results can be used to provide quantitative decision support to key role players when deciding where and when to open new clinics. Trade-offs can be made between the accessibility for communities and the cost of opening new clinics from a healthcare provider perspective or between the average or total travel distance and the total cost of opening the clinics. These trade-offs can be analysed using a Pareto frontier. The Pareto frontier will assist the decision maker to choose a solution most suitable to their strategic needs as the decision maker will be able to see the impact and make trade-offs based on that to ultimately determine a feasible accessibility strategy.

4. Accessibility and Budget Analysis

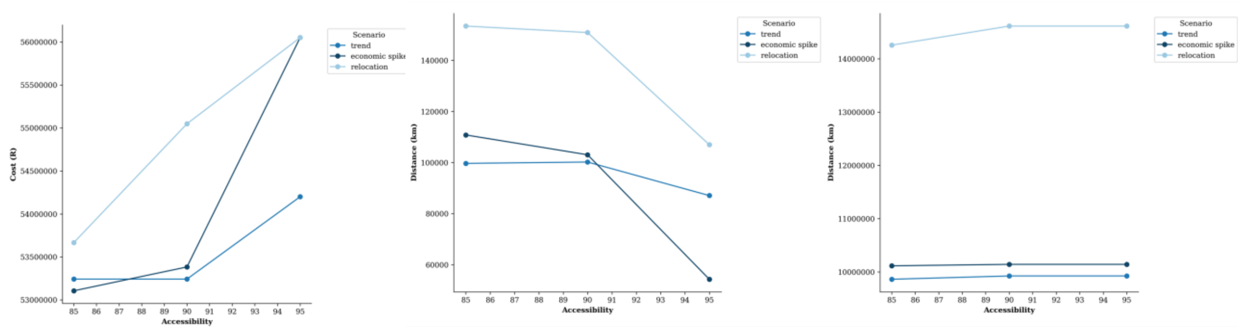
Opening most of the clinics in the first year is not a realistic representation of reality. Therefore, additional analyses are conducted to investigate the impact of different accessibility percentages and the implications of modelling a more realistic scenario on model results.

4.1 Accessibility Analysis

Three accessibility percentages (85%, 90%, and 95%) are used to determine the impact of accessibility on other variables. All the individual models and the robust model are solved for these changing accessibility percentages. Ignoring operating costs in the model skewed the results to open most clinics in the first year to minimize the total cost. However, minimizing the total distribution distance ensures that some clinics, especially on the outskirts, are opened in later years.

Individual Scenarios

Results for the individual scenarios are depicted in Figure 5 and indicate that the total cost increases with increased accessibility percentages, as more clinics are built. The increase in total cost for the relocation scenario from 85% to 90% and from 90% to 95% have similar gradients. However, for the trend and economic spike scenarios, the jump from 90% to 95% is more significant than from 85% to 90%. Moving from 85% accessibility to 90% accessibility could be advisable when considering only the costs since it results in only a slight increase in total cost.



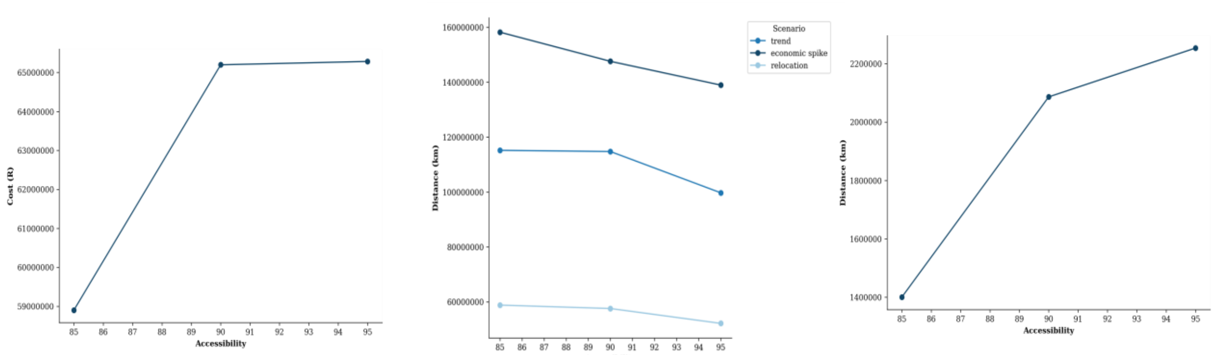
(a) Total cost to open clinics (b) Travel distance to clinics (c) Travel distance from DC

Figure 5. Comparison of the scenario solutions for changing accessibility percentages.

