Aalborg University

Master Thesis

Mathematics-Economics

# A Multi-Objective Approach for Allocating Fire Stations in Aalborg

**28/05/2025**

AALBORG UNIVERSITY
STUDENT REPORT

**Institute of Mathematical Sciences**
Thomas Manns Vej 23
9220 Aalborg Øst
www.math.aau.dk/

AALBORG UNIVERSITY
STUDENT REPORT

**Title:**

A Multi-Objective Approach for Allocating Fire Stations in Aalborg

**Project:**

Master Thesis

**Project period:**

3. February 2025 - 28. May 2025

**Project group:**

MO10-04

**Author:**

Mikkel Søby Kramer

**Supervisors:**

Inkyung Sung
Alex Elkjær Vasegaard
Magnus Berg Ladefoged

**Pagecount: 56 + front page**

**Abstract:**

This thesis is dedicated to optimizing the fire station locations in Aalborg in cooperation with the emergency department of Northern Jutland. This is done by first analyzing GIS data from the last eight years and identifying the most prominent challenges facing the department. The problem is then formulated as a multi-objective facility location problem seeking to optimize both efficiency and equity in terms of response time. The model includes objectives from both the MCLP and $p$-center formulations.
The problem is solved using the heuristic algorithm NSGA-II, and different scenarios, where the current facility location is either kept or removed, are solved and compared. The solutions are in the form of Pareto fronts, but it is seen that general performance is improved with respect to the objectives when more stations are added.
Lastly, other objectives are explored to analyze the effect of these on the solutions. These include the Gini coefficient and backup coverage, and it is concluded that both of these behave differently than the original objectives, but still achieve an overall improvement compared to the current case.

# Preface

The project period has elapsed from 03-02-2025 to 28-05-2025. Data processing and simulations are done using the programming language `R`. The scripts can be shared by the author upon request.

## Reader's guide

Literature references given by [Source number, page number] refer to the bibliography at the end of the project, and references using letters refer to the appendix.

## Acknowledgments

The work on the project has been supervised by Inkyung Sung, Alex Elkjær Vasegaard, and Magnus Berg Ladefoged during the entire work process, and I am very grateful for their input and feedback. I would also like to thank Nordjyllands Beredskab for the possibility to work with their data and Beredskabschef Lars K. Bjørndal for helpful information and feedback as well.

# Contents

# 1 | Introduction

The fire departments in Denmark handle roughly 40 thousand emergencies each year, spanning everything from normal building fires to drowning accidents [1]. In such situations, a fast response plays a major part in minimizing loss of property and lives, since the probability of a fire creating property damage or costing lives rises by 20% every minute the response time is increased, according to Kiran KC [2]. At the same time, advances in technology and infrastructure mean that the emergency departments must be able to handle a wider variety of emergencies besides fires, such as traffic accidents, natural disasters, and chemical accidents. This, in turn, means that specialized staffing and equipment are required.

These are just some of the problems that face the emergency department in Northern Jutland. This department is responsible for the 11 municipalities in Northern Jutland and operates 36 fire stations scattered across the 7886 $km^2$ area. It is an area that contains both sparsely and densely populated areas, indicating that some areas are more prone to accidents. However, the area is also large enough that 36 stations are not enough the cover the entire area efficiently, and some places can even expect a response time of over 20 minutes. Thus, there is an inequality in the service provided, which is also an important aspect for public services [3].

The 36 stations employ fewer than 1000 workers, but only one of them houses full-time employees, namely the one in Aalborg. The rest are staffed by part-time workers, who are expected to be at the fire station within 5 minutes of an emergency call. This can also pose a challenge to the emergency department, since the fire station needs to be located close to populated areas in order to have a large enough available workforce. At the same time, the part-time employees also have to keep up with the development of the competencies needed to respond to the variety of emergencies, while balancing other work and personal life. Thus, it might seem beneficial to have more full-time stations, but it was explained by the department that upgrading a part-time station to a full-time station would require approximately 10 times the cost in salaries alone.

All of this leads to a complex situation with a variety of challenges facing the emergency department, and the purpose of this project is to analyze GIS data provided by the emergency department in order to identify and solve the most significant problems. The data is from the period 2017-2024 and contains over 50 000 individual data points spanning all of Northern Jutland. Additionally, it contains 63 descriptive variables for each emergency, including response time, geographical location, water usage, vehicle type, and much more. The locations of the emergencies are summarized as a heat map in Figure 1.1.

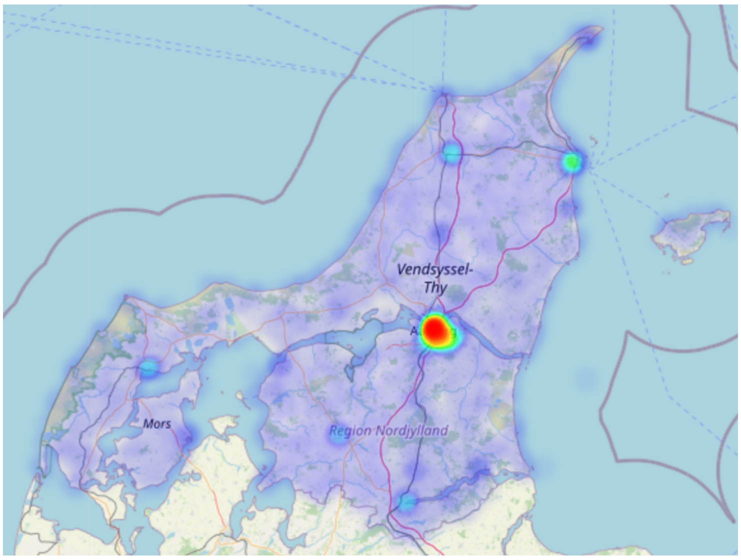


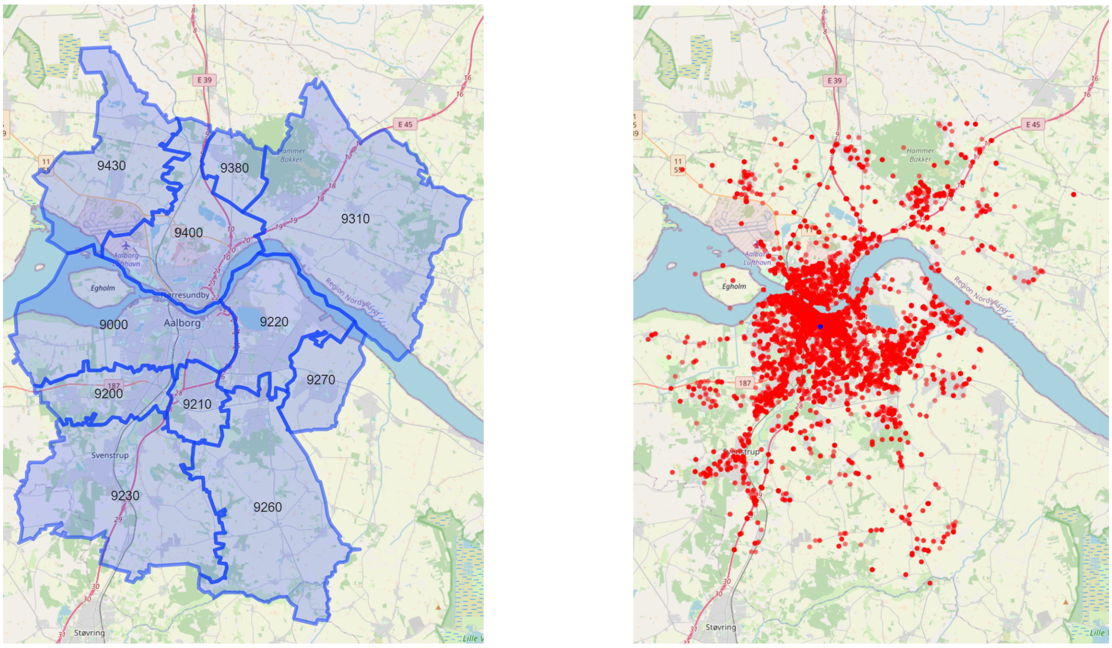

Figure 1.1: Heat map of all emergencies from 2017-2024.

From the heat map, it is clear that most of the emergencies happen in the major cities, primarily Aalborg. Indeed, approximately 35% of the incidents happened in Aalborg municipality. Initial analysis showed that each of the 11 municipalities experienced different behavior in terms of demand, response times, and types of emergencies. Thus, motivated by these differences, the amount of data, and the department's own interest in the possibility of adding another station in Aalborg, this project will focus on the data and situation related specifically to Aalborg and the surrounding area.

## 1.1 Data Exploration and Problem Identification

In order to identify the most significant problems facing the emergency department in Aalborg specifically, the data is first restricted to this area alone. This is done by only including the incidents that happened within the postal codes seen in Figure 1.2a, which uses data from [4]. This area is chosen based on the geographical location of Aalborg and the fact that it contains 92% of the incidents handled by the Aalborg NOBR fire station.

In Figure 1.2b, all of the incidents within the area explained above are shown as red dots, while the current station is shown in blue. It is clear that geographically, most of the incidents are located close to downtown, where the population is also higher. These incidents will be analyzed in the following section, first with a focus on the number of incidents, corresponding to the demand.

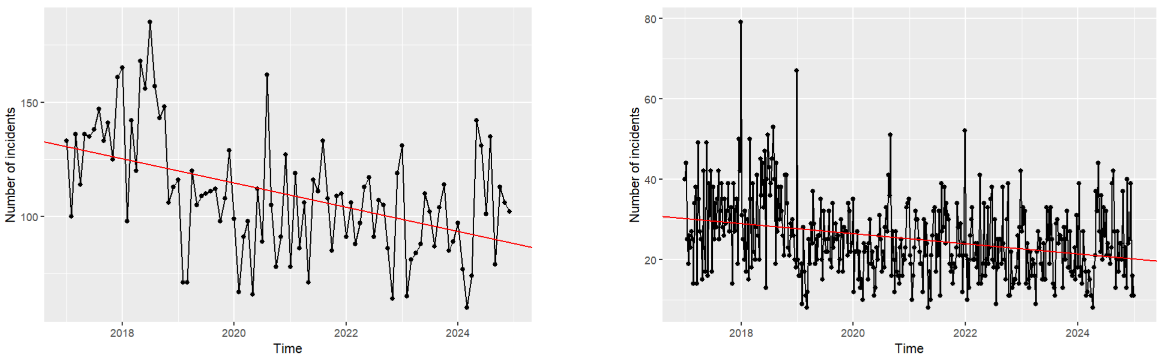



(a) Area of Aalborg defined by postal codes. (b) Incidents within the Aalborg area.

Figure 1.2

### 1.1.1 Demand Analysis

One of the overall goals for the fire department is to decrease the number of emergencies happening. At the same time, a trend over time would also be of interest in connection with possible proposals, because of the time horizon for implementing any new solutions. Therefore, the demand is plotted over time to analyze any existing trends. This is shown in Figure 1.3, where the number of incidents is summarized both monthly (a) and weekly (b).

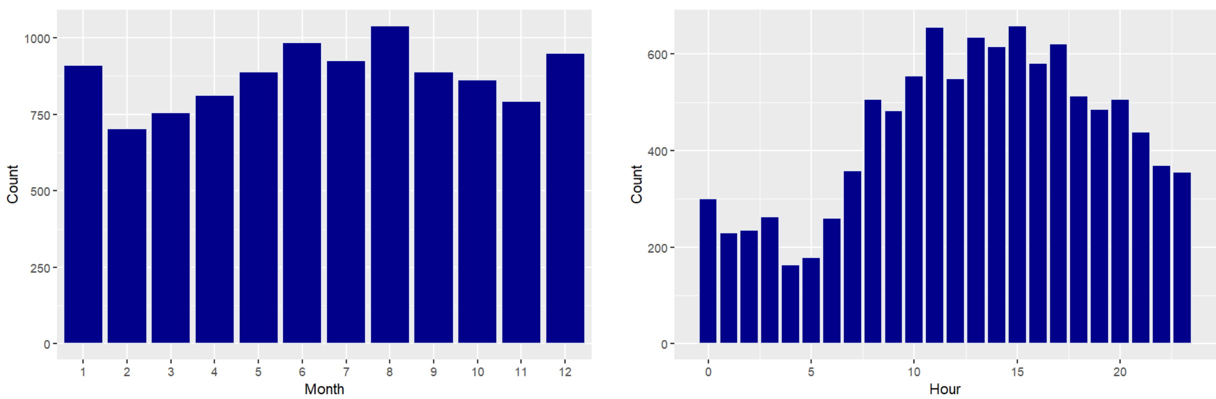



(a) Months. (b) Weeks.

Figure 1.3: Demand data divided into months and weeks with linear regression models drawn in red.

