An Optimization Tool for Facility Location in Developing Countries
Casestudy for Timor-Leste

by Joyce Antonissen (U1275292)

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Business Analytics and Operations Research

Tilburg School of Economics and Management
Tilburg University

Supervised by: Prof. dr. ir. D. den Hertog
Co-reader: Prof. dr. ir. H. Fleuren

April 23, 2021

Acknowledgements

This thesis is the final part of my academic career at the department of Econometrics & Operations Research of Tilburg University. The way to getting my degree was not always easy. Therefore, I would like to thank my mum and dad for always supporting and motivating me when needed.

Special thanks my supervisor Dick den Hertog for coming up with the magnificent subject for my thesis and his enthusiasm for this project. It has been great to be working on a project which can be of great impact for a better world. Additionally, I would like to thank him for the thoughtful comments and recommendations on this research.

My research would have been impossible without the aid and input of the World Bank, with special thanks to Parvathy Krishnakumari. Our meetings and conversations helped me to get new insights and made writing this research much easier.

Finally, many thanks to all participants that took part in the study and enabled this research to be possible.

Management summary

The World Bank is an international financial institution that provides loans and grants to the government of poorer countries for the purpose of pursuing capital projects. As an example, World Bank provides a budget for Timor-Leste to improve infrastructure development, because Timor-Leste is in a bad situation compared to other countries. In particular, Timor-Leste gets separate budgets to build new schools, invest in new hospitals and improve or construct new roads. In this research, we focus on developing a generic model to optimize locations for new facilities, where the application is on locating new hospital facilities in Timor-Leste. The model is used to support decisions for country ministries of finance and health sector authorities, when World Bank is helping in finance.

To get to the desired results, we formulate a model and implement it into Python. The objective of this model is to maximize the number of inhabitants that can reach a facility within a maximum desired travel distance. To ensure the model is generic, the input parameters of the model can be changed. In this model it is possible to modify the maximum travel distance, the number of located hospitals and the distance matrix, which contains all distances from every household to every (possible) facility location. The variability in this model ensures that it can be used for various facilities in different countries.

In order to optimize the locations of new hospitals in Timor-Leste, we need to collect the necessary data. We extract the data of the existing hospitals, the locations of households, the roads and the set of possible locations to build new hospitals. We extract this data for the whole country and for the district Dili. In this district, the capital city is located, which ensures it is more developed compared to the rest of the country, so the data is more complete in this area. Currently, there are 139 hospitals located in Timor-Leste, 7 of them are located in Dili.

Next, the data is analysed by looking to three measures in each district: the average population that needs to be served per hospital, the average squared kilometers one hospital has to cover and the longest distance that has to be traveled to get to a hospital. We expect 3 additional hospitals are needed in Dili, which results in a total of 10 hospitals in Dili. In the whole country, we predict 37 additional hospitals are needed, which results in 176 hospitals in Timor-Leste.

After this first analysis, we implement the extracted data into the formulated model. Currently, there are 138 hospitals located in Timor Leste excluding Atauro and Sar of Oecusse. We exclude those areas because they are not connected to the rest of the country, so they can be seen as separate optimization problems. The 138 existing hospitals cover approximately 61% of the population when we allow the inhabitants to travel 5 kilometers to a hospital. If we optimize hospital locations in Timor-Leste without taking into account the

already existing hospitals in Timor-Leste, for a maximum travel distance of 5 kilometers, only 17 hospitals are needed to get at least the same coverage. With a travel distance of 10 kilometers, we only need 22 hospitals to cover the same 81% of inhabitants as the 138 existing hospitals do. In order to get 95% coverage in Timor-Leste, a total of 180 or 147 hospitals are needed, taking into account the already existing hospitals with maximum travel distances of 5 or 10 kilometers, respectively.

However, we have used Open Street Map data to get to the above mentioned results, which might not be complete. Moreover, data is always different from reality and constantly improving, especially in developing countries. Since the designed model is easily adjustable, we can import new data whenever improved data is available and update the results. Additionally, this makes it possible to use the model for various countries and different types of facilities as well, which is especially helpful for other developing countries facing the same problems.

The scalability of the model is good for now. The running time of the optimization is within a few minutes, which means that the optimization model is fast enough to be used as an interactive tool. However, when the data set is enlarged, the running time of the model will be slower and therefore, it is necessary to take some steps to speed up running time.

Instead of adding hospitals to the network, it might be profitable to include more roads to the hospitals. For Dili, we measured the impact of a different road network. The second road network contains 6.2% additional kilometers of roads compared to the original network, with most additional roads in rural areas. The improvement in coverage is only 0.01%. However, Dili contains a large urban area with a complete road network, so the obtained results might not be generalizable. Therefore, we recommend to extend this research to other areas.

In order to deepen this research, multiple avenues for future research are proposed in this report. To make it more likely that a hospital is placed in a needy area, weights can be added to those areas. Secondly, capacity restrictions are not taken into account in this research, as well as personnel restrictions. It can be interesting to include this in the model. Moreover, other kinds of facilities bring other needs. This has to be cleared out before the model is used for other facilities. However, the model developed in this research is not suitable to optimize road networks. To optimize a road network, different optimization techniques should be used. Lastly, it can be interesting to combine the optimization models of various facilities and the optimization model of the road network in order to make a good consideration between different aspects of infrastructure improvement.

Contents

1 Introduction

The World Bank is an international financial institution that provides loans and grants to the governments of poorer countries for the purpose of pursuing capital projects. These loans and grants support a wide array of investments in areas such as education, health, public administration, infrastructure, financial and private sector development, agriculture, and environmental and natural resource management. World Bank undertakes analytical and advisory activities to inform country, regional and global development agenda in line with its commitment to the twin goals of eliminating extreme poverty and boosting shared prosperity in a sustainable manner. Development projects are implemented by borrowing countries following certain rules and procedures to guarantee that the money reaches its intended target.

As an example, World Bank provides a budget for Timor-Leste to improve infrastructure development. Therefore, the optimization of hospitals in Timor-Leste is used as a case study in our research. Timor-Leste is a developing country which is in a bad situation compared to other countries. The country suffered economically from the battle to become independent, which succeeded in 2002. Timor-Leste is with a GDP per capital of 2000 dollars one of the poorest countries in the world. Even though Timor-Leste has a market economy that has a lot of revenue from offshore oil and gas reserves, only little of it has gone to develop villages, which still rely on subsistence farming.

Despite considerable government investment on health care, the health infrastructure in rural Timor-Leste is very underdeveloped. There are five district referral hospitals, 66 government owned health are centres, 42 maternity clinics and 193 health clinics. The current number of health posts is far short of the planned 442 [2]. In order to help the development of the country, World Bank releases a budget for infrastructure development, which is the construction of basic foundational services in order to stimulate economic growth and quality of life improvement. In particular, Timor-Leste gets separate budgets for building new schools, investing in new hospitals and improve or construct new roads. In this research, we optimize the locations of new facilities, where the application is on locating new hospital facilities in Timor-Leste. The goal of the World bank is to develop a generic tool in which input variables can be adjusted such that it can be used to support facility location decisions for other kinds of facilities and other countries too. This tool is used to support decisions for country ministries of finance authorities, when World Bank is helping in finance.

1.1 Research goals

Since the World Bank releases separate budgets for building schools, hospitals and roads, we can see their location optimization as three separate problems. In this research, we focus on the optimization of locating new hospitals. The main goal of this research is to develop an optimization model that can optimize the number of inhabitants that can reach a hospital within a desired maximum travel distance, given the number of new located hospitals. Moreover, the model should be such that it can be solved fast enough for different input variables. This is necessary such that the model can be used as an interactive tool.

1.2 Deliverables

This report is delivered in which this research is described. It contains an introduction to the problem, a review of basic facility models, the model used to solve the problem, a description of the data used and results of the research. Additionally, a prototype of software for an interactive tool is delivered with the aid of World Bank to optimize facility locations. The aim of the research is to deliver a generic optimization software that can be adapted so it can be used to optimize hospital locations for other countries as well. Additionally, we want the model to be easily adjusted so it can be used for other types of facilities, for example schools.

1.3 Project approach

First, some literature needs to be studied to get a better idea about what kind of problems World Bank solves and the current situation of Timor-Leste. Additionally, we need to study some literature about optimizing facility locations. Next, we can formulate a model for the optimization of the hospital locations. To do this accurately, we need to make some assumptions too.

The next step is to collect the necessary data. To get to the desired results we need the existing road network, as well as the already existing locations of hospitals. Additionally, we need the distribution of population of Timor-Leste. First, we use descriptive statistics to analyze this data. Next, taking the extracted data as input variables, we can write a program to determine the optimal locations for the new facilities.

In order to measure the effect of the chosen road network on our model, we extract a second road network. Next, the optimization of hospital locations in Timor-Leste is done for both road networks, while the other input variables are unchanged. Afterwards, the results of those optimizations are compared in order to evaluate the differences in optimal solutions by changing the road network.

One of the goals of this research is to deliver a tool which can optimize the hospital network. Once this tool is available for optimizing the hospital network in Timor-Leste, we want to be able to adapt the tool to optimize different facilities for various countries to help other countries defining their optimal facility locations. Since there are other developing countries facing the same struggles in locating facilities, it may be of great interest to use our model to optimize facility locations in other countries as well.

In the next section, Section 2, a review of basic facility models is given. Section 3 contains the model used to optimize the hospital locations. Additionally, the assumptions and the implementation of the model are elaborated in that section. Section 4 contains an explanation of the data collection and a first descriptive analysis of the data. The results of this research are shown in Section 5, which are followed by the conclusion, limitations, recommendations and future research in Section 6.

2 Basic facility models

In this section, we review basic facility models. The general objective of all location problems is to optimize the location of new facilities given the locations of demand. It is given where demand is located, so which points in the network have to be served by the facilities. The already existing facility network is given, as well as the underlying network.

In general, location problems may be either continuous or discrete. A discrete facility location problem is characterized by the assumption that all demand points can be aggregated into a finite number of points. Similarly, the set of possible facility locations needs to be finite. When a location problem is continuous, the facilities may be located anywhere in the feasible region. In this research, we focus on discrete problems since we assume there is a finite set of demands and a finite set of possible hospital locations.

Daskin [6] classified discrete location problems into three broad categories: covering-based problems, median-based problems, and other problems. These classes of location problems are reviewed and we discuss whether they can be of any interest in this research. If a class of location problems is interesting, we further elaborate that class and review the formulation of the foundational models in that class.

In this section, discrete basic facility location models are presented with some notes about the application of the models to our research problem. In the first subsection, coverage-based problems are discussed. In the second, third and fourth subsection, the set covering location problem, the maximal coverage location problem and the p-center location problem are further elaborated, respectively. The fifth subsection contains a review of median-based problems and the sixth subsection gives a literature review on applications to developing countries.

2.1 Coverage-based problems

The aim of covering-based problems is to optimize the coverage of demand when locating facilities. Demand is said to be 'covered' when it is located within a specified coverage distance from facilities which serve them. The class of coverage-based problems includes three basic types: the set covering location problem, the maximal covering location problem and the p-center location problem. In our research, we would like to have as many people as possible to be able to reach the facilities within a predetermined distance. Since the methodology of coverage-based problems is in line with the problem we have for Timor-Leste, we elaborate this class of location problems.

2.2 Set covering location problem (SCLP)

The first mathematical model in covering problems was developed by Toregas, Swain, ReVelle and Bergman [14], which was a model for the set covering location problem (SCLP). For more recent papers on this problem with applications to health care, we refer to the papers written by Ablanedo-Rosas et al. [7] and Dekle et al. [1].

The objective of the set covering location model is to minimize the number of located facilities or to minimize the total location cost, such that all of the demand nodes are covered within the predetermined travel distance. This problem finds the number of facilities that needs to be located and their location such that all demand points are within a specified travel distance of the facilities that serve them. We write the set covering problem as the following discrete integer model:

$$\text{Minimize} \quad \sum_{j \in J} c_j x_j \tag{1}$$

$$\text{Subject to} \quad \sum_{j \in N_i} x_j \geq 1 \quad \forall i \in I \tag{2}$$

$$x_j \in \{0,1\} \quad \forall j \in J, \tag{3}$$

where: I = the index set of all demand nodes, indexed by $i = 1, ..., n$
J = the index set of all candidate facility sites, indexed by $j = 1, ..., m$
$N_i = \{j \in J | d_{ij} \leq S\}$, the set of all candidate locations that can cover demand point i
d_{ij} = the travel distance from demand point i to facility j
c_j = the fixed cost of locating a facility at candidate location $j \in J$
S = the maximum acceptable travel distance from demand i to facility j (the cover distance)

and additionally, we have the following decision variable

$$x_j = \begin{cases} 1 & \text{if a facility is established at candidate location } j \\ 0 & \text{otherwise.} \end{cases}$$

The objective (1) minimizes the location costs of the facilities which are needed to cover all demand points. Sometimes, the costs (c_j) are set equal to one. This results in minimizing the number of facilities needed to cover all demand points. The first constraint (2) ensures that each demand node is covered by at least one facility. Constraint (3) is a standard integrality condition, which forces a facility either to be opened or closed.

There are a few problems that arise when using this set covering location problem. First, covering all demands mostly results in an enormous amount of costs if (1) is used as the objective function. If the costs (c_j) are set equal to one as discussed above, the number of facilities that have to be opened to cover all demands is extremely large in most cases. Second, the model fails to distinguish the size of the demand nodes. It does not take into account whether a node produces a large or small amount of demand, which can cause problems.

In our research, we would like to maximize the number of people that are able to reach a facility within a desired preset distance. Therefore, the set covering problem is not suitable because it does not take into account the size of the demand at nodes and it is not possible to set a budget for building the necessary facilities.

2.3 Maximal coverage location problem (MCLP)

The two problems of the set covering location model motivated Church and ReVelle [4] to formulate the maximal coverage location problem (MCLP). For more recent papers on the MLCP with applications to health care, we refer to the papers by Jia et al. [9] and Murali et al. [10].

The MCLP addresses planning situations which have an upper limit on the number of facilities to be sited. The problem seeks the maximum population which can be served within a stated service distance or time given a limited number of p facilities. The maximal covering location problem can be formulated as follows:

$$\text{Maximize} \quad \sum_{i \in I} v_i z_i \tag{4}$$

$$\text{Subject to} \quad \sum_{j \in J} x_j = p \tag{5}$$

$$z_i \leq \sum_{j \in N_i} x_j \quad \forall i \in I \tag{6}$$

$$x_j \in \{0, 1\} \quad \forall j \in J \tag{7}$$

$$z_i \in \{0, 1\} \quad \forall i \in I, \tag{8}$$

where: I = the index set of all demand nodes, indexed by $i = 1, ..., n$
J = the index set of all candidate facility sites, indexed by $j = 1, ..., m$
$N_i = \{j \in J | d_{ij} \leq S\}$, the set of all candidate locations that can cover demand point i
d_{ij} = the travel distance from demand point i to facility j
v_i = the volume of demand at node i
S = the maximum acceptable travel distance from demand i to facility j (the cover distance)
p = the number of candidate locations to be established

and additionally, we have the following decision variable

$$x_j = \begin{cases} 1 & \text{if a facility is established at candidate location } j \\ 0 & \text{otherwise} \end{cases}$$

$$z_j = \begin{cases} 1 & \text{if demand at node } i \text{ is covered} \\ 0 & \text{otherwise.} \end{cases}$$

The objective function (4) maximizes the population that is covered. Demand is covered if there is an open facility that serves the demand within the predetermined desired distance. The first constraint of this model (5) ensures that exactly p facilities are to be located. Constraint (6) states that demand points are only covered by open facilities. Constraint (7) and (8) are standard integrality constraints.

In our research, we would like to have as many people as possible to be able to reach the facilities within a predetermined distance. This means we can use the objective that is used in maximal coverage location problems. We do not set a limit on the number of facilities as in constraint (5), but we have to keep within a budget. We can easily change constraint (5) to a budget constraint in order to satisfy this requirement. Additionally, in order to use this model, it is easier to adapt it to a linear integer model if possible.