With more clinics being built as the accessibility level is increased, the total distance travelled by households to the nearest clinic reduces as more households have shorter distances to travel to the nearest clinic. Moving from 85% accessibility to 90% accessibility has little impact on the total distance travelled for all three scenarios. However, when increasing the accessibility level to 95%, there is a significant decrease in the total distance travelled by the households, especially in the economic spike scenario and the relocation scenario. This change forces the model to locate clinics in less dense areas significantly reducing the number of households that are further than 10km from the nearest clinic. Therefore, a sharp decrease in the total distance travelled by households is seen. The distance travelled from the distribution centre to clinics increases as the accessibility constraint is increased, mainly due to more clinics that must be serviced.

Robust Solution

When comparing the robust solutions of these different accessibility percentages, the change seen in the individual scenarios is still present as illustrated in Figure 6.



(a) Total cost to open clinics (b) Travel distance to clinics (c) Travel distance from DC

Figure 6. Comparison of the robust solutions for changing accessibility percentages.

A move from 85% to 90% accessibility leads to a massive increase in cost. The cost difference between 90% accessibility and 95% accessibility is much smaller. Therefore, many more clinics are required to move from 85% to 90% accessibility than from 90% to 95%. The feasibility of the large cost difference will have to be weighed against the benefits gained for the community members and other factors.

The total distance travelled by households reduces as the accessibility target increases. This decrease is expected as a smaller portion of the households have to travel more than 5km to the nearest clinic. The benefit of moving from 85% accessibility to 90% is relatively small in all three scenarios and would not justify the significant cost increase. There is a greater impact on the total distance travelled by households when moving from 90% accessibility to 95% accessibility. This improvement will come at little cost given that the accessibility is already at 90%.

The total distance from the distribution centre to the clinics increases as more clinics are opened that have to be serviced as the accessibility target increase. Moving from 85% accessibility to 90% accessibility leads to a steep increase in the total distribution distance. This increase in the distribution distance can impact the prices paid by the patients for medication as the distribution costs are ultimately passed on to them. This large increase in distance can be attributed to the fact that more clinics are opened on the outskirts of the municipalities further away from the distribution centre. Improving the accessibility to 95% has a smaller impact on the total distribution distance as fewer clinics have to be opened to reach it when looking at the total cost.

4.2 Budget Analysis

A budget constraint is added to the model to test if the model is responsive to a restrictive budget, limiting the number of clinics that can be built per year. A random generator is used to set a budget per year. An additional constraint is added to enforce the budget where the total cost for the year has to be less than or equal to the budget for that year.

The accessibility constraint is relaxed to be reached by the end of 2030 and no longer after the first five years as in the original case. By adding the budget constraint, the impact of the omitted operating costs is less significant, forcing the model to open the clinics more gradually. However, some clinics may still be opened sooner than necessary if the budget allows and if it is not far from the distribution centre, keeping the distribution distance as small as possible.

The configuration for each scenario with the adjusted parameters is determined by running the model with the added budget constraint and the relaxed accessibility constraint. In this more realistic case, the relocation scenario still has the most clinics opened to cater for the fast-growing demand, followed by the trend scenario. When comparing the actual costs against the budget, such a large budget is not necessary. A maximum budget of about R 12 000 000 should be sufficient for all the years. By 2030 all the scenarios reached 90% accessibility. The location of the clinics does not change much from the original scenario configurations; it is mostly the year in which the clinics are opened that changes by introducing the budget constraint. The demand is still at the same nodes, and therefore clinics are still required in the same locations as the base case.