The trend is analyzed through linear regression models shown in red in Figure 1.3, and it can be seen that both monthly and weekly data indicate a decreasing trend. Model diagnostics for each of these can be seen in Appendix A.1. Considering first the monthly data, the trend coefficient is -0.44, corresponding to a yearly decrease of 5.28 incidents. The trend coefficient from the weekly model is only -0.024, corresponding to a yearly decrease of 1.25. Performing the $t$-test with a 0.05 significance level shows that both of these trends are statistically significant. Thus, it seems that the fire department's goal of lowering the number of emergencies is being fulfilled. However, from Figure 1.4(a) it appears that a higher number of emergencies were experienced in 2017 and 2018, which might be the cause of the decreasing trend. Removing these two years from the data and fitting a new linear model leads to a trend coefficient of $-0.00083$, which can be seen graphically in Appendix A.2. This $p$-value with respect to the $t$-test for this coefficient is 0.86, indicating that it is not statistically significant and can be assumed to be 0. Thus, in recent years, the demand appears relatively stable.

Another way the demand could cause problems is if it varies periodically. This is investigated below by summarizing the data by month and time of day, as seen in Figure 1.4.

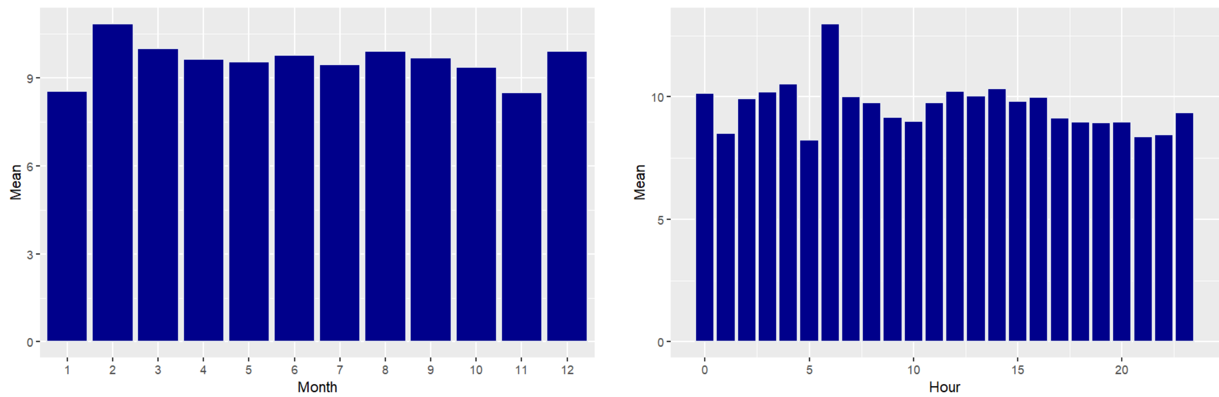



(a) Months.

(b) Hour of the day.

Figure 1.4: Number of incidents counted at different time intervals.

From the monthly summary, it seems that the demand tends to be slightly higher in the summer and winter. The varying demand is more clear in the hourly summary, which shows an increase during the normal work hours. This can be explained by the increase in activity during these hours, which in turn leads to a higher chance of accidents. These variations could be a problem if the fire station is not staffed accordingly. Such a problem would be indicated by varying response times. However, this does not seem to be the case when looking at Figure 1.5, which shows the mean response time in minutes across different periods.




(a) Months.

(b) Hour of the day.

Figure 1.5: Mean response times at different time intervals.

When fitting a linear model to response times and using the hour of the day as explanatory factors, there are only three coefficients that are significantly different from zero according to a $t$-test, and these are all between 0.8 and 1.1 minutes, which is a relatively low difference. Looking instead at the $F$-test (ANOVA), the conclusion is that statistically, the coefficient can not be assumed to be the same. This indicates that there is a difference in the response time throughout the day. However, this would also be the conclusion if only one hour had a different coefficient than all the others.

### 1.1.2 Response Time Analysis

With respect to response times, one of the emergency department's main goals is to minimize the average response time. This is important since an increase in response times leads to an increased risk for property damage and loss of lives.

One way to explore the response times is to look at their distribution. Figure 1.6 shows a histogram of the response times, where the horizontal axis is cut off at 50 minutes. This only removes the top 0.8% of the response times and is done to make the histogram more readable.

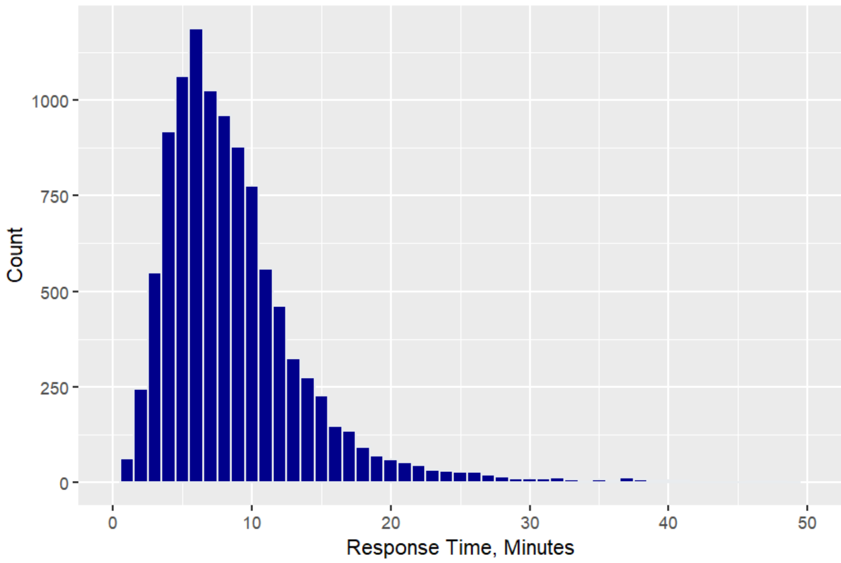



Figure 1.6: Distribution of response times in minutes.

From Figure 1.6 it is seen that most of the distribution is centered around 7-9 minutes. Indeed the mean response time over the last eight years is 9.59, with a standard deviation of 12.41. Much of the recent literature agrees upon a response time of around 5 minutes to be satisfactory, according to Aleisa [5]. Thus, compared to these standards, the response time in Aalborg is quite high. However, these standards are from all around the world, and might not apply to Aalborg or even to Denmark. Luckily, the given data set includes a specific *response time service goal* for each of the emergencies.

The service goals vary between 5 and 25 minutes, but in 71.6% of the cases it is exactly 10 minutes. Thus, for most emergencies, the department strives to be at the site within 10 minutes after receiving the emergency call. However, considering the individual cases, it appears that in 39% of the emergencies, the service goal is not satisfied, meaning they arrive later than desired. These cases could very well be the ones forming the large right tail in Figure 1.6. Thus, it appears that the department struggles to keep the response times within its own acceptable limits.

The response times can also be analyzed spatially, as in Figure 1.7, where the incidents are again plotted geographically, but now colored by response time such that a light color indicates a low response time.

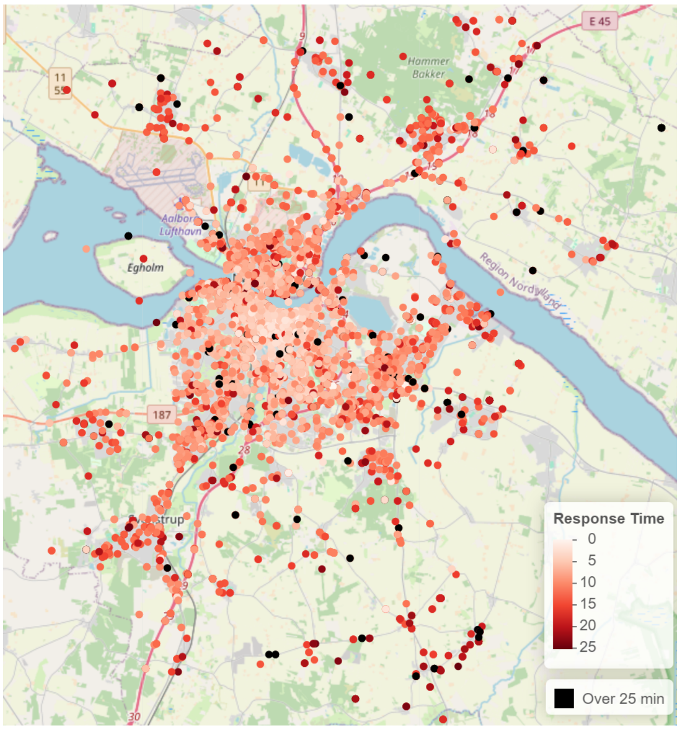



Figure 1.7: Map of the emergencies in Aalborg colored by response times.

From the figure above, it is apparent that citizens in the more rural areas tend to experience a longer response time than those living in the city center. Assumably, this is related to the longer distance to the fire station, which is analyzed with a linear regression model in Figure 1.8. The distances used in this regression are actual road distances calculated using the Routes API from the Google Cloud Service [6].

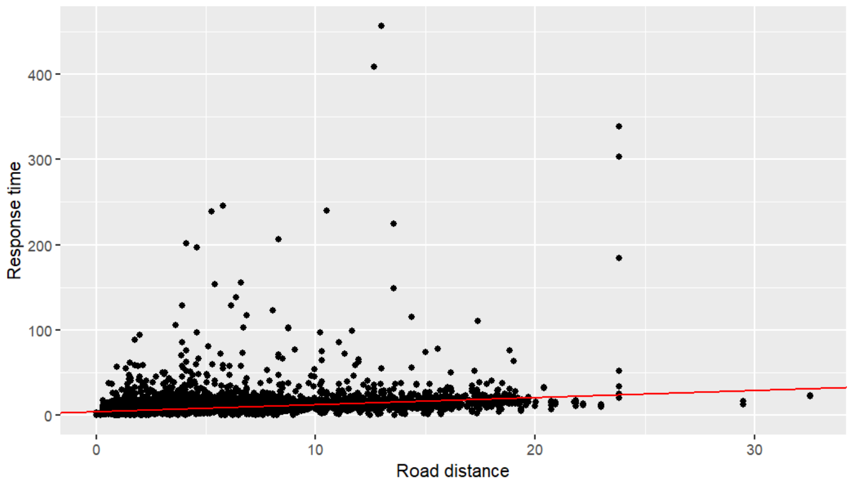



Figure 1.8: Linear regression of response times with road distance as explanatory variable.

The regression model in Figure 1.8 indicates that the response time increases by 0.83 for each kilometer with an intercept of 4.9 minutes. Again, model diagnostics for the regression model can be seen in Appendix A.1. The coefficient is statistically significant based on the $t$-test and same significance level as earlier. This indicates that if the response times are to be minimized, the location of the fire station and thereby the proximity to the emergencies have a large impact.

Furthermore, it is clear from both Figure 1.6 and Figure 1.7 that not all citizens receive the same service level with respect to response times. This is seen both in the longer right tail of the histogram, which means that some emergencies had longer response times, but also in the geographical plot, where the emergencies in the rural areas tend to have a longer response time. Especially the latter can be a problem, since it indicates that a specific group receives worse service.

Equity is not a specific goal mentioned by the emergency department, but is still an important aspect to consider, since the fire stations provide a public service [3]. At the same time, the importance of equity in both public and emergency services has received a lot of attention in recent literature. Indeed, Cepiku and Mastrodascio [7] found that the number of articles concerning equity in public services has increased considerably since 2010. Specifically in connection to emergency services, Hassler and Ceccato [8] argued that the citizens in rural areas received worse medical care that could potentially be life-threatening compared to people living closer to the hospitals, and made a case study of the impact in Scandinavia. A similar study was done in New York by Chung et al. [9], where they identified specifically vulnerable sections of the city. The main arguments used in these sources focus on the fact that the public institutions should provide a fair and even service to all citizens, while others argue that some services can have a large personal or monetary influence, which should not be affected by where you live. Other sources, such as [10] and [11], indicate that inequity might also influence the trust in public emergency service providers, at least in the USA and the UK. Thus, even though it is not a current goal for the emergency department in Northern Jutland, it is still an important aspect and could provide challenges if not considered.

A last important aspect of the response and action times is whether more incidents can happen at the same time, i.e. is it possible that another emergency call is received while the station is already handling one emergency. Figure 1.9 shows a timeline of different incidents and their action times.

From the graph, it seems that there is very little chance of overlap between different cases, since the action times are relatively short compared to the time between emergency calls. Indeed, the average action time is 0.83 hours, which only corresponds to a 7.4% chance of overlap when compared to the distribution of inter-arrival times of alarms. In that regard, multiple emergencies happening at the same time do not seem to pose a large challenge for the department.