2.4 p-center location problems (PCLP)

The third type of covering-based problems are the p-center location problems (PCLP). The first mixed integer programming (MIP) formulation for the discrete p-center problem was proposed by Daskin [5].

The aim of these problems is to minimize the average travel distance among all demand points and their facilities, assuming each demand point is covered. When not taking into account capacities, demands are allocated to their closest open facility. The p-center problems minimizes the maximum distance between demand and its nearest facility. Therefore, this problem is also known as the minimax problem. The basic PCLP is formulated as follows:

$$
\begin{aligned}
\text{Maximize} \quad & L && (9)\\
\text{Subject to} \quad & \sum_{j \in N_i} y_{ij} = 1 \quad \forall i \in I && (10)\\
& \sum_{j \in J} x_j = p && (11)\\
& \sum_{j \in N_i} d_{ij} y_{ij} \leq L \quad \forall i \in I && (12)\\
& y_{ij} \leq x_j \quad \forall i \in I, \forall j \in N_i && (13)\\
& x_j \in \{0,1\} \quad \forall j \in J && (14)\\
& y_{ij} \in \{0,1\} \quad \forall i \in I, \forall j \in N_j && (15)\\
& L \geq 0, && (16)
\end{aligned}
$$

where: I = the index set of all demand nodes, indexed by $i = 1, ..., n$
J = the index set of all candidate facility sites, indexed by $j = 1, ..., m$
$N_i = \{j \in J | d_{ij} \leq S\}$, the set of all candidate locations that can cover demand point i
d_{ij} = the travel distance from demand point i to facility j
S = the maximum acceptable travel distance from demand i to facility j (the cover distance)
p = the number of candidate locations to be established

and additionally, we have the following decision variables

$$
x_j = \begin{cases} 1 & \text{if a facility is established at candidate location } j \\ 0 & \text{otherwise} \end{cases}
$$

$$
y_{ij} = \begin{cases} 1 & \text{if demand point } i \text{ is assigned to the facility at candidate location } j \\ 0 & \text{otherwise.} \end{cases}
$$

In this model, L is an auxiliary variable. This variable is used to compute the maximum distance between a demand point and the nearest facility allocated to it.

The objective function (9) minimizes the maximum demand-weighted distance through the auxilary variable L. The first constraint of this model (10) makes sure each demand point is covered by exactly one facility. Constraint (11) ensures that there are p facilities established. Constraint (12) is used to compute L, which is used to compute the maximum distance. Constraint (13) is needed such that demand is only subscribed to open facilities. Constraint (14) and (15) are integrality constraints and constraint (16) is there to be sure that the maximum distance is greater or equal to zero.

This sum minimizes the maximum distance between a point and the nearest facility located to it. Although it is good to minimize the longest travel one has to make, this means that the objective does not take into account how many people can or cannot reach the facilities within a predetermined maximum travel distance. In our problem, we want that as many people as possible are able to reach the facilities within the desired distance. Therefore, this model is not suitable for our research.

2.5 Median-based problems

Median-based problems, introduced by Hakimi [8], minimize the weighted average distance costs between the demand nodes and the facilities to which they are assigned to determine the optimal location for the facilities. These optimal locations are the medians of the network. In this class, the p-median problem and the fixed charge location problems the most common problems. The p-median location problems minimize total demand-weighted travel distance. The fixed charge location problems does the same, but also takes into account the traveling costs next to the costs of opening the facilities. Median-based problems do not look to individuals but pool distances. Therefore, these models are not accurate to use for our research since we want as many people as possible being able to reach the facilities within a desired distance.

2.6 Literature review on applications to developing countries

The paper by Rahman and Smith [11] reviews the use of location-allocation models in health care in developing countries. This paper tries to examine the suitability of the already existing methods, as well as showing their relevance to overall development. There are a lot of papers written about location-allocation applied to developing countries. We mention a few of them in this subsection.

The paper written by Radiah Shariff et al. [13] uses a maximal coverage location problem (MCLP) to study the healthcare in one of the districts in Malaysia. A new solution approach is proposed based on genetic algorithm to examine the percentage of coverage within the allowable distance specified by the Malaysian government. The solution was used to analyze the past location decisions and to identify the need for additional facilities or the need for additional increase in capacity. Since this problem is corresponding to our research problem, the algorithm developed in this paper can be useful for our research too.

The paper written by Rahman and Smith [12] describes the problem of finding suitable sites for additional health facilities in a rural area in Bangladesh. The goal is to improve the accessibility of people to the health care system given the existing set of facilities. The problem of Bangladesh is considered as a

MCLP. It addresses a maximum travel distance between facilities and villages in Bangladesh. Additionally, it descrips the importance of location-allocation modelling in developing countries. This paper is helpful to convince people that optimization techniques can be usefull in locating facilities. The example stated in this paper clearly shows better results when using optimization techniques.

The paper by Chu and Chu [3] focuses on matching the supply and demand of public hospital beds in Hong Kong for the future years. To get to the desired result, the entire territory of Hong Kong is partitioned into clusters before the Hospital Location with Service Allocation Planning (HLSAP) model is used. This paper addresses the planning issues of hospital locations and service allocations, which include new services distribution as well as existing services redistribution. This paper uses a totally different model, which makes the paper less relevant for our research.

3 Mathematical model

The locations of the facilities providing public health are very crucial in ensuring that the chosen location network serves the purpose of maximizing the benefits of people. Therefore, we use a mathematical model to optimize the locations of hospitals in Timor-Leste. The main goal in this model is that as many people as possible should be able to reach a hospital facility within a desired maximum travel distance.

Using the theory we have seen in Section 2, we formulate a mathematical model to solve the problem of this research. In the first subsection, this mathematical model proposed. In the second subsection some general assumptions for the model are given. In the third subsection is explained how this model is implemented and solved for Timor-Leste. How to speed up the running time of the model is described in the fourth section.

3.1 Model formulation

The following formulation is used to model the problem:

$$\text{Maximize} \quad \sum_{i \in I} \sum_{j \in J} v_i y_{ij} \tag{17}$$

$$\text{Subject to} \quad \sum_{j \in J} c_j x_j \leq B \tag{18}$$

$$\sum_{i \in I} y_{ij} \leq n x_j \quad \forall j \in J \tag{19}$$

$$\sum_{j \in J} y_{ij} \leq 1 \quad \forall i. \tag{20}$$

$$x_j, y_{ij} \in \{0,1\} \quad \forall i \in I, \forall j \in J, \tag{21}$$

where:
- I = the index set of all demand nodes, which are corresponding to the villages in Timor-Leste, indexed by $i = 1, ..., n$
- J = the index set of all facility sites, which are corresponding to the possible locations for a hospital at the island, indexed by $j = 1, ..., m$
- v_i = the volume of the demand at node i
- c_j = the cost of opening facility j
- d_{ij} = the travel distance from demand i to facility j
- S = the maximum travel distance from demand i to facility j
- B = the budget released by World Bank

and additionally, we have the following decision variables:

$$x_j = \begin{cases} 1 & \text{if the hospital is sited at } j \text{ needs to be build} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ij} = \begin{cases} 1 & \text{if the demand at node } i \text{ is served by facility } j \text{ and } d_{ij} \leq S \\ 0 & \text{otherwise.} \end{cases}$$

The objective function (17) maximizes the population assigned to a facility within the desired distance S. When the facility is not within the desired range of S kilometres, y_{ij} is equal to zero. This means that the demand at node i is not served by facility j in that case. The first constraint (18) makes sure that building the new hospitals stays within the budget. At this point, we do not have a budget for building the hospitals in Timor-Leste. Therefore, we replace constraint (18) with the following constraint:

$$\sum_{j \in J} x_j \leq p \tag{22}$$

where: p = total number of hospitals located.

This constraint limits the total number of hospital facilities located. This maximum number of hospitals include the already existing hospitals. If the budget and costs are available at a later point of time, this constraint can be replaced by the budget constraint (18) again.

The second constraint (19) of the model ensures that no one is assigned to a facility if the facility is not opened. If facility j is not open, x_j is equal to zero, so this means that for each i, y_{ij} has to be zero too. Constraint (20) restricts the number of hospitals one household can be served from. This constraint is added to be sure that each household is only served by one hospital. If we omit this constraint, every hospital within the maximum travel distance of a household serves that household, which increases the objective value. Therefore, this constraint is important to get the right results. The last constraint (20) makes sure the decision variables are integer.

3.2 Assumptions

We make some general assumptions for the model formulated in the previous subsection. First, we assume the possible locations for building a hospital are known. The set of possible hospital locations should be a finite set. Second, we assume for now that capacity is not taken into account. Every hospital can serve as much people as needed. This is because we have no data about the capacity per existing hospital location. Since we do not take into account capacity in the model, we assume a person is randomly assigned to a hospital if one can reach multiple hospitals within the maximum desired travel distance.

3.3 Solving the model

In order to solve the mathematical model, we use the Gurobi Optimizer in Python. The Gurobi Optimizer is a commercial optimization solver which can be used for mixed integer linear programming (MILP). In Section A.2 of the Appendix, the Python file can be found. As mentioned before, the input parameters of this problem are:

- the desired maximum travel distance,
- the number of extra hospitals which needs to be located,
- the distance matrix.

After the importation of those input parameters, we add the decision variables and the objective to the model in Python. We add a binary decision variable x_j for every (possible) hospital location j. Additionally, we add a binary decision variable y_{ij} for every i and every j. We need that the variable y_{ij} can be equal to one if $d_{ij} \leq S$, so the distance from node i to facility j needs to be less or equal to the maximum desired travel distance. We add this requirement to the constraints. Next, all constraints of the model in Section 3.1 are added to the model in Python.

The model returns an optimal objective value after running the program. This optimal objective value is equal to the number of households that are served by a hospital within the desired distance. Lastly, we export the matrices containing x_j and y_{ij}. These matrices show which facilities are opened and which households are served by those facilities. By re-running this model for different numbers of added facilities, we can merge the outcomes to a Pareto curve for the coverage against the number of hospitals opened. This makes it easier to analyse the results.

3.4 Speed up running time

Because the data set used for our model can be enormous, it is important that the running time of the model is reasonably good. In this subsection, we discuss ways to speed up the running time of the model.

Only add the decision variable if the travel distance is less than the maximum travel distance

If we add the variable y_{ij} to the optimization for every i and every j, there are $I * J$ decision variables in the optimization problem. For the whole country, the number of variables is equal to $37,379 * 1,138 = 42,537,302$, which is a very high number of variables in the model. Next, a constraint is added for each y_{ij} which states that the variable can only be equal to one if $d_{ij} \leq S$. This number of constraints is too much for the Python program, which causes the execution to fail using a large data set. Therefore, the variable y_{ij} is only added if the distance from node i to (possible) hospital location j is less than or equal to the maximum travel distance. This results in a lower number of variables and constraints which ensures the running time decreases. The outcome of the model remains unchanged. In Figure 1 below, the different running times for the hospital location optimization of Dili with a maximum travel distance of 5 kilometers including the already existing hospitals are presented. The blue line represents the running time of the model for which the variables are added if they meet the before mentioned requirement. The orange line represents the running time of the model where all variables are added. Note that the optimization problem for Timor-Leste is much bigger than the problem for Dili, so the running time will increase significantly.

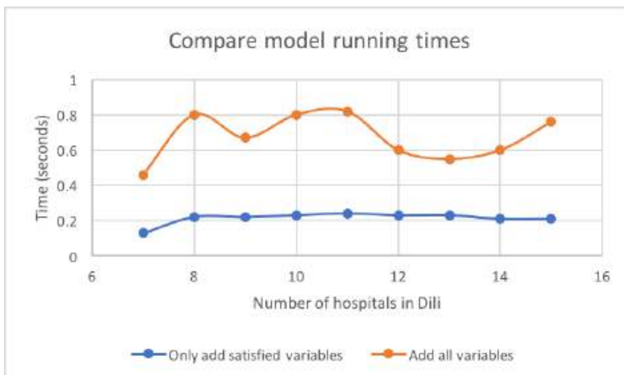


Figure 1: Running time of the model by adding all variables or only adding variables that satisfy the constraint

Make a continuous decision variable of y_{ij}

In order to speed up the running time, we set y_{ij} to a continuous decision variable instead of a binary decision variable. This means that one household can be connected to multiple hospitals, and distribute its visits among those hospitals. This does not change the objective value or which facilities are opened, but changing the type of the variable y_{ij} changes the running time of the model significantly. In Figure 2, the running time for the different types of variables are represented for the hospital location optimization of Dili with a maximum travel distance of 5 kilometers and taking into account the already existing hospitals. Note again that the data set of Dili is small compared to the data set of the whole country, so where the running time seems small in this example, it will enlarge quickly when the data set gets bigger.

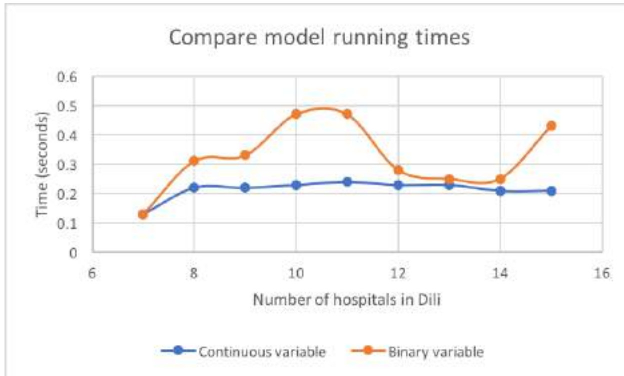


Figure 2: Running time of the model with continuous or binary variables

Cluster the households

Another possibility to speed up the running time is to cluster the households. If a few households are located close to each other, they can be clustered together. Because of this, a weight is added to the objective for the number of houses in the corresponding cluster. When we take the variables y_{ij} as binary and add a variable y_{ij} for all i and all j, clustering the houses has a positive effect on the running time. Because the running time improved already very much by changing the variable y_{ij} to a continuous variable and only adding the variables y_{ij} when the distance is less than the maximum travel distance, the improvement in time of adding clusters is nil or the running time is even slower.

4 Data collection and analysis

In order to optimize the locations of new hospitals in Timor-Leste, we need to collect the necessary data. To get to the desired results, we extract the input data from QGIS. We have used QGIS 3.10.5, which can be downloaded from the QGIS website[1]. Additionally, some plug-ins are downloaded and used, which are mentioned in the explanation of the data extraction. In this section, it is explained how the data is extracted for existing hospitals, households, possible locations to build a hospital and roads. Additionally, an extra road network is extracted such that we can measure the effect of using a different road network on our model.

The data is extracted for the whole country and for the district Dili. We consider the district Dili because this is the district where the capital city is located. It is a little more developed which makes it easier to collect data, so the data is more complete compared to the rest of the country. In the optimization of Timor-Leste, we will neglect the island Atauro and Sar of Oecusse. They are not connected to the rest of the country, so they can be seen as separate optimization problems. Therefore, the numbers for Timor-Leste excluding those areas are shown as well. Next, the data is analysed by looking to three measures:

- the population per hospital in a district,
- the squared kilometers one hospital has to cover in a district,
- the longest distance that has to be traveled to get to a hospital in a district.

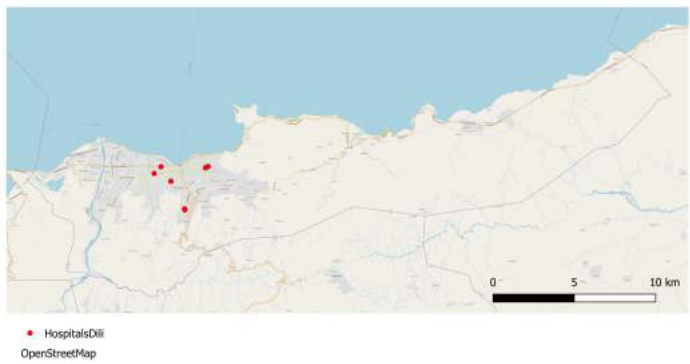
Note that all data in QGIS is open source. Therefore, the data we extract might not be complete. We take the extracted data as representative for the real data.