The results of the robust solution are provided in Table 2. The robust model also adheres to the budget, forcing a more equal distribution of clinics to be opened over the years. The total cost for the robust model is around 8% higher than the scenario cost in any of the other scenarios, showing that robustness comes at a cost. The difference between the scenario values and the values given the robust configuration is within 10%. This 10% difference between the scenario values and the robust values is better than the 20%

difference in the original case. The smaller difference is because the restricting budget ensures a more even placement of the clinics over the years and not at a surge placement in the first year. Therefore, even with additional constraints and changing parameters, the model can find a robust solution.

The model can be used to provide decision support for the Healthcare Department to identify feasible locations for clinics that will improve the overall accessibility of primary healthcare by finding the best locations and years to open clinics to cope with the growing demand. With a constrained budget, the model provides a strategic plan of when and where to open clinics in order to improve the overall accessibility of healthcare for lower-income households. The available budget for new clinics can be used as input in the models to determine the best robust configuration of clinics to provide the desired level of accessibility.

Table 2. Budget scenario robust configuration yearly solutions.

Year	Budget (ZAR)	Total Cost (ZAR)	Trend scenario total distance to nearest clinic (km)	Economic spike scenario total distance to nearest clinic (km)	Relocation scenario total distance to nearest clinic (km)	Total distance from distribution centre to opened clinics (km)
2018	0	0	16 689 239	16 689 239	16 689 239	0
2019	10 000 000	7 823 499	10 019 787	10 129 638	15 542 152	42 867
2020	11 400 000	9 017 471	8 899 738	89 62 004	13 494 607	20 648
2021	12 898 000	7 999 785	8 572 304	8 553 000	12 727 732	23 265
2022	14 500 860	9 462 603	8 388 241	8 219 435	12 084 745	21 053
2023	16 215 902	7 957 458	8 564 024	8 279 381	12 067 717	25 034
2024	18 051 003	9 503 041	8 618 613	8 341 883	12 120 807	19 806
2025	16 102 006	6 775 668	8 452 000	8 223 803	11 872 826	26 207
2026	15 014 600	11 084 775	8 427 909	8 235 647	11 803 367	27 313
2027	15 014 600	11 472 742	8 214 731	8 036 109	11 505 891	22 793
2028	12 115 620	11 389 623	8 052 017	7 918 599	11 266 171	27 737
2029	14 363 701	14 547 214	8 010 768	7 886 244	11 162 252	27 589
2030	16 706 918	0	8 010 768	7 886 244	11 162 252	0

5. Conclusion

An opportunity was identified to investigate the location of low-cost container clinics in lower-income communities. There is a lack of dynamic location models that consider the changes in the problem environment over time, such as patient population and population migration. This paper, therefore aimed to address this gap by using robust optimization and goal programming to locate health care facilities in an uncertain environment using multiple urban development scenarios.

The research considers three metro municipalities in Gauteng, South Africa. Three future development scenarios were created for this study using a synthetic population and urban growth simulation model, which estimated the population distribution from 2018 to 2030 for all three of the scenarios. Using associative forecasting, the primary healthcare demand was forecasted. These forecasts were used as input into the facility location models with an accessibility target of 90%. The models were used to determine good network configurations for the three scenarios in isolation. Thereafter, the results from each scenario were used as the goals in the goal programming model to determine a robust configuration that will work relatively well for all of the scenarios given the uncertainty of the future development of the municipalities.

Results indicated that the robust model was able to find locations that provided a relatively good solution to all the scenarios, providing key role players with quantitative decision support during network design under uncertain development scenarios. An accessibility analysis was conducted to investigate the impact of the accessibility percentage on the variable values and a budget analysis tested the impact of a more realistic case with a budget constraint. In this case, there was a more gradual placement of the clinics due to the restricting budget. Results of the analyses indicated that the model is sensitive to changes in parameters and that the model can be used for network planning whilst considering different development scenarios, thereby addressing the need for decision support tools integrating network design and future urban development scenarios for health facility placement.