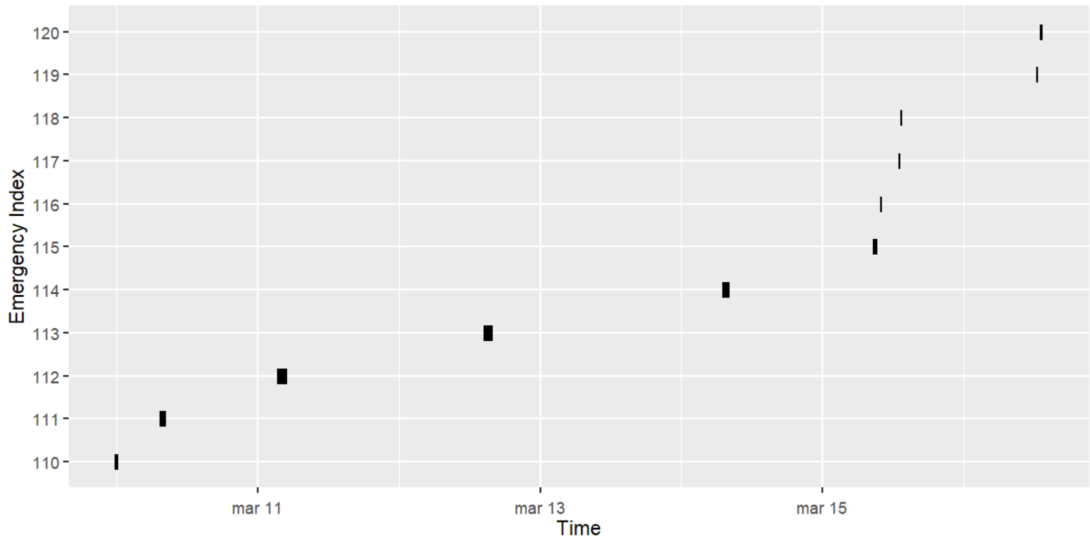



Figure 1.9: Action times shown on a timeline for a random time period.

### 1.1.3 Summary

To summarize, it appears that the problems are mostly related to the response times. In this regard, it was seen that the response times were relatively high compared to standards in the literature, while at the same time nearly 40% of the cases did not satisfy the department's service goal. Thus, there is room for improvement in this respect.

At the same time, it was seen from the distribution and spatial mapping of the response times that there is a large inequality in the service received. As mentioned, this inequity is an important aspect for public service providers and could pose challenges if ignored completely. At the same time, improving equity might also raise the overall service level and response times.

Thus, multiple challenges face the emergency department, and due to the clear connection between travel distance and response times, most of these challenges are connected to the location of the fire stations. This corresponds well with the department's own interest in placing another fire station in Aalborg, and the problems might therefore be formulated as a facility location problem. However, a challenge arises in how to properly formulate and include these objectives. Especially with respect to performance, i.e. the achieved response times, since multiple aspects pose a problem. Should the focus be on minimizing the average response times or increasing the number of satisfied service goals? Or should both be included simultaneously?

## 1.2 Literature Review

In the current literature, multiple ways of defining facility location problems (FLP) are used, each including and modeling different aspects. Overall, the approach can be divided into *static* and *dynamic* models [12]. The former is the original case where no changes happen in the setting over time, whereas the dynamic models are a later edition that allow for changing variables over time. The dynamic model has also been applied specifically for the fire station location problem by Hajipour et al. [13]. However, based on the former data analysis, where the demand and response times were seen to be relatively stable over time, the static models will be the focus for this project and literature review.

The literature can further be divided into *stochastic* and *deterministic* approaches. The former focuses on including uncertainty in the form of stochastic elements, while the latter assumes deterministic variables. Ming et al. [14] incorporates the stochastic approach for the fire station location problem and maximizes expected coverage. However, the previous analysis of the data did not show large variations and indicated that the response times were closely related to the distance. At the same time, the possibility of multiple emergencies happening at the same time was also very small. Based on these results, deterministic approaches will be considered in this project.

Another important aspect of a facility location problem is the objective used for optimization. Again, the literature proposes a wide variety of options. One class of objectives focuses on optimizing performance or efficiency, and can in general be divided into two groups: *median problems* and *covering problems*. The median problems seek to minimize the overall cost or distance to the facility, and were first introduced by Hakimi [15] as the $p$-median problem. More recent literature still uses this formulation, where Alp et al. [16] and Resende and Werneck [17] propose heuristic and meta-heuristic algorithms to solve the well-known problem.

The covering problems can further be divided into two groups. The *Location Set Covering Problem* (LSCP) was first proposed by Toregas [18] in 1970, while the *Maximal Covering Location Problem* (MCLP) was introduced in 1974 by Church and ReVelle [19]. Both of these assume a specified coverage range for the facility, but the former focuses on minimizing the number of facilities that cover the entire area of interest, while the latter seeks to optimize the covered area with a certain number of facilities. Since the emergency department of Northern Jutland does not have unlimited funds and thereby restrictions on the number of facilities, this project and literature review will focus on the MCLP.

Concerning the fire station location problem, primarily, the coverage formulation has been used in recent literature. Both Aktas et al. [20] and Idayani et al. [21] formulate an MCLP and use it in a practical application to optimize fire station locations in Istanbul and Indonesia, respectively. Motivated by a more realistic coverage Murray [22] proposes gradually decreasing coverage instead of binary coverage for fire stations. This is a version of the *generalized* MCLP also used by Berman and Krass [23] and Jabalameli et al. [24], where the latter used it to allocate cell phone towers. Others allow demand points to be covered by more than one location, either in the form of backup coverage like Karatas et al. [25] and Hogan and Revelle [26] or as cooperative coverage like Jabalameli et al. [24].

Another class of objectives concerns equity. Marsh and Schilling [27] provides a review of the most common equity measures used in facility location problems and proposes a common framework and notation. One of the most common equity models is the *p-center* problem, also called the *minimax* problem. Drezner [28] proposes different solution methods, both heuristic and optimal, for the problem, while Abubakar et al. [29] used the model formulation to optimize fire station locations in Sokoto. Some authors like Drezner et al. [30] borrow equity measures from other fields, such as the Gini-coefficient, when optimizing facility locations, while some borrow concepts from the $p$-median problem in the so-called *fringe sensitive location problem* (FSLP), like Gunnarsson et al. [31] and Untadi et al. [32].

Combining different objectives into a multi-objective model has also received a lot of attention in recent literature. In a fire station location setting, Badri et al. [33] introduced 11 different objectives and combined these in different ways in a practical application in Dubai. They found that the choice of objective had a great effect on the optimal solution. More recently, Yao et al. [34] proposed a bi-objective model for the fire station location problem that combines the objectives from the LSCP and p-median models, whereas Yu et al. [35] used an MCLP combined with workload objectives for locating fire stations in Nanjing, China. With respect to including both efficiency and equity, the literature is less common. Moche [36] explored a combination of the $p$-centre and $p$-median, though in a medical setting, while Untadi et al. [32] includes a combination of MCLP and FSLP for a fire station location problem, but focuses on incorporating socio-economic factors to weigh the demand points.

Many different solution approaches to the facility location problem have been proposed in the literature, and where most rely on linear programming software and branch-and-bound algorithms [20, 22, 23, 33, 37], some propose heuristic approaches. Indeed, Macit [38] uses a genetic algorithm to solve an MCLP for locating fire stations in Turkey, and Xia et al. [39] compares five different heuristics for the MCLP in general. The latter found that simulated annealing performed the best. However, these algorithms are not applicable in a multi-objective setting. In this case, Yao et al. [34] use the $\varepsilon$-method to find a set of Pareto-optimal solutions, while Yu et al. [35] combines the objectives into a single linear combination. Otherwise, heuristic and meta-heuristic approaches have also been applied to multi-objective optimization, where López Jaimes et al. [40] reviews a wide range of algorithms from the general literature, with a focus on evolutionary algorithms. These are used in the facility location literature as well, where Wang and Chen [41] uses the NSGA-II algorithm to solve a multi-objective facility location problem. Other heuristic methods include the multi-objective particle swarm algorithm, as used by Hu et al. [42], or problem-specific algorithms such as the *scatter search* proposed by Lopez et al. [43]. Further review of solution methods for multi-objective optimization problems will be included later in the project. A summary of the relevant literature is seen in Table 1.1 below.

| Source | Model formulation | | | | | Coverage type | | | | Solution Method |
|---|---|---|---|---|---|---|---|---|---|---|
| | $p$-median | $p$-center | MCLP | LSCP | FSLP | Binary | Gradual | Cooperative | Backup | |
| Hajipour et al. [13] | | | X | | | X | | | | Particle swarm optimization & Artificial Bee Colony |
| Alp et al. [16] | X | | | | | | | | | Genetic Algorithm |
| Resende and Werneck [17] | X | | | | | | | | | Hybrid Heuristic |
| Toregas [18] | | | | X | | | | | | |
| Church and ReVelle [19] | | | X | | | X | | | | Greedy Adding heuristic & LP Software |
| Aktas et al. [20] | | | X | X | | X | | | | LP Software |
| Idayani et al. [21] | | | X | | | X | | | | LP Software |
| Murray [22] | | | X | X | | X | X | | | LP Software |
| Berman and Krass [23] | | | X | | | | X | | | LP Software & Greedy Heuristic |
| Jabalameli et al. [24] | | | X | | | | X | X | | Simulated Annealing |
| Karatas et al. [25] | X | | X | | | | | | X | LP Software |
| Hogan and Revelle [26] | | | (X) | (X) | | | | | X | LP Software |
| Drezner [28] | | X | | | | | | | | Heuristic |
| Abubakar et al. [29] | | X | | | | | | | | |
| Gunnarsson et al. [31] | | | X | | (X) | X | | | | LP Software |
| Untadi et al. [32] | | | X | X | | X | | | | Random Search Algorithm |
| Badri et al. [33] | | | | X | | X | | | | LP Software |
| Yao et al. [34] | (X) | | | (X) | | X | | | | $\varepsilon$-constraint method & LP Software |
| Yu et al. [35] | | | (X) | | | X | | | | LP Software |
| Moche [36] | (X) | (X) | | | | | | | | LP Software |
| Gendreau et al. [37] | | | X | | | | | | | LP Software |
| Macit [38] | | | X | | | X | | | | Genetic Algorithm |
| Xia et al. [39] | | | X | | | X | | | | Multiple heuristics & meta-heuristics |
| Wang and Chen [41] | | | (X) | | | X | X | X | | NSGA-II |
| Hu et al. [42] | | | | | | | | | | MO Particle Swarm Optimization |
| Lopez et al. [43] | (X) | (X) | (X) | | | | | | | Scatter Search |

Table 1.1: Summary of FLP-related literature, classified by model formulation with coverage type and solution method specified where available. Multi-objective formulations are marked with (X).

## 1.3 Problem Statement

As summarized in the problem identification, multiple challenges face the emergency department in Northern Jutland. Motivated by the literature review, these problems can be incorporated in a static and deterministic facility location problem, and it seems beneficial to include multiple objectives to represent the different aspects of the problem. Based on the results of the problem identification, the challenges are related to both efficiency and equity of the response times, and it would therefore be sensible to include objectives representing both aspects.

With respect to equity, it was seen in Section 1.1 that some groups, primarily citizens in the rural areas, received worse service than others. Thus, equity should be included and optimized in the model. The $p$-center formulation seeks to minimize the maximal distance and thereby improve the bottom line for the performance. Optimizing this objective would improve the service in the worst cases, and thereby improve the overall equity. At the same time, the $p$-center objective also benefits from an easy interpretation and might therefore make the model more accessible to the decision makers at the emergency department.

In terms of efficiency, it was stated in Section 1.1, that the fire department was challenged by high response times as well as a high failure rate with respect to the service goal. Specifically, 39% of the emergency responses did not satisfy the service goal. In order to improve this performance, coverage should be included as an objective for the model as well. As motivated in the literature review, the MCLP formulation is the best representation of the current situation and will therefore be used to represent the coverage. It is noteworthy that this is not completely in line with the department's goal of minimizing response times, since coverage is not an exact representation of response times. However, they are still connected, and optimizing coverage might still improve the overall response time performance.

The MCLP formulation uses binary coverage, but as mentioned in the literature review, gradual coverage might be more representative in a fire station setting, since it does not ignore the uncovered demand points. However, the inclusion of the $p$-center objective means that these demand points are not disregarded entirely and thereby still affect the solutions. Therefore, binary coverage will be assumed for this project.