4.1 Existing hospitals

First, the data about the already existing hospitals is extracted from QGIS. Note that a hospital in Timor-Leste is not comparable to a hospital in the Netherlands. Since Timor-Leste is less developed, hospitals in that country are smaller and less developed as well. To extract the hospitals, we use the *QuickOSM* plug-in. This plug-in builds a query to extract the selected open street map data. We use *key = amenity* with *value = hospital* to get all hospitals. When using this plug-in we can choose whether we would like to extract all hospitals in Timor-Leste or just the hospitals in Dili. This plug-in results in a so called layer of multi-polygon features and layer of point features. To transform the multi-polygons to points, the *polygon centroids* plug-in is used. Next, the *merge vector layers* plug-in is used to merge the centroids of the multi-polygons and the points to get one set of points containing the locations of the hospitals.

[1] https://www.qgis.org/nl/site/forusers/download.html

This results in a points layer in QGIS containing all hospitals. In total, there are 139 hospitals located in Timor-Leste. Seven of them are located in the district Dili. There is only one hospital in Sar of Oecusse, and there are none on the island Atauro. Therefore, in Timor-Leste excluding Atauro and Sar of Oecusse, there are 138 existing hospitals. In Figure 3 below, the locations of the hospitals are shown on the map.



(a) In Dili

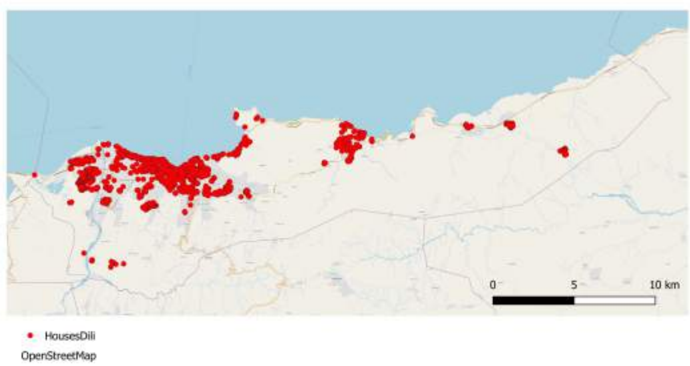
(b) In Timor-Leste

Figure 3: Existing hospitals

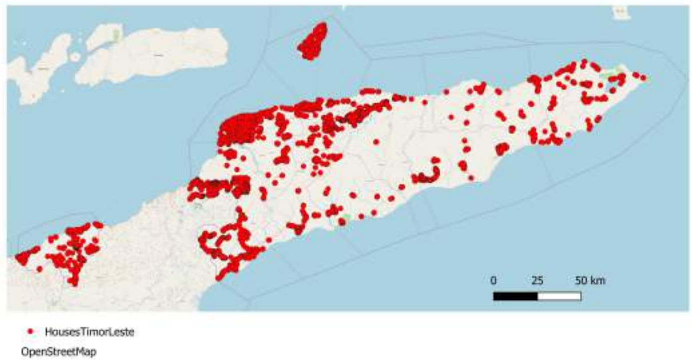
4.2 Households

Next, the locations of the households in Timor-Leste are extracted from QGIS. To do this, the same plug-in is used: *QuickOSM*. Now, we use *key* = *building* with *value* = *house*. Again, we convert the multi-polygons to points by using *polygon centroids* and then merge these layers to one set of houses with the *merge vector layers* plug-in. We can decide to extract the houses in the whole country or we can extract all houses located in Dili.

This results in a set of 915 houses in Timor-Leste. This is very little, compared to the population of 1.2 million. The reason seems to be that a lot of buildings have not been identified to any classification in the data set of QGIS. Therefore, we extract the houses in a different way. We extract all buildings in Timor-Leste by using *QuickOSM* plug-in and with the *query builder* we exclude the buildings that are classified in 'hotel', 'commercial', 'industrial' or 'retail' to be left with residential buildings. Again, we convert multi-polygons to points and merge the layers. This results in a total of 45,608 houses, which can be taken as representative for the whole data set. For Dili, we get a total of 3,861 houses. In Timor-Leste excluding Atauro and Sar of Oecusse, there are 37,379 households. The locations of the households are plotted in Figure 4 below.



(a) In Dili



(b) In Timor-Leste

Figure 4: Residential buildings

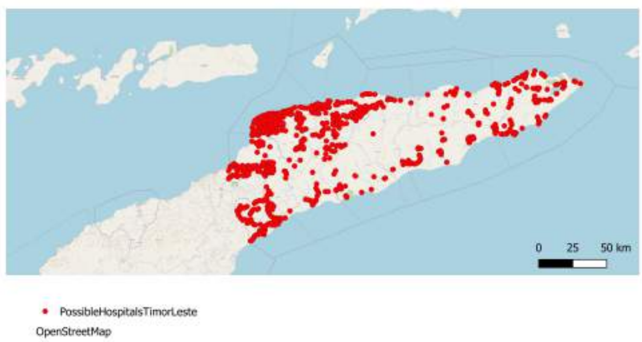
4.3 Possible hospital locations

To get possible locations for hospitals, we clustered the locations of the houses. To do this, we use the *K-means clustering* plug-in. This plug-in aims to partition the features into k clusters in which each feature belongs to the cluster with the nearest mean. The mean point is represented by the barycenter of the clustered features. In this plug-in, we need to define the desired number of clusters. The plug-in returns an additional row in the attributes table containing the number of the cluster it is placed in. Next, the *Means Coordinate(s)* plug-in is used to get the center location for each cluster.

To get the possible hospital locations for Dili, we set the desired number of clusters to 50. This means we have 50 possible hospital locations for Dili, additional to the already existing hospital locations. If we decide to open an additional hospital on all 50 extra locations, everyone is Dili can reach a hospital within 5 kilometers. Therefore, we decide 50 possible hospital locations is sufficient for Dili.

For the whole country, much more possible hospital locations are needed. We take 1000 possible hospital locations for Timor-Leste, not taking into account Atauro and Sar of Oecusse. With a maximum travel distance of 5 kilometers, 1,098,624 of the 1,105,456 inhabitants of Timor-Leste excluding Atauro and Sar of Oecusse can be served. This equals a coverage of 99,38 %. With a maximum travel distance of 10 kilometers, 1,099,393 inhabitants can reach a hospital when all 1,138 hospitals are located, which equals 99,45%. The coverage when using all hospitals is so close to 100% coverage that we assume 1,000 possible hospital locations is enough. The possible hospital locations for Dili and Timor-Leste are plotted in Figure 5 below.

(a) In Dili



(b) In Timor-Leste

Figure 5: Possible locations for new hospitals

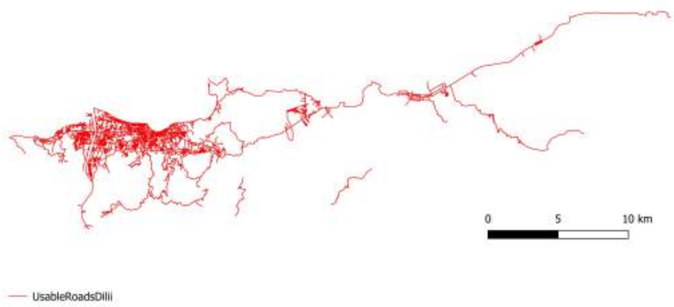
4.4 Roads

Lastly, we need to extract the roads in Timor-Leste. Again we use the *Quick-OSM* plug-in to extract the roads. We use *value=highway* to get all roads. This value has a large set of possible keys. These keys are listed in Table 1 below with a short description. To make sure that the roads in our data set can be used to travel to the hospital, the roads are analyzed using the Google Satellite. For each class, a few roads are viewed carefully and it is estimated whether the class of roads can be used to travel to the hospitals. This estimate is generalized for the whole class of roads. In Table 1 below, the last column gives a check mark if the class of roads is usable to travel to the hospital. There is a cross in that column if the roads are not usable for this purpose.

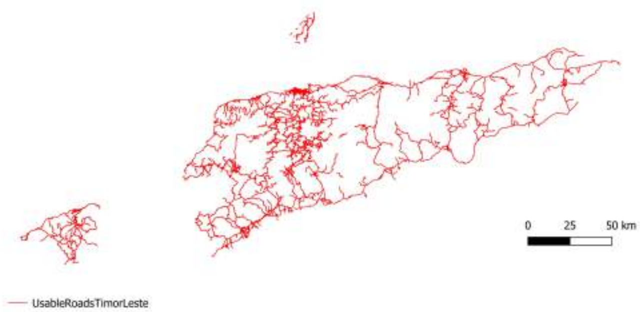
Key	**Short description**	**Usable road**
motorway and motorway_link	Major divided highway and the links to these roads	✓ ✓
trunk and trunk_link	Most important roads in a country that are not motorways and the links to these roads	✓ ✓
primary and primary_link	Often link larger towns and the links to these roads	✓ ✓
secondary and secondary_link	Often link towns and the links to these roads	✓ ✓
tertiary and tertiary_link	Often link smaller towns and villages and the links to these roads	✓ ✓
unclassified	Minor roads, serve other purpose than access to properties	✓
residential	Access to housing	✓
living_street	Residental streets	✓
service	For access roads to car park, business park etc.	✓
pedestrian	Roads used by pedestrians	✗
track	For agricultural or forestry uses	✗
bus_guideway	Not suitable for other trafic than busses	✗
escape	Enable vehicles with braking failure to safely stop	✗
raceway	For (motor) racing	✗
road	An unknown type; from footpath to motorway	✗
paths	For horseriders or walkers	✗
cycleway	To use by bicycle	✗

Table 1: Classes of roads in Timor-Leste

The roads which are marked as usable are extracted one by one using the *QuickOSM* plug-in. This results in multiple layers of roads, one layer for each class of roads. Next, these roads are merged by using the *merge vector layer* plug-in. Not every house, hospital or possible hospital location is connected to a road. Since they can be far away from the main road, it might be relevant to consider the distance a household has to travel to the road as well. To connect all buildings to the roads, we first convert the lines of the roads to points using the *Convert lines to points* plug-in. This is necessary because the plug-in we want to use to connect the buildings to the roads only connects points to points instead of points to lines. The *Distance to nearest hub (line to hub)* plug-in is used to connect every house, hospital or possible hospital location to the nearest point of this road layer consisting of points. The plug-in gives a layer of all lines connecting those points. Finally, the already extracted layer of roads is merged with these new line layers to get one layer of all usable roads. The road network is shown in Figure 6.



(a) In Dili



(b) In Timor-Leste

Figure 6: Usable roads

4.5 Second road network

In order to compare the model for two different road networks in Dili, we extract the total network containing all roads for Dili. This results in a road network for Dili including the footpaths as well. We do this using the *Quick-OSM* plug-in with *value=highway* and *key = **. Again, we use the *Convert lines to points* plug-in to convert all roads to points. Next, we use the *Distance to nearest hub (line to hub)* plug-in to connect every house, hospital or possible hospital location to points of the road network. Finally, we merge those new lines to the road network layer to get one layer of usable roads.

This second road network is shown in Figure 7. In this figure, the red road network is the same network as shown in Figure 6, so the network containig all travelable roads, which are marked good enough to travel to the hospital. The green lines show the roads that will be added to the road network if we extract the enitre road network in Dili.

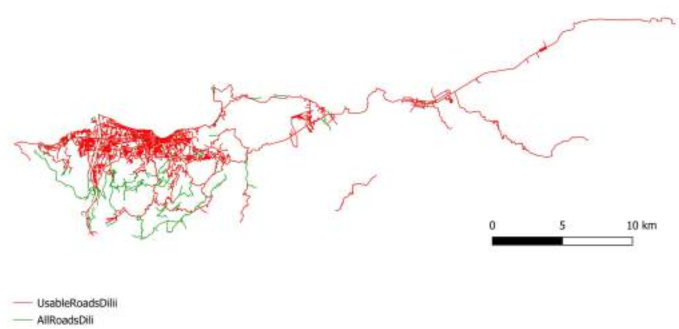


Figure 7: Second road network in Dili

As we can see in the map above, the roads that are added to the network are mostly rural roads. In the populated city centre, most of the roads are already included in the previous extracted road network of Dili in Figure 6. The new road network contains 6.2% additional kilometers of roads compared to the original road network.

4.6 Calculating the distances

In order to optimize the locations of hospitals, it is necessary to have the distance matrix. In this matrix, the distance from every household to every (possible) hospital location is given. To get to the desired distance matrix, we first use the *OD Matrix from Layers as Table (m:n)* plug-in. This plug-in searches for distances along a network between the selected origin and destination layers. We take one of the extracted road networks to be the network layer. The origin layer is the layer of the houses and as destination layer we have the layer of (possible) hospital locations. The plug-in results in a matrix with a column containing the start node, a column containing the end node and a column containing the total costs of traveling from the start to the end node, where the costs represent the distance in kilometers. Assume the number of hospitals and the number of possible hospital locations together is equal to m and the number of houses is equal to n. The size of the returned matrix is $(n * m)$ x 3. We can export this table from QGIS to an CSV file by using the *Export layer to CSV* plug-in. Next, we import this CSV file in Python and we use Python to transform the matrix into a n x m matrix, with the columns containing the (possible) hospital locations and the rows containing the houses. Below, the structure of this matrix is given, where d_{ij} represents the distance from household i to facility location j.

(Possible) hospital locations

Houses		1	2	3	...	m
	1	d_{11}	d_{12}	d_{13}	...	d_{1m}
	2	d_{21}	d_{22}	d_{23}	...	d_{2m}
	3	d_{31}	d_{32}	d_{33}	...	d_{3m}
	...			...		
	n	d_{n1}	d_{n2}	d_{n3}	...	d_{nm}

4.7 Current situation in Timor-Leste

To get an idea about the current healthcare situation in Timor-Leste, we consider the data about currently located hospitals in combination with the population of the country. For this we use the data of the population which was released by the government in 2015[2]. This is the most recent data about the population released by the government and it is very reliable. Additionally, we use the extracted data from QGIS about hospital locations, as elaborated above. Note that the data from QGIS is representative for the real data, so it might be different in real life. In this research, we treat the data like it is true and deliver a concept in which other data sets can be inserted.

We consider the data by district. In total, Timor-Leste has 13 districts. We divide the district of Dili into the part of the district on land and the island Atauro. They are separately located and not connected to each other, so we consider them as two separate 'districts'. Therefore, whenever the district Dili is mentioned, we talk about the part on land. When we discuss the island we just mention Atauro.

In Table 2, the population per district is represented, as well as the number of hospitals in that district and the average population each hospital has to serve in each district. The average population each hospital has to serve can be calculated by dividing the population by the number of hospitals in each district. In the fifth column, the surface of the country is given in squared kilometers and in the sixth column the average surface one hospital has to cover is stated. The average surface one hospital has to cover is calculated by dividing the surface of a district by the number of hospitals in that district. In the last column the average longest distance to a hospital in kilometers is given. To get this number, we estimated which houses have to travel far to a hospital. We measured the distance that these households have to travel to get the average longest

[2]http://www.statistics.gov.tl/wp-content/uploads/2016/11/Wall-Chart-Poster-Landscape-Final-English-rev.pdf

distance from a household to a hospital. We did this separately for each district.

District	Population	Number of hospitals	Average population per hospital	Surface in km^2	Average surface per hospital	Average longest distance (km)
Alieu	48,837	12	4,070	735.9	61	17.4
Ainaro	63,136	10	6,314	802.6	80	10.3
Atauro	9,274	0	-	224.0	-	-
Baucau	123,203	14	8,800	1502.2	107	14.2
Bobonaro	97,762	18	5,431	1378.1	77	12.4
Cova Lima	65,301	5	13,060	1198.6	240	18.5
Dili	268,005	9	29,778	140.1	16	29.7
Ermera	125,702	16	7,856	756.5	47	8.3
Lautém	65,240	6	10,873	1816.7	303	30,5
Liquiça	71,927	10	7,193	559.9	56	14.6
Manatuto	46,619	13	3,586	1783.3	137	29.2
Manufahi	53,691	12	4,474	1332.5	111	6.7
Oecusse	68,913	1	68,913	813.6	814	45.9
Viqueque	76,033	13	5,849	1872.7	144	11.9
Whole country	1,183,643	139	8515	15006	108	-

Table 2: Hospital statistics in Timor-Leste

The first thing we notice when analyzing the data is that no hospital in our data set that is located on the island Atauro. Since Atauro is an island, a boat or a plane needs to be used to get to a hospital. So we can conclude that at least one hospital should be located at Atauro to make sure the population of Timor-Leste gets better access.