The problem addressed in this paper has similarities with many of the capacitated p-median problems solved in literature. However, in the majority of these p-median healthcare-related problems, either transportation cost or building costs are minimized. In this paper, building costs, transportation costs, and distribution costs are minimized, resulting in a more realistic solution. In addition, none of these health-care related facility location studies considered future urban growth and development scenarios during modelling, whereas this research considered three potential future urban development and growth scenarios, providing a more robust facility location plan for future development.

5.1 Research Contribution

This paper addresses an opportunity in the literature to integrate facility location, especially robust facility location, and urban planning in a South African context. A proof of concept is developed, in the form of a multi-objective optimization model, to determine robust facility locations for container clinics, when urban planners and key role players consider multiple future development scenarios.

5.2 Limitations and Future Research Directions

The research presented in this paper focused on the three metros in Gauteng province, so future research should focus on expanding the study area and considering other provinces and metros in South Africa. In addition, literature indicated that there are numerous factors affecting health care demand, but only the household attributes provided by the UrbanSim model were considered in this paper and alternative attribute can be investigated. Additional future research opportunities include removing some of the model assumptions and adding other elements to enhance the model. For example, operating costs can be included to provide a model that better represents the real world. In this case, the model will no longer place most of the clinics in the first year as there will be a penalty in the form of operating costs if a clinic is opened before it is needed. Variable building costs can also be considered based on the location of the clinic as land costs differ based on location. Finally, different health care facility types can be considered with various capacities to determine a robust location and mix.

Conflict of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

Abera Abaerei, A., Ncayiyana, J., & Levin, J. (2017). Health-care utilization and associated factors in Gauteng province, South Africa. *Global Health Action*, *10*(1), 1305765.

Africa Health (2019). *Industry insights: South Africa healthcare market overview*. Africa Health.

Afshari, H., & Peng, Q. (2014). Challenges and solutions for location of healthcare facilities. *Industrial Engineering and Management*, *3*(2), 1-12.

Ali, S.A., Dero, A.A., Ali, S.A., & Ali, G.B. (2018). Factors affecting the utilization of antenatal care among pregnant women: A literature review. *Journal of Pregnancy and Neonatal Medicine*, *2*(2), 41-45.

Arostegui, M.A., Kadipasaoglu, S.N., & Khumawala, B.M. (2006). An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. *International Journal of Production Economics*, *103*(2), 742-754.

Barthelemy, M. (2011). Spatial networks. *Physics Reports*, *499*(1-3), 1-101.

Beheshtifar, S., & Alimohammadi, A. (2014). A multiobjective optimization approach for location-allocation of clinics. *International Transactions in Operational Research*, *22*(2), 312-328.

Booysen, F. (2003). Urban-rural inequalities in healthcare delivery in South Africa. *Development Southern Africa*, *20*(5), 659-673.

Cebecauer, M., & Buzna, L. (2017). A versatile adaptive aggregation framework for spatially large discrete location-allocation problems. *Computers & Industrial Engineering*, *111*, 364-380.

Cooke, G.S., Tanser, F.C., Barnighausen, T.W., & Newell, M. (2010). Population uptake of antiretroviral treatment through primary care in rural South Africa. *BMC Public Health*, *10*(1), 585. https://doi.org/10.1186/1471-2458-10-585.

Das, S.K., Roy, S.K., & Weber, G.W. (2020). An exact and a heuristic approach for the transportation-p-facility location problem. *Computational Management Science*, *17*(3), 389-407.

Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., & Gagne, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, *13*, 2171-2175.

Graham, G., Mehmood, R., & Coles, E. (2015). Exploring future cityscapes through urban logistics prototyping: a technical viewpoint. *Supply Chain Management: An International Journal*, *20*(3), 341-352.