To the best of my knowledge, no previous research has combined the objectives from the MCLP and $p$-center models. At the same time, the inclusion and optimization of both efficiency and equity objectives remain unexplored in a fire station location setting. Thus, this project tries to fill these gaps in the literature, while giving meaningful results and suggestions to the emergency department of Northern Jutland. This is done by answering the following problem statement:

> *How can the fire station location problem in Aalborg be formulated to address efficiency along with equity, and solved using multi-objective optimization techniques?*

# 2 | Mathematical Problem Formulation

The problem described in the previous chapter will be formulated as a *multi-objective facility location problem* in order specify the different aspects and enable solution methods for this kind of problem. Such a problem seeks to allocate a number of facilities in order to optimize certain objectives. The first objective that will be included in this project is from the *Maximal Coverage Location Problem*, and seeks to maximize the number of demand point, that are covered by a pre-determined number of stations with a specific coverage radius. A demand point can represent an area or specific location, which is in need of service from the facility, and it is only covered if it is located within the coverage radius of a facility. The second objective used in this project is from the $p$-center formulation, which also uses the demand points, but seeks to minimize the maximal distance to a facility across all of the demand points. Of course, the facilities in this project are the fire stations and in order to define the problem the following assumptions and constraints are assumed to hold:

- There is a finite number of demand points and possible locations for the fire stations.
- The demand and response times are static and deterministic, meaning that they are constant over time and not subject to any uncertainties.
- No more than one emergency can happen at a time, and hence the fire stations are always ready to respond.
- The demand points are assumed to require the same type of service, meaning that need for specialized equipment and staffing is ignored.
- In this regard, the stations are assumed to be homogeneous.

Such a problem can be stated as a mixed integer linear programming problem and to do this, let $\mathcal{I}$ be the set of demand points in need of coverage. Even though it was assumed that the demand points require the same type of service, it does not mean that they are identical. Some might be more prone to emergencies and thereby require more frequent service. To include this aspect, let $a_i$ represent the service load for demand point $i \in \mathcal{I}$.

Now, let $\mathcal{J}$ be the set of possible locations for the fire stations, and since the fire stations are homogeneous, let $S$ be the common coverage radius of the fire stations. As mentioned, a demand point is only covered if it is within the coverage radius of a fire station. Thus, for $i \in \mathcal{I}$ and $j \in \mathcal{J}$ let $d_{ij}$ be the distance between demand $i$ and station $j$ and let $N_i = \{j | d_{ij} \leq S\}$ represent the set of possible facility locations, which are able to cover demand point $i$. In this connection, define the following decision variables:

$$x_j = \begin{cases} 1 & \text{if a station is placed at location } j, \\ 0 & \text{otherwise.} \end{cases}$$

$$y_i = \begin{cases} 1 & \text{if demand point } i \text{ is covered,} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, since there are not unlimited funds to open new facilities, let $p$ be the desired number of fire stations to allocate. Some fire stations are already in use, and it might make sense to keep them in operation. In this regard, let $\Theta$ be the set of locations from $J$ containing an already existing fire station, and let $q$ be the number of existing fire stations to keep in operation.

Finally, in order to determine the distances related to the $p$-center objective, each demand point needs to be assigned to a specific fire station. Therefore, the following decision variables are needed as well:

$$z_{ij} = \begin{cases} 1 & \text{if demand point } i \text{ is assigned to station } j, \\ 0 & \text{otherwise.} \end{cases}$$

$$D = \text{Maximal distance from any demand point to its assigned fire station.}$$

The multi-objective facility location problem can now be formulated as follows:

$$\text{maximize} \sum_{i \in \mathcal{I}} a_i y_i, \text{ and} \tag{2.1}$$

$$\text{minimize } D, \tag{2.2}$$

Subject to

$$\sum_{j \in \mathcal{J}} d_{ij} z_{ij} \leq D \quad \forall i \in \mathcal{I}, \tag{2.3}$$

$$\sum_{j \in \mathcal{J}} z_{ij} = 1 \quad \forall i \in \mathcal{I}, \tag{2.4}$$

$$z_{ij} \leq x_j \quad \forall i \in \mathcal{I},\ j \in \mathcal{J} \tag{2.5}$$

$$y_i \leq \sum_{j \in N_i} z_{ij} \quad \forall i \in \mathcal{I} \tag{2.6}$$

$$\sum_{j \in \mathcal{J}} x_j = p, \tag{2.7}$$

$$\sum_{j \in \Phi} x_j = q, \tag{2.8}$$

$$x_j \in \{0,1\} \forall j \in \mathcal{J}, \quad y_i \in \{0,1\} \forall i \in \mathcal{I}, \quad z_{ij} \in \{0,1\} \forall i \in \mathcal{I},\ j \in \mathcal{J}, \quad \text{and} \quad D \in \mathbb{R}. \tag{2.9}$$

The objective in (2.1) seeks to maximize the weighted sum of covered demand points, as in standard MCLP, while (2.2) seeks to minimize the maximal distance to a fire station among all demand points, as in the $p$-center formulation. The latter equation only works in combination with (2.3), which ensures that $D$ is larger than all distances from demand points to their assigned fire station. At the same time, (2.4) and (2.5) ensure that all demand points are assigned to exactly one open fire station. Constraint (2.6) states that a demand point is only covered if it is assigned to an open fire station within range. The constraints in (2.7) and (2.8) make sure that the desired number of fire stations is opened, while keeping the correct number of existing ones. Finally, (2.9) imposes domain restrictions on the decision variables.

# 3 Theory on Multi-Objective Optimization

When considering single-objective optimization problems, it is possible to determine a single optimal solution with respect to the objective. However, when more objectives are added, this is no longer the case. Now, optimizing one objective might lead to a poorer performance for another objective, and so a balanced trade-off between the objectives might be the best solution. In order to handle and optimize multiple objectives, specific theory and solution methods are needed. This theory and related concepts are introduced in this chapter, which is based on [40], [44], and [45].

## 3.1 Pareto Optimality

When comparing solutions with respect to multiple objectives, *Pareto dominance* is often used and it is defined in Definition 3.1.1 below for a minimization problem. The definition holds in general, since possible maximization objectives can be transformed to minimization objectives by negating the values. This definition uses the *objective space* $\mathcal{F}$, which is a solution space, where the solutions are represented by their achieved objective values, as opposed to the *variable space* $\mathcal{X}$, where they are represented by the decision variables.

**Definition 3.1.1 (Pareto Dominance).**
Let $y = \{y_1, \dots, y_n\}$ and $z = \{z_1, \dots, z_n\}$, for $n \in \mathbb{N}$, be solutions in the objective space $\mathcal{F}$ of a minimization problem. Then $y$ is said to *Pareto dominate* $z$, denoted as $y \prec z$, if and only if

$$y_i \leq z_i \ \ \forall i \in \{1, \dots, n\} \quad \text{and} \quad \exists j \in \{1, \dots, n\} : y_j < z_j.$$

Definition 3.1.1 states that one solution Pareto dominates another if all objective values are at least as good, with one being strictly better. The concept of Pareto dominance is used to define *Pareto optimality*.

**Definition 3.1.2 (Pareto Optimality).**
A solution $y$ is said to be *Pareto optimal* if there does not exist another solution $z$, such that $z \prec y$.

Intuitively, Definition 3.1.2 states that a solution is Pareto optimal if one objective cannot be improved without worsening at least one of the others. When optimizing a multi-objective problem, one is commonly interested in finding the set of Pareto optimal solutions, called the *Pareto front* or *frontier*. The decision maker will then be able to choose a preferred solution from the set. This approach is referred to as *a posteriori*, as opposed to *a priori*, where the decision maker provides information before the optimization process. Since no preferences or information about the objectives used in the project were available beforehand, a posteriori methods will be the focus of the following sections.

## 3.2 Optimization Techniques

Many common techniques for finding the Pareto front rely on the mathematical formulation like the one in Chapter 2. This includes the *Linear Weighting* method, which combines the multiple objectives $f_1, \ldots, f_n$ into a single objective, by making a linear combination with weights $w_1, \ldots, w_n$, i.e.

$$\text{minimize} \sum_{i=1}^{n} w_i f_i(x), \tag{3.1}$$

$$\text{subject to } x \in \mathcal{X}. \tag{3.2}$$

A solution found by this method with specified weights is always going to be Pareto optimal, cf. Emmerich and Deutz [45]. This means that a Pareto front can be found by solving the problem repeatedly with varying weights. This method could be used as an a priori approach if the weights were given by the decision maker beforehand.

Another common approach is the so-called *ε-constraint* method. This method works by optimizing one objective at a time, while using the others as constraints bounded by an $\varepsilon$-value. Thus, the optimization problem is formulated as

$$\begin{aligned} \text{minimize} \quad & f_l(x) \\ \text{subject to} \quad & f_i(x) \leq \varepsilon_i \quad \forall i \neq l, \\ & x \in \mathcal{X}. \end{aligned}$$

This problem is a single-objective problem and can be solved as such. An optimal solution to the problem, where all of the $\varepsilon$-constraints are satisfied to equality, is Pareto optimal cf. Miettinen [44]. A Pareto front can then be found by varying the $\varepsilon$-values.

Despite their common use, both of these methods, along with similar methods used in the literature, have some major drawbacks, as pointed out by López Jaimes et al. [40]. First is the fact that the algorithms have to be run repeatedly to find the Pareto front, which, of course, increases the computation time. At the same time, some approaches require knowledge of the domain in order to solve the problem. This is the case for the $\varepsilon$-constraint method, where bounds for the objective values are needed to determine the possible ranges for the $\varepsilon$-values. To address these problems, a variety of alternative approaches have been proposed in the literature, including both heuristic and meta-heuristic algorithms, see [40]. Such algorithms often benefit from not requiring any prior knowledge of the solution space, while only needing to be run once to find the result. On the other hand, they are not guaranteed to produce the true Pareto front, but merely an approximation.

### 3.2.1 NSGA-II

Among the wide palette of heuristic and meta-heuristic algorithms, the *Non-Dominating Sorting Genetic Algorithm* (NSGA-II) is chosen for this project, motivated by the successful use of evolutionary algorithms in the single-objective literature and use of the algorithm itself in multi-objective facility location problems. Evolutionary or genetic algorithms are based on evolution, where a population of solutions, called chromosomes, is combined and mutated repeatedly over generations, while only the best-ranked solutions are kept across time, i.e. “survival of the fittest”. NSGA-II is such an algorithm, which has been designed specifically for multi-objective optimization. The general procedure of NSGA-II can be seen in the flowchart in Figure 3.1 and follows that of a standard genetic algorithm in a single-objective setting. Thus, detailed explanations of the fundamentals of the algorithm i.e. the generational loop, the crossover and mutation operators, and the termination criteria, will not be included in this project, but can be found in [46]. Instead, the following will focus on the incorporation of multiple objectives.

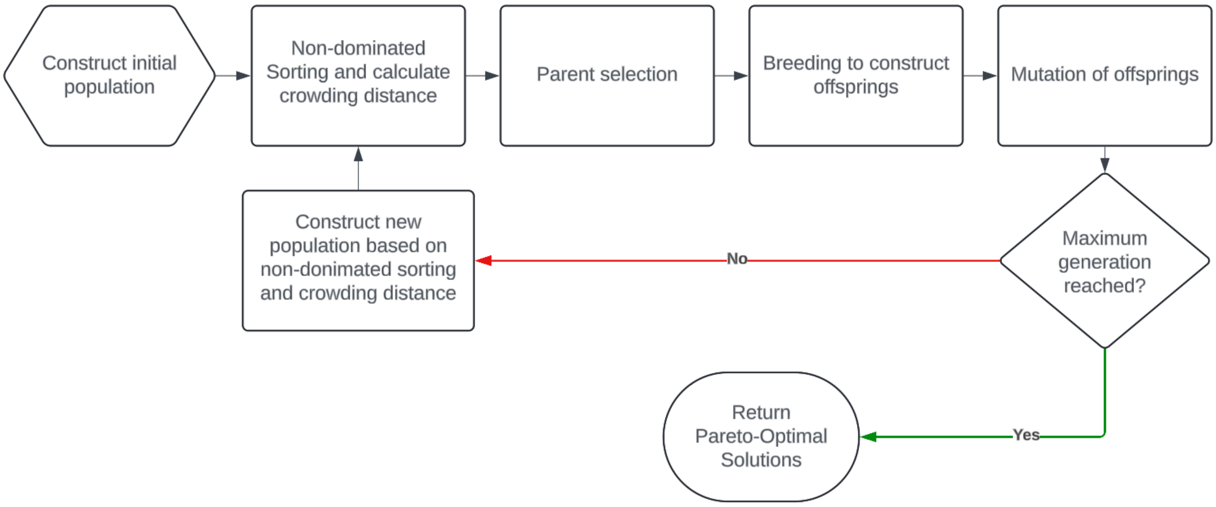



Figure 3.1: Flow chart of the NSGA-II algorithm.