If we consider to the population that has to be served per hospital, we see a very high number in Sar of Oecusse. Because there is only one hospital in this district, everyone needs to go to the same one. A reason for the small number of hospitals could be that there are a lot of hospitals located just over the border in another district, so the population could use those. Since the district Sar of Oecusse is located in between Indonesia this is not the case, so we can conclude there is a lack of hospitals in that district. In Dili, Cova Lima and Lautém, the number of people that needs to be served by one hospital is also quite high. Note that nothing is mentioned about the capacity of the hospitals, so it might be that the hospitals are designed to serve a high number of people.

If we want people to be able to reach a hospital within 5 kilometers, each hospital covers a surface of approximately 80 squared kilometers. Assuming the hospitals are perfectly distributed through the country, we see that a lot of districts have too little hospitals to meet this requirement. The districs Atauro, Baucau, Cova Lima, Lautém, Manatuto, Manufahi, Sar of Oecusse and Viqueque do not have enough hospitals to cover the whole population of their district when people are able to travel 5 kilometers to a hospital. If we extend this measure to 10 kilometers, one hospital needs to cover approximately 314 squared kilometers. With this measure, we see that the districts Atauro and Sar of Oecusse do not have enough hospitals.

In the last column, the longest travel distance to a hospital in kilometers is given. We see that in Dili, Lautém, Manatuto and Sar of Oecusse, the longest distance that has to be traveled to a hospital is more than 20 kilometers. This is quite far. Therefore, we can conclude that at least one additional hospital should be build in those districts. In every district there is at least one inhabitant who is not able to reach a hospital within 5 kilometers. If this measure is extended to 10 kilometers, only in the districts Ermera and Manufahi all people are able to reach a hospital within the desired distance. In the other districts some people of the population have to travel more than 10 kilometers. In Sar of Oecusse, the longest distance one has to travel is even more than 45 kilometers.

In conclusion, Atauro performs bad because there is no single hospital located on the island in our data set. Sar of Oecusse and Lautém perform bad on all three measures: in these districts the hospitals have to serve many people, each hospital has to cover a large area and the longest distance one has to travel to a hospital is too high. Note that bad performance can have few reasons. It can be that there is a hospital just over the border in another district. Differently, as mentioned before, it can be that the real data is slightly different from the data we extracted from QGIS. Since it is still a developing country, it can also be that there is some data missing.

4.8 Descriptive statistics

In this subsection, we describe whether an extra hospital is needed in a district and if so, how many hospitals are needed. We summarize this to a description about the number of additional hospitals needed in the whole country. We describe the need of additional hospitals using the statistics of the health care in Timor-Leste, shown in Table 2. This means we predict where hospitals are needed without using difficult mathematical models.

For the description, a figure of the current situation in Timor-Leste can be helpful too. Therefore, we combine the road network, the existing hospital locations and the locations of the households in Figure 8 below. Likewise, we add the administration levels of the districts in this picture, which is called admin_level_4. Recall that if one hospital covers a region with a radius of 5 kilometers, the covered area is approximately 80 squared kilometers. If one hospital covers a region with a radius of 10 kilometers, the area covered by the hospital is approximately 314 squared kilometers.

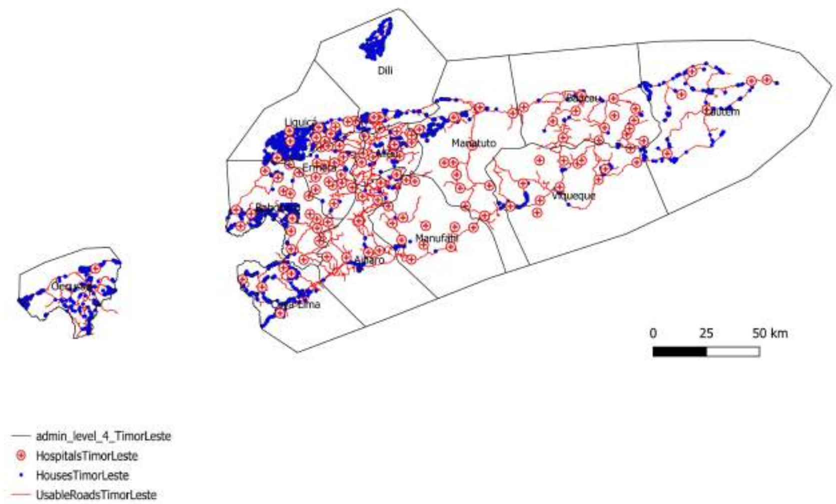


Figure 8: Infrastructure in Timor-Leste

Alieu

The first district we consider is Alieu. From Figure 8 we see that it is located in between other districts, which might be an advantage because if a person has to travel far in its own district, the hospital in another district might be nearby. The statistics of Alieu given in Table 2 are positive. The population that needs to be served per hospital is quite low and the average area each hospital has to cover is in the radius of 5 kilometers. However, the longest average distance that has to be traveled to get to a hospital is almost 20 kilometers. In Figure 8 we can see that there is a lack of hospitals on the top left of the district. Therefore, we conclude that one extra hospital should be located in this district.

Ainaro

The statistics of Ainaro are quite good. The population per hospital is not too high and if the hospitals are perfectly distributed among the district and the road structure is good, everyone in the district should be able to reach a hospital within 5 kilometers. Because the hospitals are not perfectly distributed and the roads are never straight from a house to a hospital, the longest distance that has to be traveled to a hospital equals approximately 10 kilometers. By looking to Figure 8, we see that at the bottom of the district, there is a group of houses a little further away from the hospital. Therefore, it is necessary to add one hospital in that neighbourhood.

Atauro

On Atauro, the island which is located in the district of Dili, there are no existing hospitals. The surface of the country is 224 squared kilometers, which means one hospital should be able to cover the whole island when allowing a radius of 10 kilometers around the facility. To cover the whole island with a radius of 5 kilometers around the hospitals, three facilities are needed. To ensure that health care is available for everyone on the island, it is necessary to build three hospitals. The number of inhabitants, which is equal to 9,274, supports the decision of building three hospitals rather than one. The population is too high to be treated by only one hospital.

Baucau

As we can see from Figure 8, this district is stretched out. There are a few regions where no houses are located. Therefore, it is hard to measure whether there are enough hospitals by looking to the area one hospital has to cover. The average number of people that has to be treated by one hospital is a little higher than the average of the whole country. From Figure 8, we see a few areas with a lack of hospitals: the top left of the district, the middle left of the district and the top right of the district. In order to improve the accessibility to hospitals in Baucau, we conclude it is necessary to add a hospital to each of those three areas in Baucau.

Bobonaro

The statistics for this district are very good. The number of people that each hospital has to serve as well as the area one hospital has to cover are low. The longest distance the inhabitants have to travel to a hospital is just over 10 kilometers. By looking at Figure 8, we conclude one hospital can be added in the middle of the district to lower the longest distance that has to be traveled to a hospital.

Cova Lima

This district clearly needs additional hospitals. The population per hospital is far above average and the region that has to be covered by one hospital exceeds the region which is covered with a radius of 5 kilometer around the hospitals. Note that there is an area in the middle of the country that is uninhabited, so the measure that hospitals should cover the whole district is inefficient to use. From Figure 8, we see that in the right bottom of the district there are a lot of houses but there is no hospital. We can even conclude that three hospitals are necessary in that region. Additionally one hospital needs to be added at the bottom left of the district. In order to serve the people living at the upside of the district, it is necessary to build three hospitals there such that everyone can reach a hospital within a reasonable travel distance. So we can conclude that seven hospitals should be added to this district to be sure the population of Cova Lima can more easily reach a hospital facility.

Dili

Since the capital city is located in Dili, the district has a lot of inhabitants and the area is densely populated. Therefore, it is important that there are enough hospitals to serve the population. We see that on the right side of the district, there are no hospitals located. We conclude it is necessary to add at least two hospitals on that side of the district. On the left side of the district, there is a big group of houses without a hospital. In that area an extra hospital is needed. Therefore, we conclude that in the district Dili three additional hospitals are needed.

Ermera

This district performs very well on all measures. The population that has to be served per facility is below the average of the country, the area one hospital has to cover is only 47 squared kilometers and the longest distance that has to be traveled to get to a hospital is below 10 kilometers. Additionally, from Figure 8 we see that the hospitals are fairly distributed among the district. Therefore, we can conclude that this district does not need an additional hospital.

Lautém

We can see in Table 2 that the district performs bad on all measures: there are a lot of people that have to be treated by one hospital, the surface one hospital has to cover is large and the average longest distance that an inhabitant has to travel is over 30 kilometers. From Figure 8 we see that it is necessary to add one hospital in the top left of the district, one at the top right of the district, one in the middle of the district, one in the middle left of the district and one hospital should be located in the bottom right of Lautém. By adding 5 hospitals, the number of people that has to be served by one facility becomes reasonable. One hospital still has to cover an area bigger than 80 squared kilometers, but since the district is stretched out, there parts of the district without any houses. Therefore, we conclude that adding 5 hospitals would be necessary.

Liquiça

Liquiça performs quite well on all measures in Table 2. However, from Figure 8 we see that a lot of houses are located on the left side of the district, but most hospitals are located on the right side of the district. This causes that the longest distance that has to be traveled to reach a hospital is almost 15 kilometers. Therefore, one additional hospital can be located on the left side of the district to be sure the hospitals are better accessible for the population of Liquiça.

Manatuto

By looking at Figure 8, we directly see that there is a lack of hospitals on the top left of the district. When we consider the statistics of Manatuto, we see that the average longest distance that has to be traveled to get to a hospital is almost 30 kilometers. This is way to long. Therefore, we can conclude that there should be added 3 hospitals to the top left of this district, such that everyone can reach a hospital within a reasonable distance.

Manafahi

This district is very stretched out. As we can see in Figure 8, there are multiple areas in the district without any houses. Therefore, the measure about the area one hospital has to cover can be neglected. The population that has to be served by one hospital is low and the longest distance that has to be traveled to a hospital is just above 5 kilometers. Therefore, we conclude that an additional hospital is not needed in Manufahi.

Sar of Oecusse

As we can clearly indicate from Figure 8, there is a lack of hospitals in the district Sar of Oecusse. There is only one hospital located on the right side of the district. As we can see in the table of statistics, there are a lot of people that have to be treated by the same hospital. Additionally, one hospital has to cover the whole district, which means the hospital has to cover 814 squared kilometers. In order to serve the population within 10 kilometers, the surface one hospital should cover is 314 squared kilometers. Applying this measure, we can conclude that at least two hospitals should be added to this district. When we take the same measure but with a range of 5 kilometers, there should be a total of 11 hospitals in Sar of Oecusse. Since there are a couple of areas without houses, 11 hospitals might be too many. By looking at the population, there should be at least seven hospitals in the district to make sure that each hospital has to treat less than 10,000 people. When we want the number of people per hospital be below the average of 8,515, the total number of hospitals needs to be equal to or bigger than nine. Therefore, we conclude that there should be another eight hospitals located in this district, to get a total number of nine hospitals in Sar of Oecusse. This ensures that both measures are reasonably good and everyone in the population should be able to travel to a hospital within a reasonable distance. These new hospitals can be spread over the district.

Viqueque

This district is very stretched out. We see that at the top left of the district, there are no houses located. Therefore, the area one hospital has to cover is not significant. By looking at Figure 8, we see that at the bottom left of the district there are a few houses with no hospitals around. The average population one hospital has to serve is far below the average of the country. We can conclude that one hospital should be located in the neighborhood of the houses at the bottom left in Viqueque.

The whole country

For each district, we describe how many additional hospitals are necessary to ensure the population can more easily reach a hospital. Adding up those numbers results in an extra 37 hospitals in Timor-Leste. So, this results in a total of 176 hospitals located in Timor-Leste including those extra hospitals. The average area one hospital covers is then equal to 85 squared kilometers. Recall that when the hospitals are perfectly distributed and the roads from households to the hospitals were straight, it would mean that one hospital has to cover 80 squared kilometers. Since there are parts that are uninhabited, this coverage seems reasonable. When considering the same measure but with a radius of 10 kilometers, we see that there are fairly enough hospitals located in Timor-Leste. In Figure 9 below, the predicted locations for the newly located hospitals are shown.

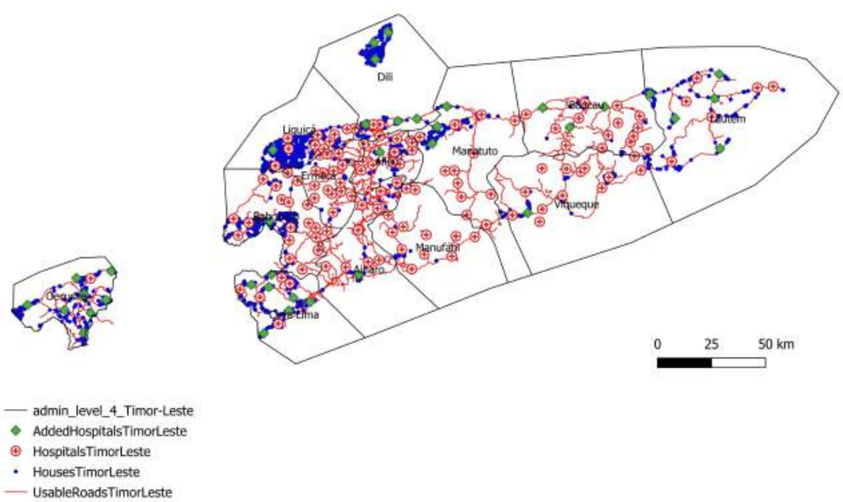


Figure 9: Infrastructure in Timor-Leste including additional hospitals

5 Numerical results

In this section, the results are shown for optimizing the facility locations using the model presented in Section 3. We do this for two different maximum travel distances. We allow people to travel 5 kilometers or 10 kilometers from their houses to a hospital. Additionally, two different scenarios are elaborated:

- The optimization of new hospital locations taking into account the already located hospitals.
- The optimization of new hospitals from scratch. This means we take the locations of the already existing hospitals as possible hospital locations. The already existing hospitals can be chosen as optimal hospital locations in the optimization but this is not necessary.

These optimizations are done for Dili, as well as for the whole country. When optimizing the whole country, we do not take into account Sar of Oecusse and the island Atauro, since they can be seen as separate optimization problems.

Additionally, we want to measure the effect of a different road network on our model. We consider this comparison for the district Dili. To do this, we will run the same model with the second, complete road network of Dili shown in Figure 7 and compare the results to the original road network of Dili shown in Figure 6.

In the first subsection, the results of the facility location optimization in Dili are shown. In the second subsection, the results of the facility location optimization in Timor-Leste, excluding the areas mentioned before, are shown. In the third subsection, we show the impact of using a different road network. In the fourth subsection, we describe the interactive tool that can be used with this model.

5.1 Optimization of hospital locations in Dili

We extract the data as explained in Section 4. As mentioned before, a total of 7 hospitals are already located in Dili. There are 3,861 households in Dili and we have computed 50 possible hospital locations. When we locate a hospital facility at all 57 (possible) hospital locations, everyone in Dili is able to reach a hospital within 5 kilometers, and therefore, also within 10 kilometers.

The population of Dili equals 268,005. This means than on average 69 people live in one house, according to our data. Of course, this number is much lower in reality. However, since the capital city might be densely populated, houses can lay close to each other. Therefore, this data can be seen as representative for reality. So if one household is covered, 69 people are able to reach a hospital within the desired maximum travel distance.