Hill, C.F., Powers, B.W., Jain, S.H., Bennet, J., Vavasis, A., & Oriol, N.E. (2014). Mobile health clinics in the era of reform. *American Journal of Managed Care*, *20*(3), 261-264.

Hotchkiss, D.R. (1998). The tradeoff between price and quality of services in the Philippines. *Social Science & Medicine*, *46*(2), 227-242.

Kim, D.G., & Kim, Y.D. (2013). A Lagrangian heuristic algorithm for a public healthcare facility location problem. *Annals of Operations Research*, *206*(1), 221-240.

Masiye, F., & Kaonga, O. (2016). Determinants of healthcare utilisation and out-of-pocket payments in the context of free public primary healthcare in Zambia. *International Journal of Health Policy and Management*, *5*(12), 693-703.

McLaren, Z.M., Ardington, C., & Leibbrandt, M. (2014). Distance decay and persistent healthcare disparities in South Africa. *BMC Health Services Research*, *14*(1), 541. https://doi.org/10.1186/s12913-014-0541-1.

Meskarian, R., Penn, M.L., Williams, S., & 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., & Barbosa-Povoa, A.P. (2015). Location-allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research*, *240*(3), 791-806.

Mitropoulos, P., Mitropoulos, I., & Giannikos, I. (2013). Combining DEA with location analysis for the effective consolidation of services in the health sector. *Computers & Operations Research*, *40*(9), 2241-2250.

Mwabu, G., Ainsworth, M., & Nyamete, A. (1993). Quality of medical care and choice of medical treatment in Kenya: An empirical analysis. *Journal of Human Resources*, *28*(4), 838-862.

Nteta, T.P., Mokgatle-Nthabu, M., & Oguntibeju, O.O. (2010). Utilization of the primary health care services in the Tshwane region of Gauteng province, South Africa. *PloS one*, *5*(11), e13909.

O'donnell, O. (2007). Access to healthcare in developing countries: breaking down demand side barriers. *Cadernos de Saude Publica*, *23*(12), 2820-2834.

Paul, P., & Chouhan, P. (2020). Socio-demographic factors influencing utilization of maternal health care services in India. *Clinical Epidemiology and Global Health*, *8*(3), 666-670.

Sarma, S. (2009). Demand for outpatient healthcare: empirical findings from rural India. *Applied Health Economics and Health Policy*, *7*(4), 265-77.

Shariff, S.R., Moin, N.H., & Omar, M. (2012). Location allocation modeling for healthcare facility planning in Malaysia. *Computers & Industrial Engineering*, *62*(4), 1000-1010.

Stats SA (2020). Mid-year population estimates 2020. South Africa.

Talbi, E. (2009). *Metaheuristics: from design to implementation*. John Wiley & Sons, Hoboken, New Jersey.

Wabiri, N., Chersich, M., Shisana, O., Blaauw, D., Rees, H., & Dwane, N. (2016). Growing inequities in maternal health in South Africa: A comparison of serial national household surveys. *BMC Pregnancy and Childbirth*, *16*(1), 256. https://doi.org/10.1186/s12884-016-1048-z.

Waldeck, L., Holloway, J., & Van Heerden, Q. (2020). Integrated land use and transportation modelling and planning. *Journal of Transport and Land Use*, *13*(1), 227-254.

Wellay, T., Gebreslassie, M., Mesele, M., Gebretinsae, H., Ayele, B., Tewelde, A., & Zewedie, Y. (2018). Demand for health care service and associated factors among patients in the community of Tsegedie District, Northern Ethiopia. *BMC Health Services Research*, *18*(1), 697. https://doi.org/10.1186/s12913-018-3490-2.

Yantzi, N., Rosenberg, M.W., Burke, S.O., & Harrison, M.B. (2001). The impacts of distance to hospital on families with a child with a chronic condition. *Social Science & Medicine*, *52*(12), 1777-1791.

Zimmerman, S. (1997). Factors influencing Hispanic participation in prostate cancer screening. *Oncology Nursing Forum*, *24*(3), 499-504.