The main way in which NSGA-II distinguishes itself from the normal single-objective genetic algorithms is in the evaluation and comparison of objective values. The NSGA-II is focused on finding Pareto fronts and, as such, considers all objectives simultaneously. This is done in the *non-dominated sorting* module of Figure 3.1. In this module, the population is sorted into fronts $\mathcal{F}_1, \ldots, \mathcal{F}_k$, with the same number of Pareto-dominating solutions. This is done by first finding the solutions that are not Pareto-dominated by any other solutions and adding them to $\mathcal{F}_1$. The next front is then found by ignoring the solutions from $\mathcal{F}_1$ and finding the new non-dominated solutions. This procedure repeats until all solutions are assigned to a front.

Within each front, the chromosomes are ranked based on *crowding distance*. The crowding distance for solution $i$ is defined as

$$c_i = \sum_{j=1}^{n} \frac{f_j^+(i) - f_j^-(i)}{f_j^{\max} - f_j^{\min}}, \tag{3.3}$$

where $f_j$ represent each of the $n$ objective functions. For solution $i$, $f_j^+(i)$ and $f_j^-(i)$ represent the objective values of the closest solutions, within the front, with respect to objective $j$. In the same way, $f_j^{\max}$ and $f_j^{\min}$ represent the maximum and minimum of the objective values achieved within the front. Intuitively, the formula in (3.3) can be explained as follows: If there are other solutions close to $i$, then the numerator will be small and thus decreasing the crowding distance, whereas if solution $i$ is in a less populated area, then the numerator and the crowding distance increases. A large crowding distance is preferable, since it promotes more exploration of the solution space.

The non-dominated sorting and crowding distance is primarily used in the construction of the new generations. In this case, both the parent population and the child population are combined into one and sorted into fronts. The best fronts are added to the new population one by one until the desired population size is reached. If the last added front is too large and adding it would exceed the population size limit, then the crowding distance is used to select the best solutions, such that the correct population size is reached. This procedure is explained graphically in Figure 3.2.

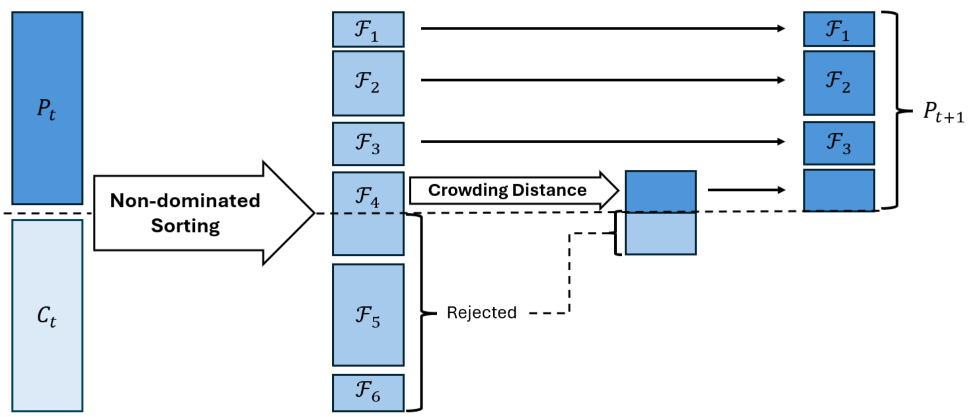



Figure 3.2: Graphical explanation of how the new population, $P_{t+1}$, is constructed from current parent population, $P_t$, and child population $C_t$.

After the new generation has been constructed, the next iteration of the generational loop can begin. When the pre-determined maximal number of generations has been reached, the Pareto optimal solutions in the final populations are returned. To evaluate the performance of the algorithm, the approximated Pareto front can be compared to the true Pareto front, if this is known. If this is not the case, other measurements can be used to compare Pareto fronts, which is the focus of the following section.

## 3.3 Quality Indicators

In a single objective setting, different solutions can easily be compared based on the objective value. However, in a multi-objective setting, the result is not a single solution, but an entire Pareto front. In order to compare different fronts, so-called *Quality Indicators* (QI) are used instead, see [47] and [48]. Different QIs measure different aspects of a front, where mainly four aspects or qualities are of interest. *Convergence* describes the distance to the true Pareto front, where closeness is preferred. *Spread* is related to the area covered by the front, while *Uniformity* measures how evenly the solutions are distributed. A wide and evenly distributed front is preferable. The last quality is *Cardinality*, and it is concerned with the number of solutions on the Pareto front, where a larger number is generally preferred.

A large set of quality indicators exists in the literature, and many of them are concerned with a single aspect. For example, convergence of fronts is often measured through dominance-based QIs, where two fronts are compared based on Pareto dominance, or distance-based QIs, where the solutions are compared to a pre-known element, like the ideal point or a reference set. Spread and uniformity might be investigated by measuring the distance between neighboring solutions, while cardinality can be measured simply by counting the number of solutions. However, in this project, the QI known as *hypervolume* is used, due to its common use in literature and the fact that it measures both convergence, spread, and cardinality. Another benefit of this quality indicator is its easy interpretation in the sense that it assigns a real value to each front, and a larger value indicates a higher quality. Thus, it can be used to compare multiple fronts as opposed to only pairwise comparisons, which is the case for other QIs. At the same time, it does not require any knowledge of the solution space, which is also the case for some methods.

For a solution front $A$, the hypervolume is defined as

$$HV(A) = \lambda \left( \bigcup_{a \in A} (x | a \prec x \prec r) \right), \tag{3.4}$$

where $\lambda$ is the Lebesgue measure and $r$ is a reference point. In a discrete two-dimensional setting, the hypervolume of a Pareto front is the area of the union of rectangles with corners at the reference point and each of the solutions on the Pareto front. An example can be seen in Figure 3.3, where both objectives have to be minimized. With this in mind, the reference point is often chosen as 1.1 times the nadir point, i.e. the combination of the worst observed values across the front, to ensure consistency in the "direction" of the rectangles.

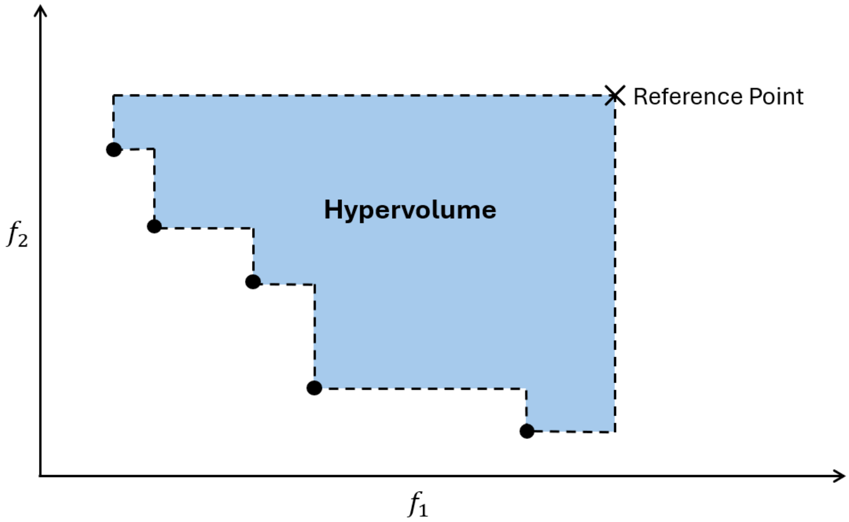



Figure 3.3: Illustration of the hypervolume of a Pareto front.

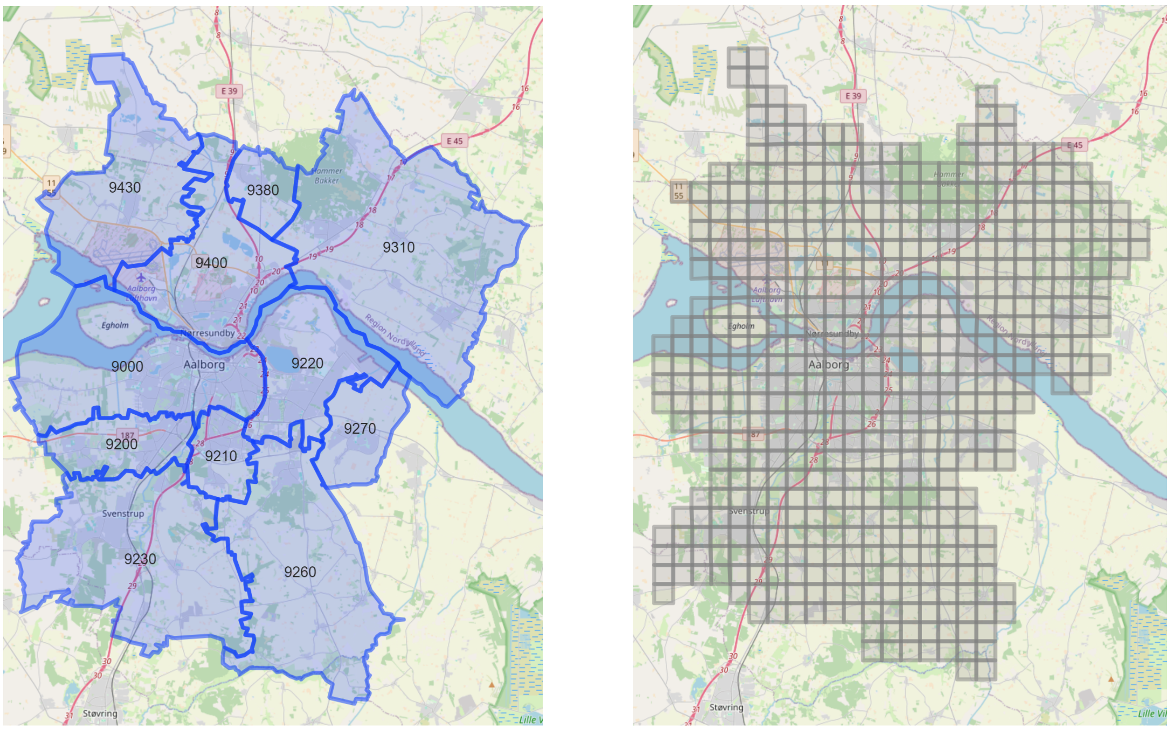
It is easy to see from Figure 3.3, that an improvement in either of the three qualities will lead to an increase in the hypervolume. The convergence might be improved by moving one of the solutions closer to the origin, which would increase the size of the area. Spread and cardinality could both be improved by adding a solution to the Pareto front, which would also add to the blue area and thereby increase the volume. A last important notion is that in order to get a meaningful comparison, a common reference point should be used for all fronts.

# 4 | Solution Approach

It was mentioned in the literature review that multiple ways of solving multi-objective optimization problems exist, and that some are better than others depending on the scenario. The purpose of this chapter is to motivate and explain the approach used in this project. To do so, the data is first processed in order to define the necessary parameters explained in Chapter 2 and to determine the size of the problem.

## 4.1 Data Preparation

The area of interest has already been defined in Figure 1.2a, but to solve the problem in Chapter 2, a set of demand points is needed. This is obtained by using data from [49] to divide the area into $1 \times 1$km squares, forming a grid. The grid contains 523 squares and is seen in Figure 4.1, along with the map of the postal codes for easier comparison.

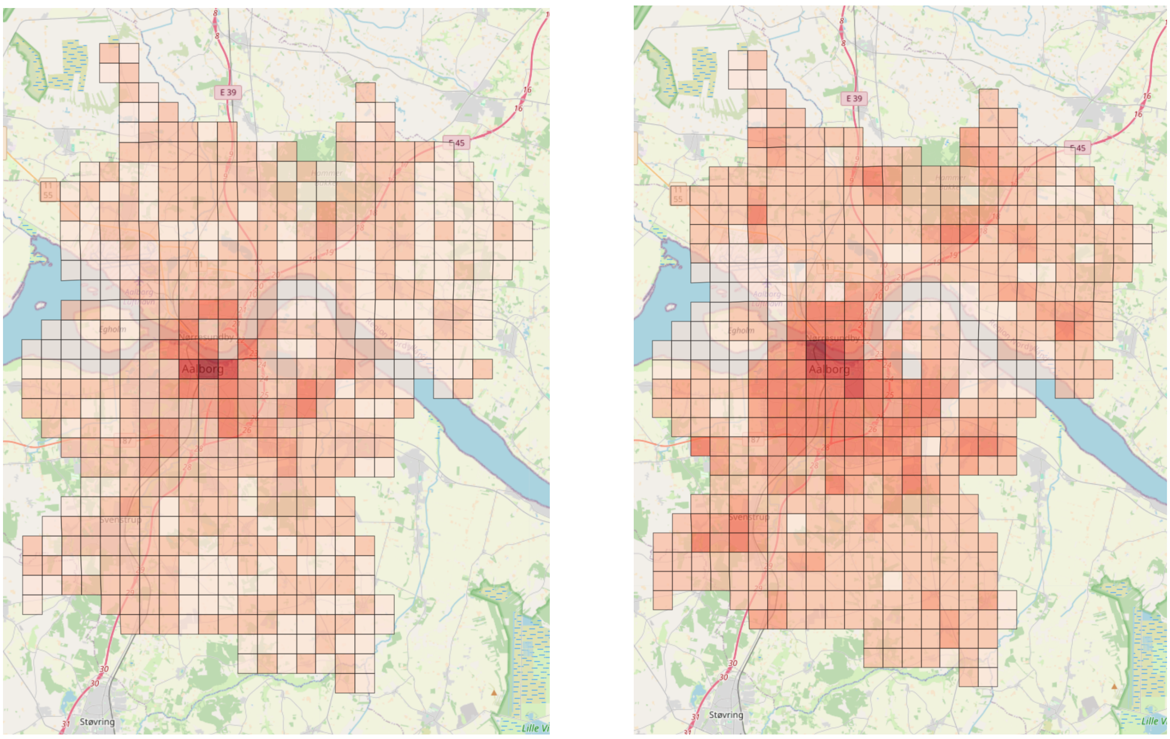



(a) Area of Aalborg defined by postal codes. (b) Grid with 523 cells.