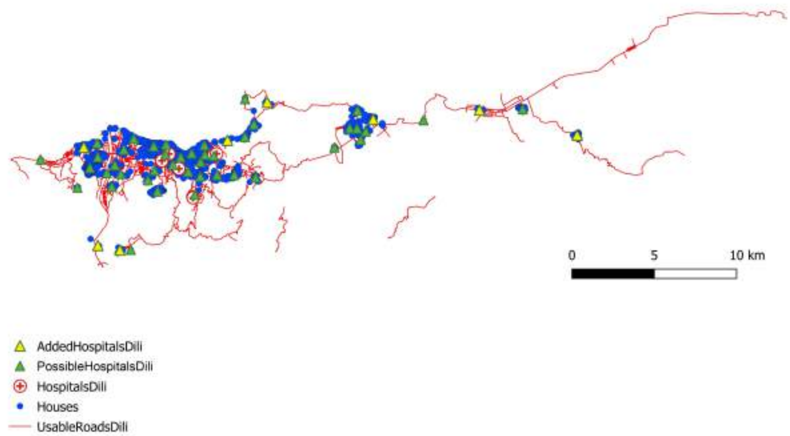
Optimization when taking into account the already existing hospitals

The 7 existing hospitals ensure that approximately 230,175 inhabitants of Dili are able to reach a hospital within a travel distance of 5 kilometers, which is equal to 79.9%. If we extend the maximum travel distance to 10 kilometers, approximately 244,890 inhabitants of Dili can reach a hospital. This equals a coverage of 91.4%. To ensure the entire population of Dili is able to reach a hospital within 5 kilometers, 8 additional hospitals are needed. To ensure everyone of the population is able reach a hospital within 10 kilometers, 3 additional hospitals are needed in Dili. The results of optimizing the hospital locations in Dili for various total numbers of hospitals are shown in Table 3 below. The optimization is done for a maximum travel distance of 5 and 10 kilometers. The number of hospitals in Dili mentioned in the first column of the table represents the total number of hospitals in Dili, including the existing ones.

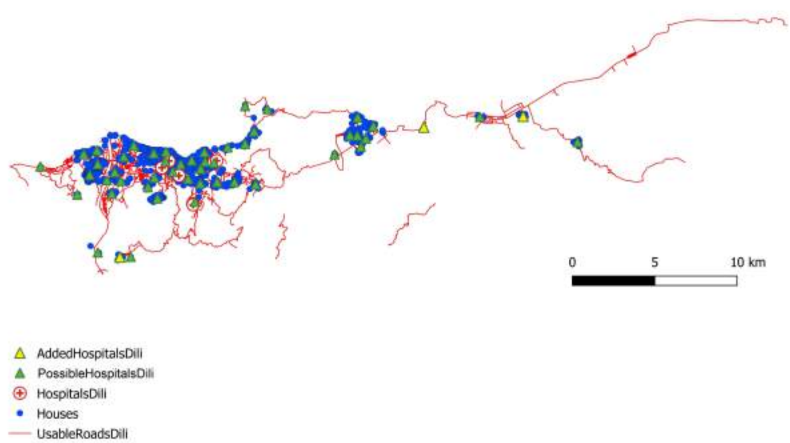
	Maximum distance of 5 km		**Maximum distance of 10 km**	
Hospitals in Dili	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
7	230,175	85.9%	224,890	91.4%
8	247,112	92.2%	265,228	99.0%
9	260,022	97.0%	267,380	99.8%
10	263,771	98.4%	268,005	100.0%
11	265,853	99.2%	268,005	100.0%
12	266,478	99.4%	268,005	100.0%
13	267,103	99.7%	268,005	100.0%
14	267,727	99.9%	268,005	100.0%
15	268,005	100.0%	268,005	100.0%

Table 3: Coverage in Dili taking into account already existing hospitals

In order to get a visualisation of the results, we show the results of optimizing the locations of hospitals on a map. In Figure 10 below, the locations of the extra hospitals needed to ensure everyone in Dili can reach a hospital are shown. The results are shown for a maximum travel distance of 5 and 10 kilometers. As mentioned above, for a maximum travel distance of 5 kilometers we add 8 hospitals to the already existing network. For a maximum travel distance of 10 kilometers, we add 3 hospitals to the already existing hospital network in Dili.



(a) For a maximum travel distance of 5 kilometers



(b) For a maximum travel distance of 10 kilometers

Figure 10: The optimal locations of new hospitals to ensure the coverage is equal to 100% in Dili, while taking into account the already existing hospitals

The optimal locations for a maximum travel distance of 10 kilometers do not match the optimal locations for a maximum travel distance of 5 kilometers. We see that the location at the bottom of the map is chosen to be optimal in both optimizations. At the right side of the map, some hospital locations are chosen in both optimizations. However, the exact location of those additional hospitals depends on the chosen maximal travel distance.

The running time to optimize the model is less than one second. Loading the data into the model takes some time as well, but this only needs to be done once. The optimization of the model with a maximum travel distance of 10 kilometers takes longer than the optimization of the model with a maximum travel distance of 5 kilometers. This is because more houses are within a distance of 10 kilometers from a hospital, so there are more variables and thus more constraints implemented into the model. However, the running time of the optimization is still below one second.

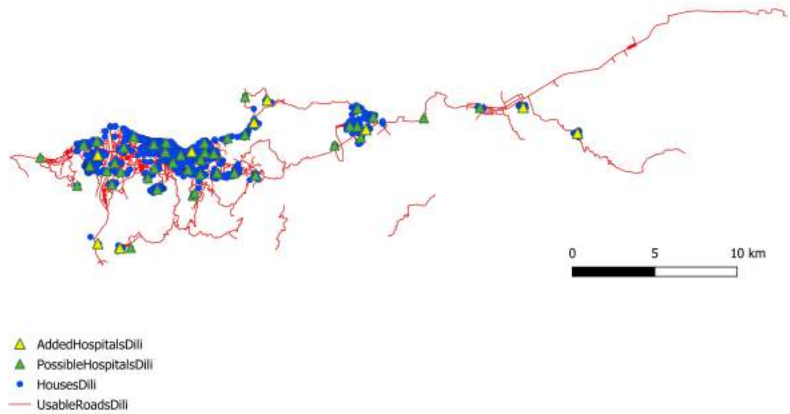
Optimization of hospital locations from scratch

By optimizing from scratch we mean that we do not take into account the already existing hospitals. The locations of the existing hospitals are added to the possible hospital locations. This easily means that with zero hospitals, none of the inhabitants can reach a hospital. To cover at least as many inhabitants as the already existing hospitals do within a travel distance of 5 kilometers, 2 hospitals are needed in Dili. If we consider a maximum travel distance of 10 kilometers, only one hospital can cover approximately the same percentage of inhabitants as the 7 existing hospitals do. To cover all inhabitants in Dili, 9 hospitals are needed when people are able to travel 5 kilometers to a hospital and 4 hospitals are needed when this measure is extended to 10 kilometers. The results are shown in Table 4 below.

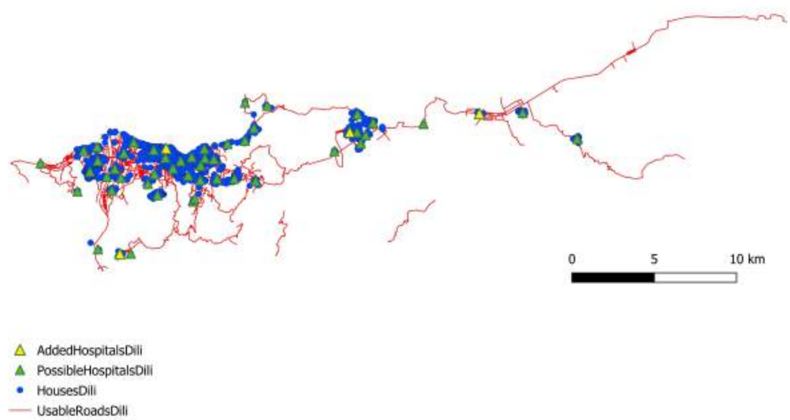
	Maximum distance of 5 km		**Maximum distance of 10 km**	
Hospitals in Dili	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
0	0	0%	0	0%
1	214,140	79.9%	244,335	91.2%
2	241,975	90.3%	264,951	98.9%
3	258,912	96.6%	267,380	99.8%
4	262,660	98.0%	268,005	100.0%
5	264,742	98.8%	268,005	100.0%
6	266,478	99.4%	268,005	100.0%
7	267,102	99.7%	268,005	100.0%
8	267,727	99.9%	268,005	100.0%
9	268,005	100.0%	268,005	100.0%

Table 4: Coverage in Dili from scratch

Again, we show the results for optimizing hospital locations on a map to make the results visible. The optimization is done from scratch, which means we do not take into account the already existing hospitals. For a maximum travel distance of 5 kilometers, a total of 9 optimal hospital locations are selected such that everyone is covered. For a maximum travel distance of 10 kilometers, 4 hospitals are needed to ensure everyone is covered. Those optimal locations are shown in Figure 11.



(a) For a maximum travel distance of 5 kilometers



(b) For a maximum travel distance of 10 kilometers

Figure 11: The optimal locations of new hospitals to ensure the coverage is equal to 100% in Dili, from scratch

Again, there are a lot of differences in the optimal locations of the hospitals for the different maximum travel distances. The location at the bottom of the map is used again for both maximum travel distances. At the right side of the map, at least one hospital location is needed in order to get an optimal result, but not the same location is used for the different travel distances. We can conclude that at least one hospital location is needed in the middle of the map, but again, for different maximum travel distances, the optimal location is different.

The running time from scratch is a little longer compared to the running time of the model including the already existing hospitals. The more hospitals we add to the model, the less nodes have to be explored so the running time is lower. The running time with a maximum travel distance of 5 kilometers is approximately 1.5 seconds. Again, we see that the running time for a maximum travel distance of 10 kilometers is longer than the running time using a maximum travel distance of 5 kilometers. The running time of the optimization with a maximum travel distance of 10 kilometers is approximately 2 seconds.

Compare results

In order to draw the right conclusion for Dili, we compare all results discussed above. We do this using a Pareto curve. Such a curve shows the percentage of coverage against the number of hospitals present in Dili. In Figure 12, four Pareto curves are shown:

- a curve representing the optimization for different numbers of hospitals including the already existing hospitals in Dili with a maximum travel distance of 5 kilometers,
- a curve representing the optimization for different numbers of hospitals including the already existing hospitals in Dili with a maximum travel distance of 10 kilometers,
- a curve representing the optimization for different numbers of hospitals while using the existing hospital locations in Dili as possible hospital locations with a maximum travel distance of 5 kilometers,
- a curve representing the optimization for different numbers of hospitals while using the existing hospital locations in Dili as possible hospital locations with a maximum travel distance of 10 kilometers.

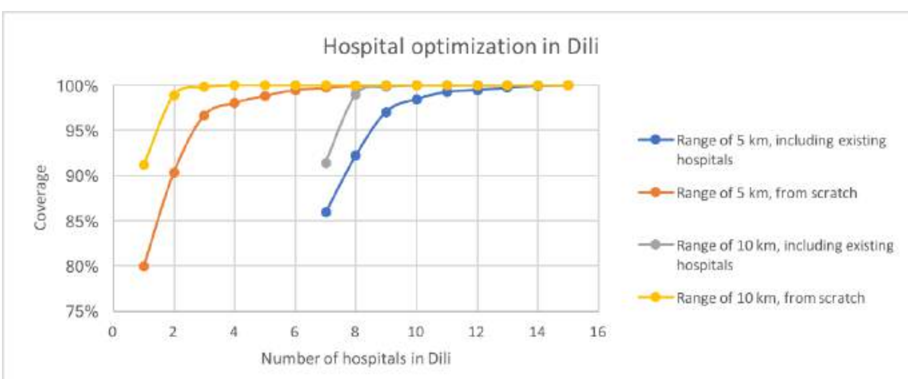


Figure 12: Pareto curves for various number of hospitals in Dili

In this figure, we can clearly see how much coverage grows when adding an extra hospital. Especially the first added hospitals in Dili make a huge difference in coverage. Using Figure 12, a consideration can be made how many hospitals are needed to what costs.

We can also compare Figure 10 to Figure 11. When we add new hospitals from scratch, we see that none of the already existing facility locations is used when building hospitals on the optimal locations. This is true for both maximal travel distances we considered.

Allocation of people to the hospitals

The model we use to optimize facility locations, shown in Section 3, assigns the population to a facility if they are within the maximum travel distance of an opened facility. In Table 5, the number of people assigned to each hospital is given for the optimization of hospital locations using a maximum travel distance of 5 kilometers in Dili. In the second column, the distribution of people is shown for locating a total of 15 hospitals including the already existing hospital locations. The third column shows the distribution of people when locating a total of 9 hospitals from scratch. Note that in both cases, the whole population of Dili is served. Only hospital locations which are optimal in one or both problems are shown in this table.

Hospital ID	**Including existing hospitals: number of inhabitants allocated**	**From scratch: number of inhabitants allocated**
Already existing - 1	18,325	-
Already existing - 2	15,549	-
Already existing - 3	23,739	-
Already existing - 4	25,891	-
Already existing - 5	55,808	-
Already existing - 6	27,557	-
Already existing - 7	63,305	-
New hospital - 4	-	157,082
New hospital - 5	347	-
New hospital - 7	12,911	-
New hospital - 25	16,937	16,937
New hospital - 26	3,748	3,748
New hospital - 28	-	56,711
New hospital - 29	2,082	2,082
New hospital - 34	625	764
New hospital - 37	625	625
New hospital - 49	555	29,778

Table 5: Allocation of the population to hospitals in Dili

The results in Table 5 show that some hospitals serve very few people and other hospitals serve very much people. There are four hospitals that serve less than

1,000 people and two hospitals that serve more than 50,000 people when including the existing hospitals in Dili. For the optimization from scratch, we see that there are two hospitals which serve less than 1,000 inhabitants and 2 hospitals that serve more than 50,000 people.

The model used in this research allocates people to an open hospital randomly when the travel distance is within the maximum travel distance. This is fine for now, since we did not set any capacity constraint. Differently to assigning the population randomly to the hospitals, as much people as possible can be assigned to their closest hospital. Likewise, people could be distributed as equally as possible among the hospitals. This can be formulated as an optimization problem.

5.2 Optimization of hospital locations in Timor-Leste

As mentioned in the introduction of this section, we optimize the hospital locations for Timor-Leste excluding the island Atauro and Sar of Oecusse because they can be seen as separate optimization problems. In the leftover part of Timor-Leste, there are 37,379 households. The total population of this part of Timor-Leste is 1,105,456. This means we have on average 30 people per household according to our data set. This can be taken as representative for the real data.

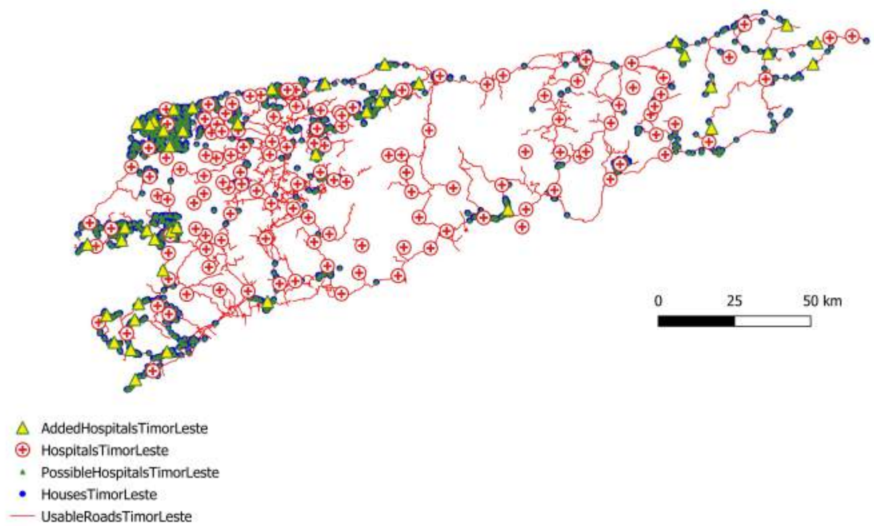
The number of already existing hospitals in this region is equal to 138. Additionally, we take 1000 possible hospital locations. If we use all possible hospital locations, 1,098,624 inhabitants of Timor-Leste can reach a hospital within 5 kilometers, which equals 99,4% of the population. If we allow people to travel 10 kilometers, the total number of inhabitants that can reach a hospital equals 1,099,393 in Timor-Leste. This is equal to a coverage of 99,5%.

Optimization when taking into account the already existing hospitals

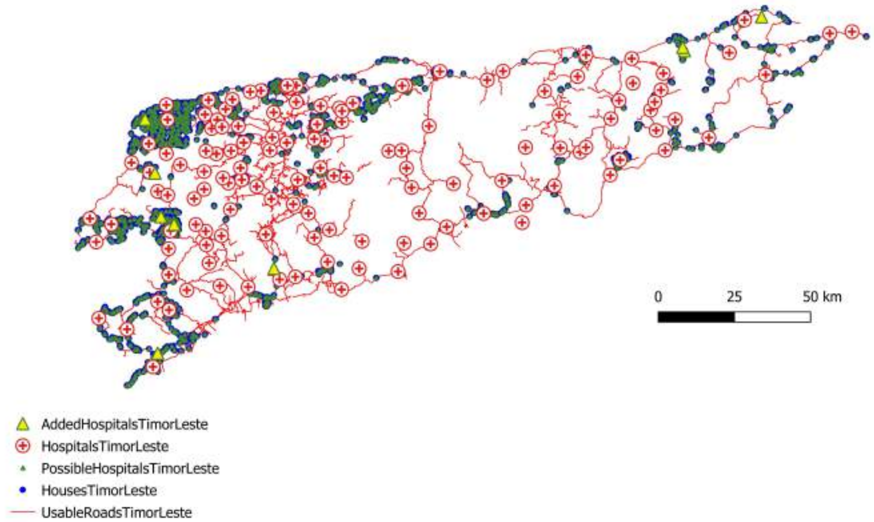
The 138 existing hospitals cover 676,629 inhabitants when allowing a maximum travel distance of 5 kilometers in Timor-Leste. This equals a coverage of 61.2%. If we extend the maximum travel distance to 10 kilometers, 893,823 people in Timor-Leste can reach a hospital, which equals a coverage of 80.9%.

To ensure 95% of the population of Timor-Leste can reach a hospital within 5 kilometers, 42 additional hospitals are needed. This results in a total of 180 hospitals in Timor-Leste. If we extend the maximum travel distance to 10 kilometers, 9 additional hospitals are needed in order to get a coverage of 95%, which results in a total of 147 hospitals in Timor-Leste. The results of optimizing the hospital locations in Timor-Leste for various numbers of hospitals are shown in Table 7 in Appendix A.1. The optimization is done for a maximum travel distance of 5 and 10 kilometers. The number of hospitals in Timor-Leste in mentioned in the first column of the table represents the total number of hospitals in Timor-Leste, including the existing ones.

In order to visualize these results, we show the optimal hospital locations in Timor-Leste on a map. To cover at least 95% of the population with a travel distance of 5 kilometers, 42 additional hospitals are needed. To cover at least 95% of the population with a maximum travel distance of 10 kilometers, 9 additional hospitals are needed. The optimal locations of those extra located hospitals are shown in Figure 13 below.



(a) For a maximum travel distance of 5 kilometers



(b) For a maximum travel distance of 10 kilometers

Figure 13: The optimal locations of new hospitals to ensure the coverage is at least 95% in Timor-Leste, including the already existing hospitals

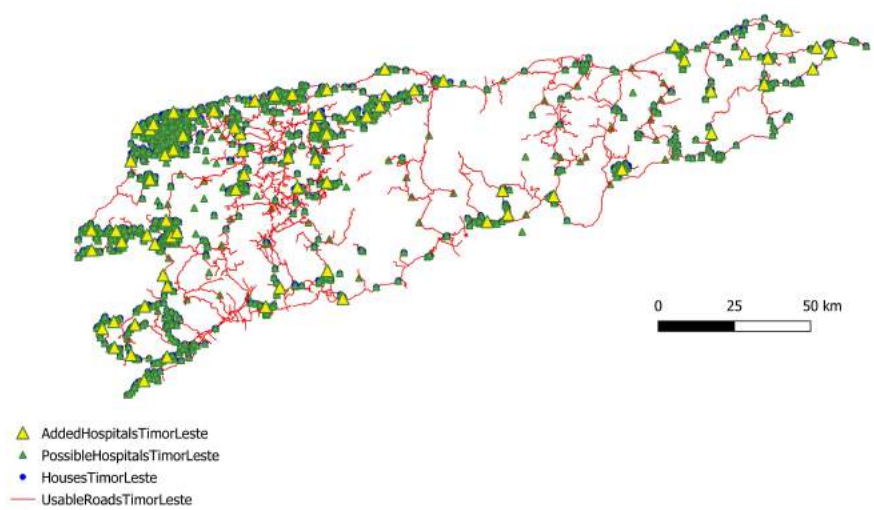
From Figure 13 we see that there are little differences in the hospital locations for a maximum travel distance of 5 kilometers and a maximum travel distance of 10 kilometers. The optimal locations for a maximum travel distance of 10 kilometers are always nearby an optimal hospital location for a maximum travel distance of 5 kilometers, but the exact location is different. For a maximum travel distance of 5 kilometers, we see that a lot of optimal hospital locations are located at the top left of the map, where the population is more densely populated than in the rest of the country.

The running time to optimize the model with a maximum travel distance of 5 kilometers including the already existing hospitals is less than one minute for most of the optimizations. For some different input variables for the total number of hospitals, the running time is a little bit longer, but never above two minutes. The optimization of the model with a maximum travel distance of 10 kilometers takes a little longer than the optimization of the model with a maximum travel distance of 5 kilometers. However, the running time of the optimization is still below two minutes.

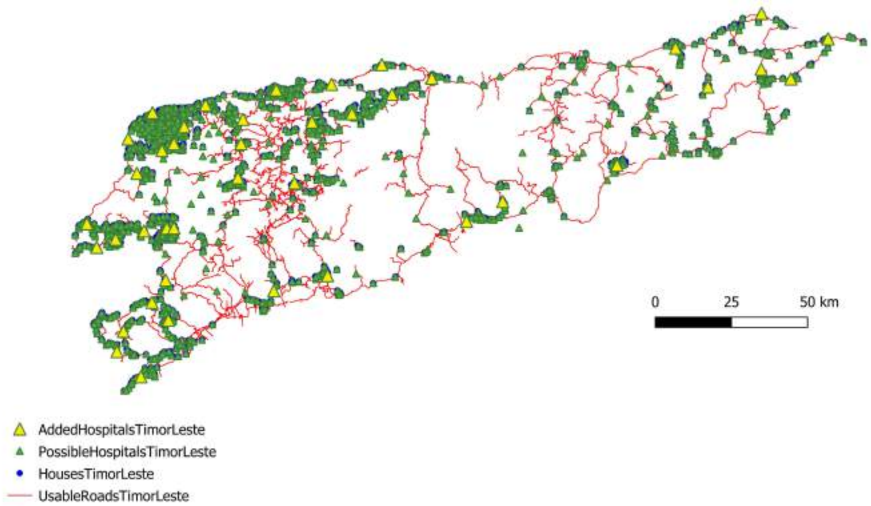
Optimization of hospital locations from scratch

Recall that by optimizing from scratch we mean that we do not take into account the already existing hospital in Timor-Leste. The locations of the existing hospitals are added to the set of possible hospital locations. This means that with zero hospitals, none of the inhabitants of Timor-Leste can reach a hospital. With a maximum travel distance of 5 kilometers, the already existing hospitals cover 61.2% of the population. To get at least the same coverage from scratch with the same maximum travel distance, 17 hospitals are needed. With a maximum travel distance of 10 kilometers, the already existing hospitals covered 80.9%. To get the same coverage from scratch with a maximum travel distance of 10 kilometers, 22 hospitals are needed in Timor-Leste. The results of optimizing the model from scratch for various numbers of hospitals in Timor-Leste are shown in Table 8 in Appendix A.1. The number of hospitals in Timor-Leste mentioned in the first column of the table represents the total number of hospitals in Timor-Leste.

To cover at least 95% of the population with a maximum travel distance of 5 kilometers, 76 hospitals are necessary in Timor-Leste. For a travel distance of 10 kilometers, 41 hospitals are needed to cover at least 95% of the population of Timor-Leste. In order to visualize those results, we again show the optimal locations for those hospitals on a map. The optimal locations from scratch for a coverage of at least 95% for 5 and 10 kilometers are shown in Figure 14 below.



(a) For a maximum travel distance of 5 kilometers



(b) For a maximum travel distance of 10 kilometers

Figure 14: The optimal locations of new hospitals to ensure the coverage is at least 95% in Timor-Leste, from scratch

We notice that with a maximum travel distance of 5 kilometers, a lot of hospitals are located at the left side of the country, which is more densely populated. For a maximum travel distance of 10 kilometers, there are a little more hospitals at the left side of the country compared to the right side of the country, but the difference is way less.

The running time to optimize the model from scratch with a maximum travel distance of 5 kilometers is on average 100 seconds, but never above 3 minutes. This is longer compared to optimizing including the already existing hospitals, which makes sense, because the model has to explore more nodes before ending up in the optimal solution. The running time to optimize from scratch with a maximum travel distance of 10 kilometers is a little longer than the optimization with a maximum travel distance of 5 kilometers. The running time is on average 2 minutes and never above 200 seconds.

Compare results

In order to draw the right conclusion for Timor-Leste, we again compare the results using a Pareto curve. In Figure 15, four Pareto curves are shown:

- a curve representing the optimization for different numbers of hospitals including the already existing hospitals in Timor-Leste with a maximum travel distance of 5 kilometers,

- a curve representing the optimization for different numbers of hospitals including the already existing hospitals in Timor-Leste with a maximum travel distance of 10 kilometers,
- a curve representing the optimization for different numbers of hospitals while using the existing hospital locations in Timor-Leste as possible hospital locations with a maximum travel distance of 5 kilometers,
- a curve representing the optimization for different numbers of hospitals while using the existing hospital locations in Timor-Leste as possible hospital locations with a maximum travel distance of 10 kilometers.

Figure 15: Pareto curves for various number of hospitals in Timor-Leste

This figure clearly shows how much coverage grows when adding an extra hospital to the network. The Pareto curves flatten when the total number of hospitals becomes bigger. Using Figure 15, a consideration can be made how many hospitals are needed to what costs.

We can compare Figure 13 to Figure 14. When adding new hospitals from scratch, only 3 of the already existing hospitals are chosen to be optimal. This is true for a maximum travel distance of 5 kilometers and a maximum travel distance of 10 kilometers.

Additionally, we can compare the results of this optimization to the description of extra hospitals to be added in Section 4.8. This means we compare Figure 9 to Figure 13. In both maps, the majority of new located hospitals is on the left side of the country. Furthermore, we see some similar optimal locations, but definitely not all of the locations are corresponding. For example in the region of Baucau (at the top middle of the country) we describe in Section 4 that it is necessary to add 3 hospitals. Using our model, there is no possible hospital location marked as optimal in that region.

5.3 Impact of a different road network

In order to measure what impact a different road network has on our model, we compute a second road network containing all roads in Dili as described in Section 4. The originally used road network contains roads which can be used to travel to the hospital. The new road network contains roads which might not be useful for this purpose, but measuring the difference might give clarity about spending money on road construction instead of spending it on building extra hospitals. The second road network contains 6.2% more kilometers than the original road network. The results of the optimization including already existing hospitals with a maximum travel distance of 5 kilometers in Dili for both road networks are shown in Table 6 below.

	Using the original road network		**Using the complete road network**	
Hospitals in Dili	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
7	230,175	85.9%	230,452	86.0%
8	247,112	92.2%	247,389	92.3%
9	260,022	97.0%	260,022	97.0%
10	263,771	98.4%	263,771	98.4%
11	265,853	99.2%	265,853	99.2%
12	266,478	99.4%	266,964	99.6%
13	267,103	99.7%	267,589	99.8%
14	267,727	99.9%	268,005	100.0%
15	268,005	100.0%	268,005	100.0%

Table 6: Coverage in Dili for different road networks, taking into account already existing hospitals

The differences in coverage between those networks are small. To cover the whole population of Dili, only 7 additional hospitals are needed using the complete road network, which is equal to 8 using the previous road network. Most of the time, the difference only equals 300 inhabitants, which is on average 0.01% of the population. The increase in coverage is not much compared to the enlargement of 6.2% of the road network.

Note that a large part of Dili is urban, and thus has a complete road network. Especially in the rural areas of Dili, roads are marked as unusable to travel to the hospital. Therefore, the second road network mostly shows additional roads in those rural areas. Since the other districts of Timor-Leste are mostly rural, improving the road network in the whole country might have a big impact on access to the facilities, especially for those living in rural areas.

5.4 Interactive tool

To get the results of this research visible for different input variables, an interactive tool has been developed in association with World Bank. This tool uses the model we developed to identify the optimal locations of hospitals. As we use different input variables, the objective value will change, as well as which facilities are opened.

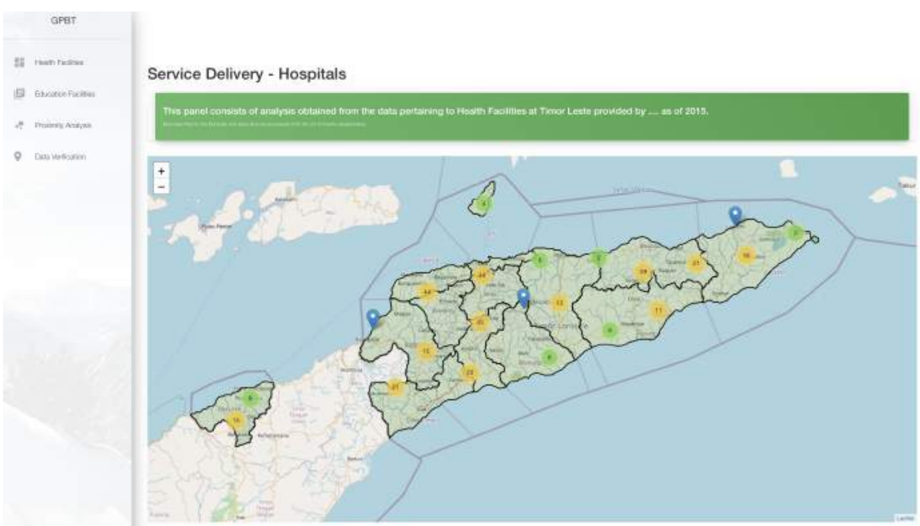
Input variables

In this interactive tool, a few input variables can be changed. The first variable which can be adjusted is the maximum travel distance. Of course, when we increase the maximum travel distance, more people can reach a facility within the desired maximum travel distance. However, when the maximum travel distance is too high, it is not realistic that people actually can reach a facility. Therefore, it is necessary to pick a realistic maximum travel distance.

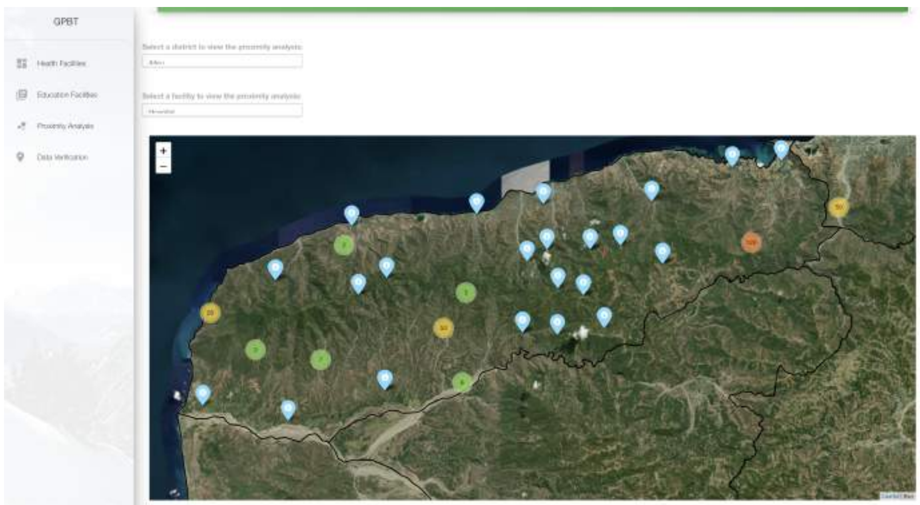
The second variable which can be changed is the number of hospitals located in the country. Recall that once the budget is available, the constraint restricting the number of variables can be replaced by the budget constraint. Comparing the coverage for different numbers of hospitals in the country makes it possible to consider what is the best coverage to what costs.