Figure 4.1

The centroid in each of these squares will be used as a demand point representing the entire square. All of these squares are not of equal importance, since some might be uninhabited while others are densely populated and thereby more prone to accidents. Thus, each of them is assigned a weight, also called a service load. The literature presents different possibilities for assigning these values, such as the population in each area [19, 50] or the number of incidents [33], while some use a combination of the two [31, 34]. Untadi et al. [32] even include socio-economic factors in a linear combination to weigh each demand point. Due to the availability of data and the wide usage in the literature, the service loads will be determined as a linear combination of the number of incidents and population size in each square. Both of these factors will be normalised to ensure equal weight of the two.

However, the construction of a new fire station is not done immediately, and it has been explained by the fire department that it can take 5-8 years before the new station is operational. To take this into account, historical population data from [49], dating back to 2006, is used to predict the population in each grid. This is done with a linear model for 2030, and these new population sizes are then used for the service load calculation. It should be mentioned that only five previous data points exist for each square and might not result in the most accurate and representative models, though they might still capture some of the trends. The number of incidents and the predicted population can be seen in Figure 4.2 (a) and (b), respectively.




(a) Incidents. (b) Population.

Figure 4.2: Grid of Aalborg coloured by different variables.

From Figure 4.2 it can be seen that the two factors seem to correlate, since the areas with high population density also experience more emergencies. It is also seen that some of the squares, colored white, have a value of zero in both factors. This is a problem since a service load of zero might lead to a demand point not being counted as covered even though it is within the coverage radius of an open facility, since it would not do anything to increase the objective value. Thus, to ensure that no demand point has a service load of zero, the minimum observed value among the two factors is added to all points, hence keeping the ranking of the squares.

Another important part of the problem formulated in Chapter 2 is the possible locations for new fire stations. Since no information about potential locations was given, all of the squares are considered possible locations, expect the ones only containing water. Additionally, the location of the current fire station is also seen as a possibility, which results in a total of 504 possible locations.

A last variable mentioned in the mathematical formulation is the distance between the fire station and the demand point. Ideally, it would be calculated along the roads as done for the model in Figure 1.8. However, these were calculated using the Google Cloud service, which has limited free usage. Therefore, it was not possible to calculate all of the distances between possible locations and demand points in this way. Instead, the Manhattan distance is used as an approximation. When calculating the Manhattan distance to each of the incidents from the data and comparing them to the road distances used in Figure 1.8, a mean absolute error of 1.82 km is obtained. Thus, when using these as an approximation, a small bias is introduced in the model. To measure coverage, these distances are compared with the coverage radius, which will be calculated using the mean travel speed across all incidents from the last eight years, which is 43.8 km/h or 0.73 km/m. This speed is then compared to the service goal of 10 minutes, leading to a coverage radius of 7.3 km. It might be argued that this speed seems low. However, it is the average speed that has to take into account acceleration, turning, and possible traffic in the city center. At the same time, using a low travel speed optimizes the worst-case performance, which is of more interest considering the uncertainty in travel speed.

## 4.2 Solution Method

As explained in Chapter 3, common mathematical programming approaches exist for solving multi-objective optimization problems. One of these, specifically the $\varepsilon$-constraint method, was initially tested on a small-scale problem with 50 possible locations. This was done with the `ompr`-package in `R` [51], which uses the *branch and bound* method. However, this resulted in a computation time over an hour. Based on the drawbacks mentioned in Section 3.2 and the long computation time, NSGA-II will be used in this project instead.

An important part of implementing NSGA-II is determining how to represent each chromosome. Here it is done as a binary vector, where each entry represents a potential location for a fire station. Having the value 1 in one of the entries therefore means that a station is opened at the corresponding location, while a value of 0 means that no station is opened at the location. The amount of ones in the chromosomes will therefore be equal to the number of fire stations to be located, $p$. With this representation, the chromosomes in the initial population are generated by randomly placing $p$ ones in a vector of zeroes. After this population has been created, the generational loop is initiated, which is also shown in the flowchart of the algorithm in Figure 3.1.

In terms of parent selection, this project uses *binary tournament selection*, where each parent is chosen as the best among two randomly selected chromosomes based on non-dominated sorting and crowding distance. This promotes elitism and more exploitation of good solutions. These parents are then used to create a child by the simple order crossover [52]. Each child then has a possibility of being mutated, which in this project is done by swapping two values in the chromosome vector. If either of these operations results in a change of the number of fire stations, a repair function is applied that either adds or removes a random fire station, depending on the situation.

## 4.3 Tuning of parameters

When implementing the NSGA-II algorithm, three parameters are chosen: the population size, the number of generations, and the mutation rate. The tuning of these parameters will be based on the hypervolume introduced in Section 3.3, while the possible ranges for each of them are based on initial testing and common practice in the literature [41, 53]. To find the common reference point, a grid search was first made across all possible combinations of the three parameters, and the worst observed objective values were combined into an overall nadir point, which was then multiplied by 1.1. Afterwards, each parameter was tuned individually by running the algorithm repeatedly and only changing one of the parameters, while keeping the others fixed. Since the algorithm includes random elements, the results were made more robust by repeating the process multiple times for each parameter and possible value, and taking the mean hypervolume across all repetitions. Different numbers of repetitions were tested, but no great changes in the mean hypervolume were seen beyond 25, as can be seen in Appendix B.1. Hence, 25 repetitions were then chosen to keep the computation time low. The result of the tuning can be seen in Figure 4.3 below.

From Figure 4.3 it can be seen that the general tendency is that a higher parameter value leads to better performance. This can be explained by the fact that an increase in any of the parameters leads to more exploration of the solution space, thereby increasing the possibility of finding more optimal solutions. However, not all increases are significant enough with respect to the added computation time, and the “elbow method” is therefore used to choose the actual parameter values, which will be used to create the main results. These are summarized in Table 4.1.

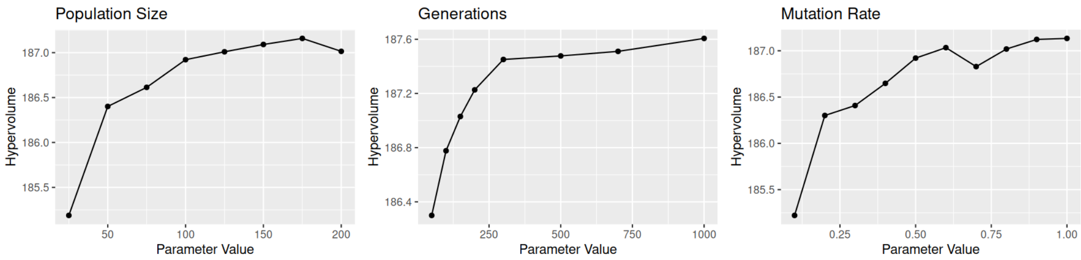



Figure 4.3: Tuning of parameters with respect to hypervolume.

| Population size | Generations | Mutation rate |
|---|---|---|
| 100 | 300 | 0.6 |

Table 4.1: Chosen parameters based on hypervolume.

Using these parameters, the algorithm approximates the true Pareto front fairly well, as can be seen in Figure 4.4. Here, the true Pareto front was found by calculating the objective values for all possible placements of two fire stations. A plot of the true Pareto front with objective values can be seen in Appendix B.2. This brute force method took about two hours to compute, while the NSGA-II with the specified parameters took just under two minutes on the same computer. This further motivates the use of NSGA-II.

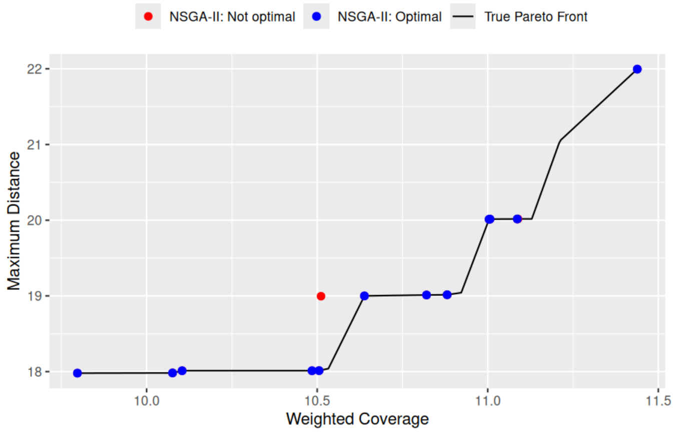



Figure 4.4: True Pareto frontier compared to the results from the NSGA-II algorithm.

In Figure 4.4, it can be seen that most of the solutions found by NSGA-II (blue points) lie on the true Pareto front, while only one of them is not Pareto optimal. At the same time, each of the "bends" in the black line represents a solution on the true Pareto front, which means that even though NSGA-II finds optimal solutions, it does not capture all possibilities. Still, it is a fairly good approximation and will be used to create results in the following chapter.

# 5 | Performance Analysis

When using the solution method described in the previous chapter, mainly two scenarios are of interest. One is the case where the fire department keeps their current location and simply adds another station in the area. The other is the case where they remove the current station and place two freely. The latter is, of course, more expensive, but might lead to a better performance. This chapter is dedicated to analyzing the performance in each of the two scenarios with respect to the efficiency and equity objectives, while ultimately proposing a set of solutions that the fire department can choose from. First, the Pareto fronts achieved in each situation are compared to the current situation in Figure 5.1.

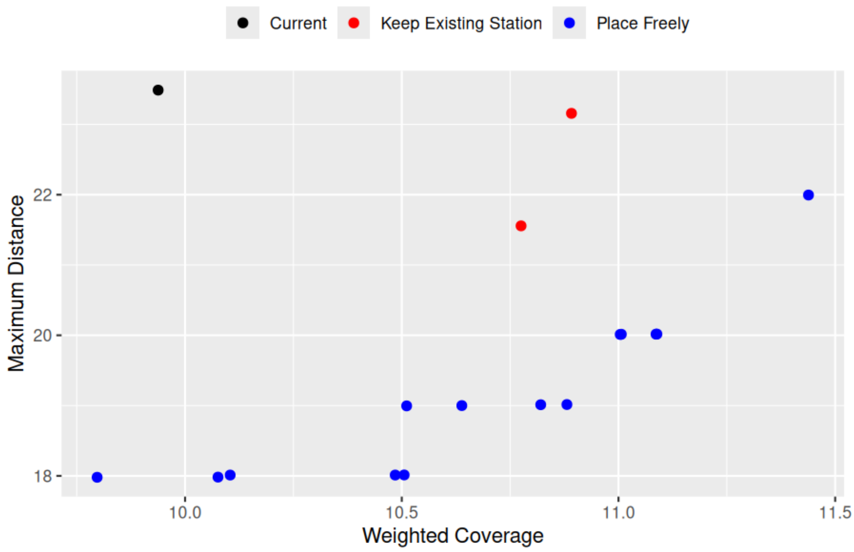



Figure 5.1: Comparison of possible Pareto-optimal solutions in different scenarios.