Output

When all necessary variables are set in the tool, the underlying model runs. Below in Figure 16, two screenshots of the tool are shown. The points marked in those screenshots are the optimal hospital locations obtained from running the model. Note that this is just an example of the interactive tool and the data can not be interpreted.



(a) Of the whole country



(b) Zoomed in on the district Alieu

Figure 16: Screenshots of the optimization tool

6 Conclusion, recommendations and future research

As put forward in Section 1, infrastructure development is very important in developing countries. Therefore, this research focuses on developing a model to optimize facility locations in order to maximize the number of inhabitants that can reach a facility within a maximum desired travel distance. The model needs to be generic such that it can be used for various facilities and multiple countries. However, the method we used and the model have some limitations.

The first subsection contains conclusions on the main goal of this research. The next four subsections include conclusions drawn about different aspects of the research. Additionally, some limitations and recommendations for these aspects are given. The following aspects will be discussed: the data set we have used in our research, the scalability of the model, the interactive tool we developed in corporation with World Bank and the effect of improving the road network. Those aspects are shown in the second to the fifth subsection, respectively. In the sixth subsection, possible avenues for future research are proposed.

6.1 Main goal

In this subsection, the achievements regarding the main goal of this research are described.

Conclusion

The main goal of this research is to develop an optimization model that can optimize the locations of facilities in order to maximize the number of inhabitants that can reach a facility within a desired maximum travel distance, given the number of new located facilities. Moreover, the model should be such that it can be solved fast enough for different input variables. This is necessary such that the model can be used as an interactive tool. As an application to our model, we optimize the locations of hospitals in Timor-Leste and interpret the results.

We can conclude that we succeed to develop a model which meets the requirements for the main goal of this research. The model we developed optimizes the locations of hospitals. The objective function of the model maximizes the number of inhabitants that can reach a facility within the a maximum desired travel distance. Additionally, the number of hospitals located in the country can be chosen. The running time of the model is only a few minutes, so the model can be used as an interactive tool.

Additionally, we implemented the data set of Timor-Leste to our model and extracted some numerical results. Therefore, we can conclude that the model is successfully applied to Timor-Leste and it can be applied to other countries in the same way. This report clearly describes how to implement the data such that the numerical results can be obtained for other areas as well.

6.2 Data set

As an application to our model, we implemented data of Timor-Leste to optimize hospital locations. In this subsection, conclusions about the data set used are given, as well as limitations and recommendations.

Conclusion

In Timor-Leste, the 138 already existing facilities cover approximately 61% of the population, considering a maximum travel distance of 5 kilometers. If we optimize hospital locations with the same maximum travel distance, but without taking into account the already existing hospitals, only 17 hospitals are needed to achieve the same coverage. Considering a maximum travel distance of 10 kilometers, approximately 81% of the population can reach one of the 138 existing hospitals in Timor-Leste. By optimizing from scratch, only 22 hospitals are needed to get to the same result.

To ensure at least 95% of the population of Timor Leste can reach a hospital within 5 kilometers, we conclude 42 extra hospitals are needed additional to the already existing ones. This results in a total of 180 hospitals in Timor Leste. If we consider a maximum travel distance of 10 kilometers, 9 additional hospitals are needed to cover at least 95% of the population. This results in a total of 147 hospitals in Timor-Leste.

Those results are promising for future investment decisions. The results of our model on the application of Timor-Leste show that using a model is very helpful in finding the optimal locations. The model ensures that the investment is used to maximize the benefits of the inhabitants of developing countries.

Limitations and recommendations

Despite the good intention to extract the best available data, we were restricted in this phase of the research. The extracted networks are constructed using GIS (Geographic Information Systems) data available from open-source data sources. This results in data which is representative for the real data, but the set of existing hospital locations, houses and roads is not complete. Moreover, data always differs from reality, especially in developing countries. Therefore, it is important that the problem will be re-optimized when more complete or better data is available.

Especially in developing countries, data collection is shortcoming compared to the rest of the world, which results in a huge gap between data and reality. To be sure that geographical optimization models reach their intended goals, it is important that the data is as close to reality as possible. Therefore, it is recommended that the user of this model is actively engaged in data collection in the developing countries. Better data makes optimizing the locations of facilities more accurate which ensures maximization of the benefits of the population.

6.3 Scalability of the model

An important aspect of our research is to formulate a model which can be easily adjusted. Therefore, the scalability of the model is important. In this subsection, conclusions, limitations and recommendations about the scalability of the model are given.

Conclusion

The aim of the developed model is to use it as an interactive tool. This means that the running time of the model should be limited to a few minutes. In Section 3.4, it is given how to speed up the running time of the model, which is successfully applied. As mentioned before in Section 5, the running time of optimizing the hospital locations without taking into account the existing hospitals for a maximum travel distance of 10 kilometers takes most time of all optimizations we have done in this research. The running time of those optimizations are always within a few minutes. Therefore, we can conclude the performance of the model combined with the data set we used is good.

Limitations and Recommendations

The performance of the model is very good at this point, but the problem of Timor-Leste is not as big as it is when the data set is complete. At the moment, we optimized the model with approximately 38,000 households, whereas there will be 280,000 households if we assume there live on average 4 people in one house in Timor-Leste and every household is included in our data set. This means the number of variables and constraints in our optimization model expands with factor 3.5, not even taking into account for example a bigger set of (possible) hospital locations. Since we want to use the model as an interactive tool and not wait hours and hours for our optimization model to be finished, we need to find a way to improve the running time of our model.

A solution can be found by diving deeper into the model and try to use a different formulation of the model to let the optimization run faster. Differently, we can terminate the solver when the gap between the upper and the lower bound is small enough. Another solution can be to cluster the houses in the data set. By doing this, the number of demand nodes and thus the number of constraints in the model will decrease, such that the running time of the optimization will speed up. However, a weight has to be added to the objective of the model. If a cluster with a certain number of people in it is assigned to a facility, the weight will ensure that objective value will increase with this number as well.

6.4 Interactive tool

With the aid of World Bank, we develop an interactive tool to optimize locations of facilities. This subsection describes conclusions about the interactive tool and explains limitations and recommendations.

Conclusion

This study results in developing an interactive tool by World Bank in which multiple scenarios can be implemented. This tool can be deployed by World Bank to support the decision of the ministry of finance in optimizing facility locations for various facilities in different countries, if the running time is kept low as discussed before. The variables can easily be adjusted in the tool to ensure that a consideration can be made in for example the coverage against the investment.

To make a good consideration, it is necessary to be able to change the input variables in an easy way. This to be sure that people without knowledge of optimization techniques can also use output of the model. The input variables are limited to the maximum desired travel distance and the number of hospitals to be located, including the already existing hospitals. The output of the model includes a map with the optimal hospital locations.

Limitations and recommendations

At the moment, the tool is easily usable for everyone, even without any knowledge. However, it might be interesting to expand the input variables. This makes the tool harder to use, but more realistic and it can be used for multiple purposes. We can include the decision whether the already existing hospitals need to be included in the optimization or if we want to optimize from scratch. Differently, we can add variables for the road network used in the model, such that you can add or neglect each class of roads. When the road network is easily changeable, the impact of a different network can directly be observed.

6.5 Improving the road network

To measure the effect of a different road network on our model, we extract a second road network. Conclusions about using a different road network are given in this subsection, as well as limitations and recommendations.

Conclusion

In Section 5.3, we import a different road network into the model to measure the impact of a change in road network to our model. We measure the effect of the road network on the optimization of facility locations in Dili, assuming a maximum travel distance of 5 kilometers and taking into account the already existing hospital locations. The second road network especially has additional roads in rural areas compared to the original network, which makes it easier for the people living in those areas to reach a facility. The road network in the urban area remains approximately unchanged. The second road network contains 6.2% more kilometers compared to the original network.

Using the original road network, a total of 15 hospitals are needed to cover the whole population of Dili with a maximum desired travel distance of 5 kilometers. Importing the second road network into our model results in a little improvement. Using this second road network, only 14 hospitals are needed to ensure everyone of the population in Dili can reach a hospital, where the already existing hospitals are included again. The overall effect of the improved road network is not much. Only an extra 0.01% of the inhabitants are able reach a hospital using the enlarged road network.

Limitations and recommendations

Since we only measured the impact of a different road network on the district Dili, the results might not be generalizable. The district Dili has a huge urban area, which means that most of the roads in that area are already present in the original road network. To prove a different road network has effect on our model, it can be interesting to measure the effect on a road network to another part of the country, where more rural roads are present. Additionally, one can compare a marginal investment on improving the road networks versus building a new facility. Hereby, a decision can be made whether it is better to revise the road network or build new hospital facilities.

6.6 Future research

In this subsection, possible directions for future research are proposed.

Adding weights to areas

Currently, the whole population has the same weight in our model. However, based on diseases or more elderly people in certain regions, there are some areas which require hospitals more compared to other areas. By adding a weight, the model is more likely to assign a hospital to needy areas. Future work can be based on determining which areas have a lot of elderly people or which areas are in great need of a hospital in order to add a weight to the right areas, and make it more likely that a hospital is assigned to those areas.

Including capacity constraints

In this research, we did not take into account capacity. This is because for the already existing hospitals, capacity is not known. The percentage of people in the need of a hospital needs to be examined as well. Future work can be based on the collection of the necessary data to include capacity. The model has to be adjusted such that the capacity can be implemented. This can easily be done by including a constraint restricting the capacity of each facility and possibly make the costs of building a hospital partly dependent of the capacity of that facility. Instead of only considering building new hospitals, the decision also can be made to expand one of the existing hospital facilities. It is not very complicated to include capacity into the model and it makes the model much more realistic.

Include personnel restrictions

In developing countries, not everyone of the population has access to good education. Therefore, hospitals can be expanded and added to the hospital network in order to increase capacity and serve more people, but if there is no personnel to take care of patients, this makes no sense. It is important to consider the availability of staff in certain areas before adding a hospital. Differently, it can be examined whether there is a great lack of hospital staff. If this is the case, future work could be based on improving education centers for health care in order to simulate the population to study in this field.

Other kinds of facility locations

In this paper, the model is applied to locating hospitals in a country. The model is formulated in such a way that it can be used for other facilities as well. In most of the developing countries, there is a lack of primary care units. Instead of focusing on hospitals in developing countries, it might be interesting to examine locations for primary care units first, to be sure that everyone of the population is able to reach a doctor.

The model can also be used for facilities in different sectors. However, using the model for other facilities might bring other assumptions. For example, when optimizing locations for schools, we have to take into account that kids probably walk to schools in developing countries. This ensures the maximum desired travel distance to schools will be much lower than the maximum desired travel distance to a hospital. In this research, we considered roads which can be used by a four-wheel vehicle. If kids walk to school, a footpath can be marked as travalable as well. This example shows that for other types of facilities, making different assumptions is important.

Road network optimization

As mentioned above, the model developed in this research van be used to optimize other facilities as well. Additionally, we can import a different road network into our model to measure the impact of that network to the optimization of facility locations. However, to determine exactly where new roads are needed or which roads should be renovated, other optimization techniques are required, because those optimizations involve a network structure with dependencies between roads. Therefore, the model used in this research is not suitable to optimize the road network.

Integration of the models

In order to make a good consideration between the different aspects of infrastructure to invest in, a subsequent step of this research can be to integrate the different optimization models. In particular, the model to optimize hospital locations, the model to optimize school locations and the model to optimize the road network can be combined. Between those models, a connection can be made to ensure a good consideration is made on how to invest in infrastructure in a country in order to maximize the benefits of the inhabitants.

References

[1] J.H. Ablanedo-Rosas et al. "Allocation of emergency and recovery centers in Hidalgo, Mexico". In: *International Journal of Service Sciences, 2 (2), 206-218* (2009).

[2] A.D. Asante et al. "Retaining doctors in rural Timor-Leste: a critical appraisal of the opportunities and challenges". In: *Bulletin of the World Health Organization 2014, 92, 277-282* (2018).

[3] S.C.K. Chu and L. Chu. "A modeling framework for hospital location and service allocation". In: *International Transactions in Operational Research, 7, 539-568* (2012).

[4] R. Church and C. ReVelle. "The maximal covering location problem". In: *Papers of the Regional Science Association, 32, 101-118* (1974).

[5] M.S. Daskin. *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Blackwell, 2013.

[6] M.S. Daskin. "What you should know about location modeling". In: *Naval Research Logistics 55, 283-294* (2008).

[7] J. Dekle et al. "A Florida county locates disaster recovery centers". In: *Interfaces, 35 (2), 133-139* (2005).

[8] D.M. Hakimi. "Optimum locations of switching centers and the absolute centers and medians of a graph". In: *Operations Research 12, 450-459* (1964).

[9] H. Jia, F. Ordóñez, and M.M. Dessouky. "Solution approaches for facility location of medical supplies for large-scale emergencies". In: *Computers Industrial Engineering , 52 (2), 257-276* (2007).

[10] P. Murali, F. Ordóñez, and M.M. Dessouky. "Facility location under demand uncertainty: response to a large-scale bio-terror attack". In: *Socio-Economic Planning Sciences, 46 (1), 78-87* (2012).

[11] S. Rahman and D.K. Smith. "Use of location-allocation models in health service development planning in developing nations". In: *European Journal of Operational Research, 123, 437-452* (2000).

[12] S.U. Rahman and D.K. Smith. "Deployment of rural health facilities in a developing country". In: *The Journal of the Operational Research Society, 50 (9), 892-902* (1999).

[13] S.S.R. Shariff, N.H. Moin, and M. Omar. "Location allocation modeling for healthcare facility planning in Malaysia". In: *Computers and Industrial Engineering, 62, 1000-1010* (2012).

[14] ReVelle C. Toregas C.S.R. and C. Bergman. "The location of emergency service facilities". In: *Operations Research, 19, 1363-1373* (1971).

A Appendices

A.1 Numerical results for Timor-Leste

Table 7: Coverage in Timor-Leste taking into account already existing hospitals

	Maximum distance of 5 km		**Maximum distance of 10 km**	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
138	676,629	61.2%	893,823	80.9%
139	761,981	68.9%	980,446	88.7%
140	784,191	70.9%	995,972	90.1%
141	802,882	72.6%	1,005,702	91.0%
142	820,597	74.2%	1,014,959	91.8%
143	834,556	75.5%	1,023,683	92.6%
144	847,687	76.7%	1,032,408	93.4%
145	858,748	77.7%	1,041,073	94.2%
146	868,773	78.6%	1,048,289	94.8%
147	877,261	79.4%	1,055,416	95.5%
148	885,571	80.1%	1,060,651	95.9%
149	893,852	80.9%	1,065,176	96.4%
150	902,133	81.6%	1,069,494	96.7%
151	910,000	82.3%	1,073,812	97.1%
152	917,278	83.0%	1,077,952	97.5%
153	925,704	83.7%	1,081,057	97.8%
154	932,979	84.4%	1,083,956	98.1%
155	940,195	85.1%	1,086,617	98.3%
156	947,234	85.7%	1,088,362	98.5%
157	953,888	86.3%	1,090,018	98.6%
158	960,306	86.9%	1,091,349	98.7%
159	965,806	87.4%	1,092,650	98.8%
160	971,278	87.9%	1,093,567	98.9%
161	976,601	88.3%	1,094,484	99.0%
162	981,747	88.8%	1,095,371	99.1%
163	986,804	89.3%	1,096,258	99.2%
164	991,832	89.7%	1,097,501	99.2%
165	996,830	90.2%	1,097,944	99.3%
166	1,001,562	90.6%	1,098,181	99.3%
167	1,006,027	91.0%	1,098,417	99.3%
168	1,010,404	91.4%	1,098,595	99.4%
169	1,014,367	91.8%	1,098,743	99.4%
170	1,018,271	92.1%	1,098,861	99.4%
171	1,022,027	92.5%	1,098,950	99.4%
172	1,025,606	92.8%	1,099,009	99.4%
173	1,029,095	93.1%	1,099,127	99.4%
174	1,032,496	93.4%	1,099,068	99.4%

Continued on next page

Table 7 – continued from previous page

	Maximum distance of 5 km		Maximum distance of 10 km	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
175	1,035,779	93.7%	1,099,127	99.4%
176	1,039,062	94.0%	1,099,157	99.4%
177	1,042,345	94.3%	1,099,186	99.4%
178	1,045,450	94.6%	1,099,216	99.4%
179	1,048,407	94.8%	1,099,245	99.4%
180	1,051,158	95.1%	1,099,275	99.4%
181	1,053,642	95.3%	1,099,305	99.4%
182	1,056,037	95.5%	1,099,334	99.4%
183	1,058,226	95.7%	1,099,364	99.4%
184	1,059,971	95.9%	1,099,393	99.5%
185	1,061,657	96.0%	1,099,393	99.5%
186	1,063,254	96.2%	1,099,393	99.5%
187	1,064,851	96.3%	1,099,393	99.5%
188	1,066,418	96.5%	1,099,393	99.5%
189	1,067,838	96.6%	1,099,393	99.5%
190	1,069,168	96.7%	1,099,393	99.5%
191	1,070,470	96.8%	1,099,393	99.5%
192	1,071,771	97.0%	1,099,393	99.5%
193	1,073,013	97.1%	1,099,393	99.5%
194	1,074,255	97.2 %	1,099,393	99.5 %
195	1,075,438	97.3 %	1,099,393	99.5 %
196	1,076,592	97.4 %	1,099,393	99.5 %
197	1,077,686	97.5 %	1,099,393	99.5 %
198	1,078,573	97.6 %	1,099,393	99.5 %
199	1,079,431	97.6 %	1,099,393	99.5 %
200	1,080,259	97.7 %	1,099,393	99.5 %
201	1,081,057	97.8 %	1,099,393	99.5 %
202	1,081,797	97.9 %	1,099,393	99.5 %
203	1,082,477	97.9 %	1,099,393	99.5 %
204	1,083,157	98.0 %	1,099,393	99.5 %
205	1,083,808	98.0 %	1,099,393	99.5 %
206	1,084,429	98.1 %	1,099,393	99.5 %
207	1,085,050	98.2 %	1,099,393	99.5 %
208	1,085,612	98.2 %	1,099,393	99.5 %
209	1,086,174	98.3 %	1,099,393	99.5 %
210	1,086,735	98.3 %	1,099,393	99.5 %
211	1,087,268	98.4 %	1,099,393	99.5 %
212	1,087,859	98.4 %	1,099,393	99.5 %
213	1,088,392	98.5 %	1,099,393	99.5 %
214	1,088,835	98.5 %	1,099,393	99.5 %
215	1,089,279	98.5 %	1,099,393	99.5 %
216	1,089,663	98.6 %	1,099,393	99.5 %
217	1,090,018	98.6 %	1,099,393	99.5 %

Continued on next page

Table 7 – continued from previous page

	Maximum distance of 5 km		Maximum distance of 10 km	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
218	1,090,344	98.6 %	1,099,393	99.5 %
219	1,090,610	98.7 %	1,099,393	99.5 %
220	1,090,876	98.7 %	1,099,393	99.5 %
221	1,091,142	98.7 %	1,099,393	99.5 %
222	1,091,408	98.7 %	1,099,393	99.5 %
223	1,091,674	98.8 %	1,099,393	99.5 %
224	1,091,941	98.8 %	1,099,393	99.5 %
225	1,092,207	98.8 %	1,099,393	99.5 %
226	1,092,443	98.8 %	1,099,393	99.5 %
227	1,092,680	98.8 %	1,099,393	99.5 %
228	1,092,917	98.9 %	1,099,393	99.5 %
229	1,093,153	98.9 %	1,099,393	99.5 %
230	1,093,390	98.9 %	1,099,393	99.5 %
231	1,093,626	98.9 %	1,099,393	99.5 %
232	1,093,656	98.9 %	1,099,393	99.5 %
233	1,094,070	99.0 %	1,099,393	99.5 %
234	1,094,277	99.0 %	1,099,393	99.5 %
235	1,094,484	99.0 %	1,099,393	99.5 %
236	1,094,691	99.0 %	1,099,393	99.5 %
237	1,094,898	99.0 %	1,099,393	99.5 %
238	1,095,075	99.1 %	1,099,393	99.5 %
239	1,095,253	99.1 %	1,099,393	99.5 %
240	1,095,430	99.1 %	1,099,393	99.5 %

Table 8: Coverage in Timor-Leste when optimizing from scratch

	Maximum distance of 5 km		**Maximum distance of 10 km**	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
0	0	0%	0	0%
1	263,743	23.9%	296,866	26.9%
2	359,860	32.6%	403,215	36.5%
3	389,966	35.3%	449,677	40.7%
4	418,978	37.9%	493,890	44.7%
5	447,813	40.5%	535,915	48.5%
6	474,815	43.0%	577,851	52.3%
7	501,431	45.4%	609,230	55.1%
8	524,647	47.5%	640,490	57.9%
9	546,976	49.5%	666,367	60.3%
10	567,352	51.3%	689,997	62.4%
11	587,019	53.1%	713,508	64.5%
12	605,829	54.8%	735,512	66.5%
13	623,603	56.4%	757,367	68.5%
14	641,140	58.0%	777,980	70.4%
15	657,850	59.5%	797,795	72.2%
16	673,968	61.0%	815,037	73.7%
17	689,110	62.3%	831,835	75.2%
18	703,571	63.6%	848,219	76.7%
19	717,797	64.9%	864,337	78.2%
20	730,366	66.1%	879,124	79.5%
21	742,669	67.2%	893,379	80.8%
22	754,143	68.2%	907,604	82.1%
23	765,145	69.2%	920,144	83.2%
24	775,407	70.1%	931,530	84.3%
25	785,226	71.0%	941,792	85.2%
26	794,956	71.9%	950,694	86.0%
27	804,154	72.7%	959,477	86.8%
28	813,174	73.6%	968,202	87.6%
29	822,135	74.4%	976,778	88.4%
30	830,918	75.2%	984,823	89.1%
31	839,672	76.0%	992,601	89.8%
32	848,130	76.7%	999,817	90.4%
33	856,441	77.5%	1,006,944	91.1%
34	864,722	78.2%	1,013,658	91.7%
35	872,884	79.0%	1,020,341	92.3%
36	880,958	79.7%	1,026,522	92.9%
37	888,825	80.4%	1,032,408	93.4%
38	896,691	81.1%	1,037,642	93.9%
39	903,967	81.8%	1,042,285	94.3%
40	911,183	82.4%	1,046,869	94.7%
41	918,221	83.1%	1,051,424	95.1%

Continued on next page

Table 8 – continued from previous page

	Maximum distance of 5 km		Maximum distance of 10 km	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
42	924,905	83.7%	1,055,742	95.5%
43	931,559	84.3%	1,060,030	95.9%
44	937,800	84.8%	1,063,845	96.2%
45	943,803	85.4%	1,067,276	96.5%
46	949,481	85.9%	1,070,470	96.8%
47	954,864	86.4%	1,073,427	97.1%
48	960,128	86.9%	1,075,409	97.3%
49	965,274	87.3%	1,077,360	97.5%
50	970,272	87.8%	1,079,283	97.6%
51	975,063	88.2%	1,081,146	97.8%
52	979,440	88.6%	1,082,713	97.9%
53	983,403	89.0%	1,084,015	98.1%
54	987,307	89.3%	1,085,286	98.2%
55	991,063	89.7%	1,086,558	98.3%
56	994,671	90.0%	1,087,504	98.4%
57	998,249	90.3%	1,088,392	98.5%
58	1,001,828	90.6%	1,089,279	98.5%
59	1,005,377	90.9%	1,090,137	98.6%
60	1,008,778	91.3%	1,090,817	98.7%
61	1,012,090	91.6%	1,091,467	98.7%
62	1,015,373	91.9%	1,092,059	98.8%
63	1,018,478	92.1%	1,092,621	98.8%
64	1,021,554	92.4%	1,093,124	98.9%
65	1,024,334	92.7%	1,093,626	98.9%
66	1,027,084	92.9%	1,094,129	99.0%
67	1,029,775	93.2%	1,094,602	99.0%
68	1,032,348	93.4%	1,095,075	99.1%
69	1,034,833	93.6%	1,095,460	99.1%
70	1,037,287	93.8%	1,095,756	99.1%
71	1,039,683	94.1%	1,096,022	99.1%
72	1,042,049	94.3%	1,096,288	99.2%
73	1,044,178	94.5%	1,096,554	99.2%
74	1,046,278	94.6%	1,096,820	99.2%
75	1,048,348	94.8%	1,097,057	99.2%
76	1,050,182	95.0%	1,097,294	99.3%
77	1,051,838	95.1%	1,097,501	99.3%
78	1,053,435	95.3%	1,097,708	99.3%
79	1,055,002	95.4%	1,097,885	99.3%
80	1,056,570	95.6%	1,098,062	99.3%
81	1,057,960	95.7%	1,098,269	99.3%
82	1,059,261	95.8%	1,098,447	99.4%
83	1,060,562	95.9%	1,098,595	99.4%
84	1,061,834	96.1%	1,098,743	99.4%

Continued on next page

Table 8 – continued from previous page

	Maximum distance of 5 km		Maximum distance of 10 km	
Hospitals in Timor-Leste	**Number of inhabitants covered**	**Percentage covered**	**Number of inhabitants covered**	**Percentage covered**
85	1,063,106	96.2%	1,098,861	99.4%
86	1,064,348	96.3%	1,098,950	99.4%
87	1,065,590	96.4%	1,099,038	99.4%
88	1,066,743	96.5%	1,099,098	99.4%
89	1,067,867	96.6%	1,099,127	99.4%
90	1,068,991	96.7%	1,099,157	99.4%
91	1,070,085	96.8%	1,099,186	99.4%
92	1,071,150	96.9%	1,099,216	99.4%
93	1,072,096	97.0%	1,099,245	99.4%
94	1,072,983	97.1%	1,099,275	99.4%
95	1,073,841	97.1%	1,099,305	99.4%
96	1,074,640	97.2%	1,099,334	99.4%
97	1,075,349	97.3%	1,099,364	99.4%
98	1,076,030	97.3%	1,099,393	99.5%
99	1,076,710	97.4%	1,099,393	99.5%
100	1,077,360	97.5%	1,099,393	99.5%
105	1,080,466	97.7%	1,099,393	99.5%
110	1,083,127	98.0%	1,099,393	99.5%
115	1,085,493	98.2%	1,099,393	99.5%
120	1,087,475	98.4%	1,099,393	99.5%
125	1,089,279	98.5%	1,099,393	99.5%
130	1,090,698	98.7%	1,099,393	99.5%
135	1,092,059	98.8%	1,099,393	99.5%
140	1,093,271	98.9%	1,099,393	99.5%
145	1,094,366	99.0%	1,099,393	99.5%
150	1,095,312	99.1%	1,099,393	99.5%
155	1,096,170	99.2%	1,099,393	99.5%
160	1,096,879	99.2%	1,099,393	99.5%
165	1,097,441	99.3%	1,099,393	99.5%
170	1,097,826	99.3%	1,099,393	99.5%
175	1,098,122	99.3%	1,099,393	99.5%
180	1,098,388	99.4%	1,099,393	99.5%
185	1,098,506	99.4%	1,099,393	99.5%
190	1,098,624	99.4%	1,099,393	99.5%

A.2 Python code

```
#!/usr/bin/env python2
# -*- coding: utf-8 -*-
"""
Created on Thu Jul  9 18:29:29 2020

@author: Joyce Antonissen
"""

#To clear all: %reset

#import necessary packages
import gurobipy as gb
from gurobipy import GRB
import pandas as pd
import numpy as np

#If needed: get weights for clusters
#values = pd.read_excel('~/ClusterNumbers.xlsx' )
#values_list = values['CLUSTER_ID'].tolist()
#nclusters = max(values['CLUSTER_ID']) + 1
#v = [1]

#for i in range(nclusters):
#    v.append(values_list.count(i))

#Import data: distances from houses to (possible) hospital locations
dfh = pd.read_csv(r'~/DistancesHospitals.csv')
dfc = pd.read_csv(r'~/DistancesPossibleHospitals.csv')
dfh = dfh.sort_values(by=['destination_id', 'origin_id'])
dfc = dfc.sort_values(by=['destination_id','origin_id'])

#Put into one DataFrame
df = ([dfh, dfc])
df = pd.concat(df)

#Number of (possible) hospital locations, existing hospitals and households
q = df.shape[0]
m = 138 + 1000         #total number of facilities
n = q/nhosp            #total number of households
existinghosp = 138  #number of existing hospitals

#resize the input variables to a distance matrix of nhosp x nhouseholds
dfcolTC = df[['total_cost']]
x = np.resize(dfcolTC,(nhosp,nhouseholds)).T
```

```
dist = pd.DataFrame(x) #distance matrix


#Set the maximum travel distance, the total number of hospitals.
#These numbers can be adjusted.
S = 5000                    #maximum distance
p = existinghosp + 0        #total number of hospitals to be optimized


#THE MODEL

#Decision variables
X = {}
Y = {}

#Make a model
M = gb.Model("Facility location problem")

#Add the decision variables to the model.
for j in range(m):
    X[j] = M.addVar(vtype=gb.GRB.BINARY, name="X"+str((j+1)))

#listI: which j's are added for each I to the model
#listY: which Y_ij are added to the model
listI=[]
listY=[]
for i in range(n):
    a = []
    c = []
    for j in range(m):
        if dist.iloc[i,j] <= S:
            Y[i,j] = M.addVar(lb=0.0,vtype=gb.GRB.CONTINUOUS,
                                            name="Y"+str((i+1,j+1)))
            a.append(j)
            c.append(Y[i,j])
    listI.append(a)
    listY.append(c)

#listJ: which i's are added for each J to the model
listJ=[]
for j in range(m):
    b = []
    for i in range(n):
        if j in listI[i]:
            b.append(i)
    listJ.append(b)
```

```
#Objective, only add existing Y_ij's
obj = gb.LinExpr()
for i in range(n):
    for j in range(m):
        if j in listI[i]:
            obj += Y[i,j]

M.setObjective(obj, gb.GRB.MAXIMIZE)

#Constraints
#Existing hospitals,  set to 1
#To optimize from scratch: comment those constraints
for j in range(existinghosp):
    M.addLConstr(X[j] == 1, name="Existing hospitals"+str((j)))


#Max number facilities (instead of budget) constraint
s = M.addLConstr(gb.quicksum(X[j] for j in range(m))<= p, name = "Budget")

#Assign demand to an open facility
for j in range(m):
    M.addLConstr(gb.quicksum(Y[i,j] for i in listJ[j]) <= n * X[j],
                         name = "Assign demand to open facility"+str(j+1))


#Maximal assign each person to one facility
for i in range(n):
    #M.addSOS(GRB.SOS_TYPE1, listY[i])
    M.addLConstr(gb.quicksum(Y[i,j] for j in listI[i]) <= 1,
                         name="Only assign to one facility"+str((i)))


#Optimize the model
M.optimize()

# Get x and y
var = M.getVars()

#facility opened if x_j = 1
XVAR = var[0:m]
Xvalues = np.zeros(m)
for i in range(m):
    Xvalues[i]=X[i].x
```

```
facility_site_opened = pd.DataFrame(data=Xvalues)

# Get a matrix where Y_ij = 1 if node i is served by facility j
YVAR = var[m:n*m+m]
Yvalues = np.zeros(n*m)
for i in range(n*m):
    if j in listI[i]:
        Yvalues[i]=YVAR[i].x

Yvalues.shape = (n,m)
Served = pd.DataFrame(data=Yvalues)


#Change the number of facilities, easily re-run the model
#M.remove(s)
#p = 138 + 40
#s = M.addConstr(gb.quicksum(X[j] for j in range(m))<= p, name = "Budget")
#M.optimize()
```