From Figure 5.1 it can be seen that regardless of which two-station scenario is chosen, the performance in both objectives is generally improved. Keeping the current fire station and placing one more leads to a noticeable improvement in weighted coverage, while also improving the maximum distance. Further improvement is achieved if two stations are placed freely. An exception is found in the lower left corner, where the weighted coverage surprisingly is slightly worse than in the current situation. However, this solution achieves the best performance in the $p$-center objective. For a more precise comparison, the actual objective values are shown in Table 5.1.

| | Weighted Coverage | | Maximum Distance | | Mean Response Time | |
|---|---|---|---|---|---|---|
| | Value | Impr, % | Value | Impr, % | Value | Impr, % |
| **Current** | 9.9379 | 0 | 23.4876 | 0 | 16.2368 | 0 |
| **Keep One** | 10.8916 | 9.6 | 23.1571 | 1.41 | 14.0748 | 13.32 |
| | 10.7753 | 8.43 | 21.5566 | 8.22 | 13.1087 | 19.27 |
| **Place Freely** | 9.7971 | -1.42 | **17.9803** | **23.45** | **11.9207** | **26.58** |
| | 11.4384 | 15.1 | 21.9954 | 6.35 | 12.1074 | 25.43 |
| | **11.0886** | **11.58** | 20.0172 | 14.78 | 12.2701 | 24.43 |
| | 10.5112 | 5.77 | 18.9962 | 19.12 | 12.3925 | 23.68 |
| | 10.881 | 9.49 | 19.0153 | 19.04 | 12.2094 | 24.8 |
| | 10.5055 | 5.71 | 18.0134 | 23.31 | 12.0885 | 25.55 |
| | 11.004 | 10.73 | 20.0121 | 14.8 | 11.9452 | 26.43 |
| | 10.1041 | 1.67 | 18.012 | 23.31 | 12.0678 | 25.68 |
| | 10.4849 | 5.5 | 18.012 | 23.31 | 12.0495 | 25.79 |
| | 10.076 | 1.39 | 17.9827 | 23.44 | 12.0283 | 25.92 |
| | 10.6385 | 7.05 | 19.0002 | 19.11 | 12.8505 | 20.86 |
| | 10.8206 | 8.88 | 19.0125 | 19.05 | 12.3088 | 24.19 |
| | 11.0066 | 10.75 | 20.0145 | 14.79 | 12.0106 | 26.03 |
| | 11.0866 | 11.56 | 20.0162 | 14.78 | 12.1808 | 24.98 |

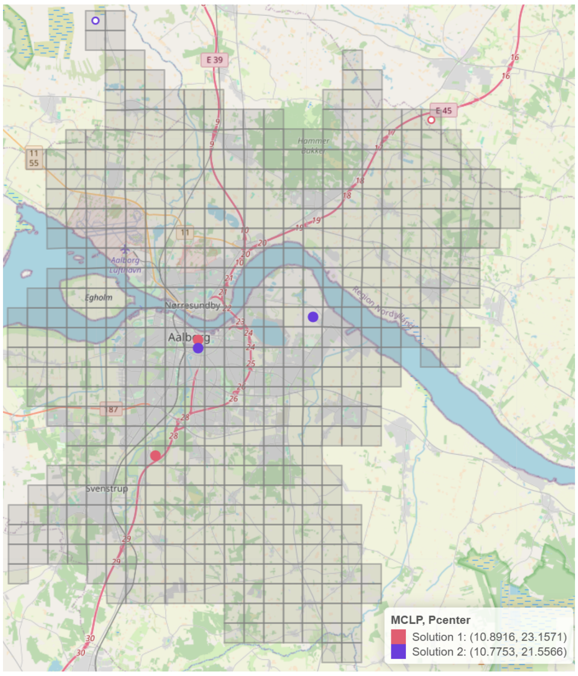
Table 5.1: Comparison of objective values for optimal solutions in the different scenarios.

In Table 5.1, the percent-wise improvement for each of the solutions and objectives when compared to the current situation is also included. The solutions that achieve the biggest improvement are marked in bold. In terms of objective values, the table shows the same general trend as Figure 5.1, where the improvement in coverage tends to vary more than in the case of distance, while the distance objective experiences the greatest improvement of 23.45%.

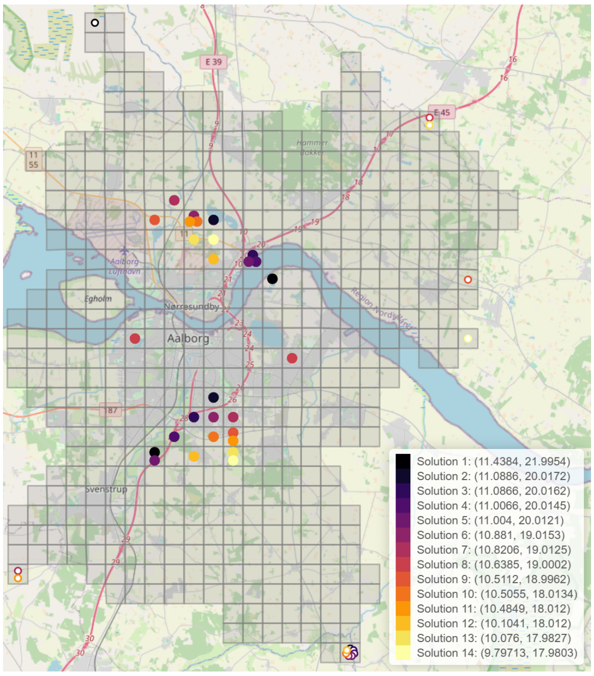
The last two columns in Table 5.1 concern the response time. Even though this was not one of the objectives that were optimized directly, it might still be of interest to the decision maker. Indeed, it can be seen that the mean response time is improved by up to three minutes, corresponding to about 20%, by just placing one more station, while placing two freely can give improvement up to nearly 27% or over four minutes. These response times are calculated using the same mean travel time as was used in calculating the coverage radius.

The main reason for including the coverage objective was to improve upon the fact that 39% of the emergency responses did not satisfy the service goals. Thus, it makes sense to investigate the performance of the proposed solutions with respect to this number. However, since the model is only an approximation of reality, it does not make sense to compare the solution directly with the 39%, but rather the performance of the current solution in the approximated scenario. Considering the current location of the fire station and the coverage radius used in the model, 60.2% of the demand points are not covered within 10 minutes. This is worse than the actual situation, but it is probably caused by the coverage radius being a worst-case scenario. The best performing solution with respect to coverage from Table 5.1 achieves a failure rate of 36.9%, which is a relatively large improvement. Indeed, all of the solutions, where two new stations are placed, achieve a failure rate between 36.7% and 46.1%, with most of them being below 39%. At the same time, the two solutions where only one new fire station is placed lead to failure rates of 48.4% and 45.5% respectively. Thus, all of the proposed solutions improve this aspect of the performance.

To analyze the distribution of these solutions geographically, they are plotted in Figure 5.2 below.




(a) Map of Pareto optimal solutions when keeping the current fire station and placing one more. The farthest points are added with white centers.




(b) Map of solutions from NSGA-II, when placing two fire stations freely. The solutions are sorted by objective values, and the farthest points are added with white centres.

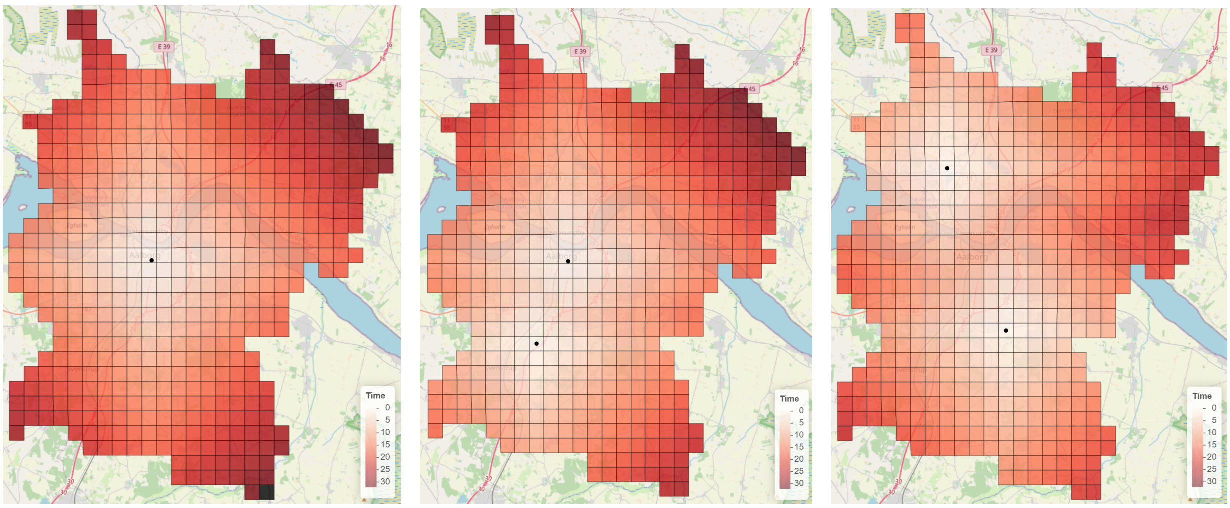
Figure 5.2

From Figure 5.2a, it can be seen that both solutions have a fire station placed near the city center, which is the location of the current station. The solutions are paired by color, and both of the suggested new locations are placed south of the Limfjord. One might expect that having one station on either side would be beneficial with respect to the added response time of crossing the Limfjord. However, the use of the Manhattan distance ignored the influence of the Limfjord, and the current fire station might therefore cover the most demanding parts of Nørresundby.

The results in Figure 5.2b behave more intuitively in the sense that most of the solutions have a station placed on either side of the fjord. This corresponds better with the fire department's expectations, and when presented with these results, they argued that the locations to the northwest are interesting, since they allow for cooperation with the airport's fire department, enabling the possibility to share the costs. At the same time, the locations to the southeast were also deemed beneficial since they are close to important infrastructure such as the new hospital.

A last noteworthy aspect of the plots in Figure 5.2 is the smaller points with white circles in the middle. These represent the farthest point from the stations in the different solutions, i.e. the point determining the value of the $p$-center objective. Especially in the second case, it can be seen that a single point determines this value in most of the cases, namely the point in the southeastern corner. This indicates that some demand points might have too much influence on this objective, which could argue against the representativeness and use of this objective.

A last way to compare the scenarios geographically is with respect to response times. Again, it is not strictly the focus for the optimization, but it still indicates some of the performance of the solutions. In Figure 5.3, the response times for three solutions from the different scenarios are shown, where darker colors indicate longer response times.



(a) Current scenario. (b) Keep existing station. (c) Place two freely.

Figure 5.3: Map of response times in different scenarios.

It is clear from Figure 5.3, that in the current situation, the city center receives the best service, while the outskirts are serviced more poorly. Placing the two stations freely seems to slightly reduce the service in the center, but improves the service in the outer areas more significantly. Thus, the two new stations seem to provide more equitable service. In general, this latter scenario performs the best, and the following analysis and results will therefore focus on this case.

## 5.1 Distance to Ideal

A common factor in the results from the previous section is that they are meant as support for the decision maker, who can choose one of the possible solutions. If the purpose instead was to propose a single solution, one method would be to incorporate specific preferences from the decision maker. However, such preferences are not available, and other methods must therefore be explored. Inspired by Marler and Arora [54], one way to choose the best solution along the Pareto front is to use the *ideal*, also called the *utopia* point. This point is a combination of the optimal achievable value for each of the objectives. In general, reaching this point is not possible, and consequently, the closest point must be the next best thing.

Before calculating the distances to the ideal for each of the solutions, the two objectives are normalized to a scale between the nadir point and the ideal point. Afterwards the *Euclidean* distance is calculated for each of the solutions on the Pareto front, and the best solution is found as the one with the shortest distance. Figure 5.4 shows the result of this process.

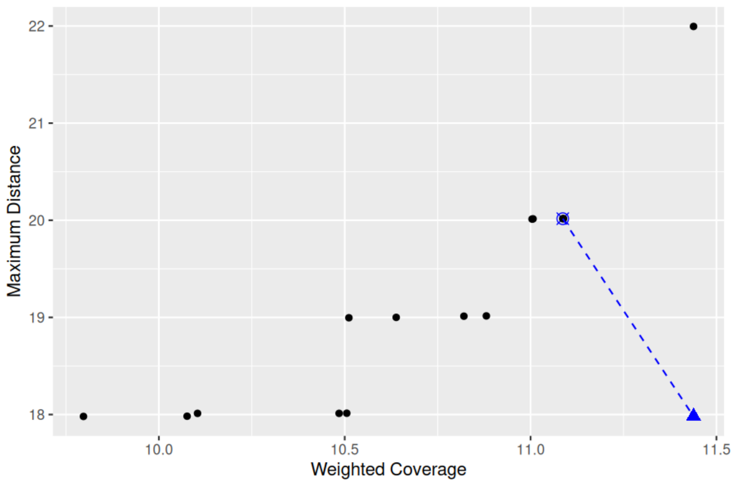



Figure 5.4: Optimal point when measuring the distance to the ideal (blue triangle).

In Figure 5.4, the ideal solution is marked by a blue triangle while the closest solution is also marked in blue. It should be mentioned that the values in Figure 5.4 are not normalized and it might therefore seem that other solutions are closer than the one marked in blue. The blue solution presents a good trade-off between the two objectives and can be seen geographically in Figure 5.5

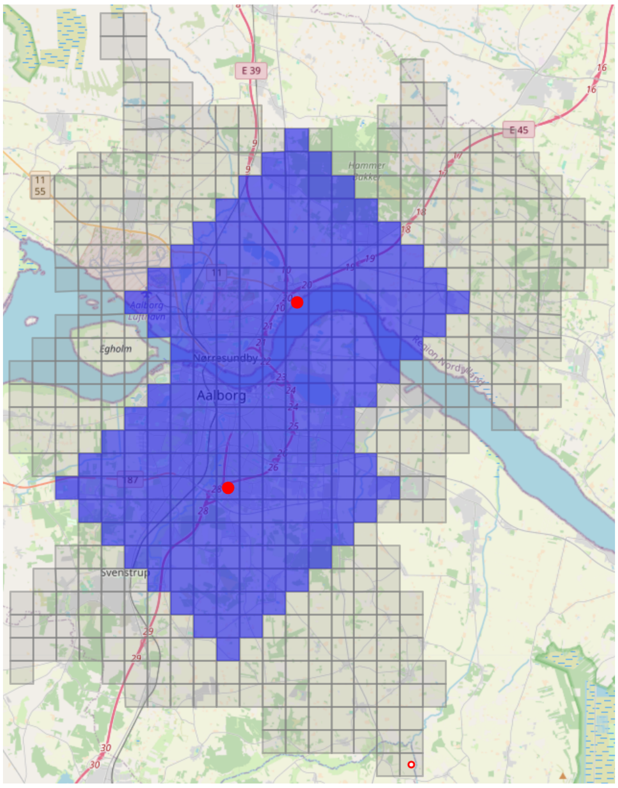



Figure 5.5: Map of the solution closest to the ideal, with coverage area in blue and stations in red. The farthest point is added with a white center.

The map in Figure 5.5 shows a solution where a station is placed on either side of the Limfjord. At the same time, both seem to be located in squares with highways and major road junctions, which, even though it was not incorporated in the model, might influence the response times. At the same time, it can be seen that the two stations cover most of the densely populated areas. However, many squares are not covered within the 10-minute boundary. A fact that could be helped by adding another station.

## 5.2 More Fire Stations

Even though the fire department has limited funds, it is still theoretically interesting to analyze what happens if more stations are added. To investigate this behavior, the NSGA-II algorithm is applied with an increasing number of stations between two and ten. A benefit of the NSGA-II is that the computation time is approximately the same, regardless of whether two or ten stations are being placed. The result of this application is seen in Figure 5.6.

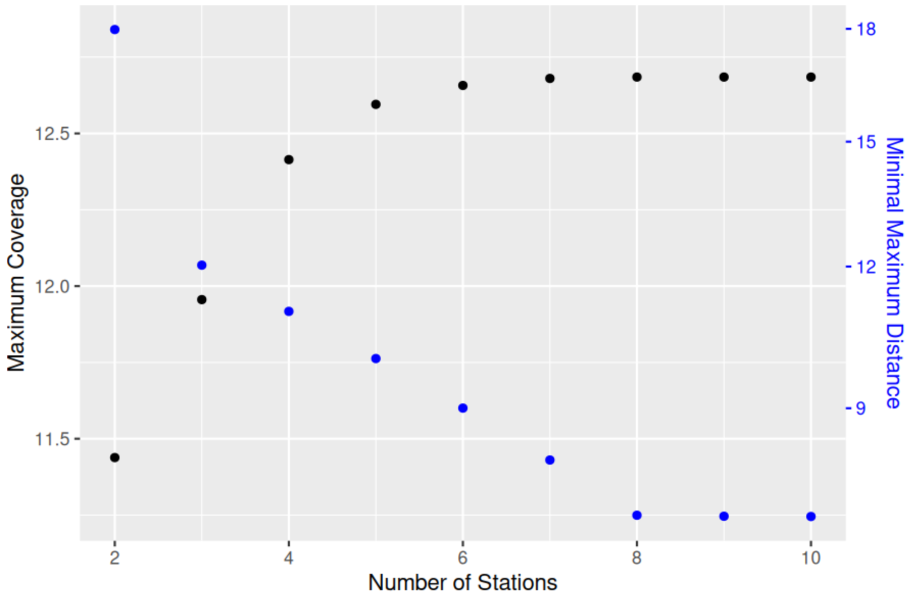



Figure 5.6: Optimal objective values for solution fronts obtained with different numbers of fire stations.

The vertical axes in Figure 5.6 represent the best achieved value in each of the objectives across the whole Pareto front. Thus, for a certain number of fire stations, the black dots represent the maximum achieved weighted coverage for any of the Pareto optimal solutions. The same applies to the blue, but with respect to the minimum value of the $p$-center objective. It seems that quite a large improvement in both objectives occurs when the number of fire stations is increased to three. For comparison, an example of such a solution is seen in Appendix C.1 along with a plot of the approximated Pareto front. From the Pareto front, it can be seen that especially the maximum distance objective is improved compared to the previous case of two stations.

As expected, both objectives seem to reach a saturation point. For the coverage, this happens at 8 stations, where all points are covered, and the objective cannot be increased. For the maximum distance, this also seems to happen at 8 stations, but small improvements still happen to the following values. In fact, when placing eight stations, the NSGA-II algorithm only found one Pareto optimal solution. This can be seen in Appendix C.2, where indeed all points are covered.

# 6 | Other objectives

It was earlier mentioned that one point controlled the value of the $p$-center objective in a lot of the cases, where it was also argued that other objectives might be more representative. This chapter is therefore dedicated to exploring other objectives and their effect on the solutions. To keep the interpretation and structure of the model, new objectives will not be added to the model, but instead replace the current ones, such that one efficiency and one equity objective is always included.

## 6.1 Gini Coefficient

In terms of equity, Marsh and Schilling [27] describes multiple objectives which has been used in relation to the facility location problem. These include the already used $p$-center objective, but also more complex objectives like minimizing the range of distances or the variance. Other possibilities include the *coefficient of variation*, which takes into account the standard deviation, or *Thiel's entropy coefficient*, which uses log-distances. However, for this project, the *Gini coefficient* is used, based on its wide usage in other fields, such as economics and social sciences, and its capability of taking all demand points into account instead of being determined by a single one. For a total of $N$ demand points, the Gini coefficient is defined as

$$\frac{\sum_i \sum_h |d_i - d_h|}{2N^2\bar{d}},$$

where $d_i$ is the distance from demand point $i$ to its assigned station and $\bar{d}$ is the mean of all these distances. The Gini coefficient always returns a value between 0 and 1, where 0 means total equity and 1 is the opposite. This easy interpretation also motivates the use of this objective.

When implementing this objective, instead of the $p$-center, the NSGA-II algorithm leads to the Pareto optimal solutions seen in Appendix D.1. In this regard, it should be mentioned that the same parameter values are used in the NSGA-II as before, which might create a small bias in the model, since it is no longer tuned to the current objectives. From Figure D.1 it can be seen that a larger number of possible solutions are found when using this new objective, which could be explained by the fact that the Gini-coefficient allows for more subtle changes in value, compared to the $p$-center. At the same time, these solutions also behave differently with respect to the objectives. It appears that there is an area in the middle where the weighted coverage is reduced significantly, but the Gini coefficient is barely improved.

To analyze the Pareto front spatially, the solutions are shown on a map in Figure 6.1, where each colored pair of stations represents a different solution. Additionally, the solutions are also sorted by objective values, such that lighter colors represent a lower and more optimal Gini coefficient.

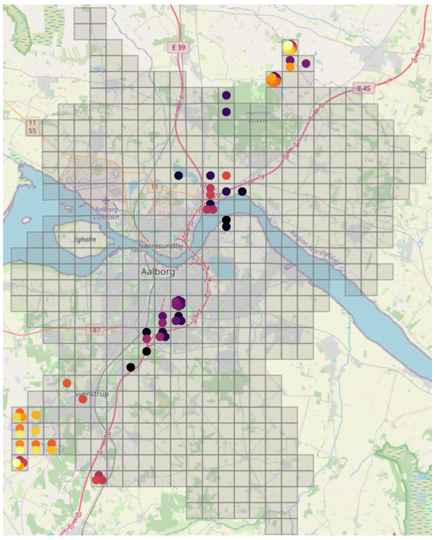



Figure 6.1: Map of Pareto optimal solutions with the Gini coefficient as equity objectives.

The optimization of the Gini coefficient shows an interesting spatial behavior. From Figure 6.1 it appears that when the Gini coefficient is decreased, the stations are moved towards the edges of the area. However, in terms of coverage, these solutions only achieve weighted coverage values close to zero. Thus, if equity is optimized, all citizens receive equally bad service, and in practice, a more fitting trade-off should be chosen instead.

## 6.2 Backup Coverage

In relation to the efficiency objective, multiple approaches were mentioned in the literature review. One of these was the $p$-median objective, which optimized the mean distance between the fire stations and their assigned demand points. Others used the coverage formulation, but added gradually decreasing coverage instead of binary, which was used in the previous analysis. In the earlier part of the project, it was also assumed that no emergencies could happen at the same time, even though the data showed a small possibility of it happening. This section, therefore, explores the scenario where this can happen and *backup coverage* is used as an objective instead of the normal coverage objective from the MCLP. It is combined with the $p$-center objective.

The concept of backup coverage is concerned with having two stations cover each point. Strictly speaking, a point is only considered covered if it is within the coverage radius of two stations. However, it is often combined with a normal coverage objective, such that both aspects contribute to the objective value. This means that the backup coverage model is bi-objective in itself. Inspired by Hogan and Revelle [26] and in order to combine it with the equity objective and still restrict the overall model to two objectives, the normal and backup coverage is combined into one objective as a linear combination. This is done with a weight $\omega \in [0, 1]$, such that the coverage objective becomes

$$\text{maximize} \quad \omega \sum_i a_i y_i + (1 - \omega) \sum_i a_i u_i,$$

where $u_i$, for $i \in \mathcal{I}$, is a decision variable defined as

$$u_i = \begin{cases} 1 & \text{if demand } i \text{ is covered by two stations,} \\ 0 & \text{otherwise.} \end{cases}$$

The result of this model, with a weight of $\omega = 0.5$ is seen in Appendix D.2. Again, it appears that a larger number of Pareto optimal solutions are found in this case than in the first proposed model. However, these solutions seem to behave more similarly to the ones in Figure 5.1. The spatial behavior of these solutions can be seen in Figure 6.2.

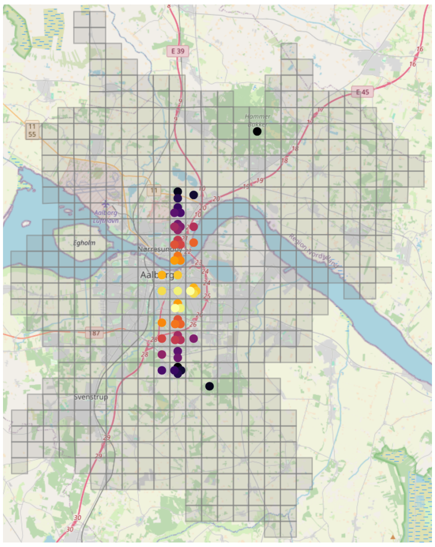


Figure 6.2: Map of Pareto optimal solutions with backup coverage and p-center objectives.

It appears from Figure 6.2 that this objective behaves oppositely to that of the Gini coefficient. Instead of pushing the stations to the boundaries, they are pulled towards the center when backup coverage is optimized. This makes sense, since the maximum number of demand points covered twice is achieved if the fire stations are in the same location. At the same time, they are placed in the center, since this is where the demand points are weighed the highest. Again, these solutions would not be very practical, since the two fire stations would require double the cost of a single station, while covering approximately the same area. Thus, a tradeoff would be a better solution, or it might be better to solve this objective by having multiple fire engines and response teams ready at the same station.

# 7 | Discussion

Throughout the project, multiple assumptions and choices have been made, affecting how representative and realistic the model and analysis are. Initially, most of these were related to the choice of model and objectives. One of the main choices was to model and optimize coverage instead of response times directly. Even though this choice was based on the data analysis and the use of service goals with high failure rates, coverage does not correspond exactly to response times, and it might therefore have given a more direct interpretation for the decision maker if the response times were optimized instead. With respect to the coverage, other aspects might also have been included to increase the realism of the model. Since the fire department does not simply ignore the emergency calls outside their 10-minute boundary, gradual coverage might have been a more realistic representation than the binary coverage. On the other hand, the not-covered demand points were not ignored completely, due to the inclusion of an equity objective.

The multi-objective approach was chosen to account for the complexity of the fire station location problem, which includes different and often conflicting goals. However, only two objectives were considered at a time, which might not capture the whole picture. Thus, more objectives might have been included simultaneously, such as backup coverage or minimizing cost. Other aspects could also have been included as objectives, such as water accessibility, equality of demand loads between the fire stations, or overlap of coverage areas. Nevertheless, the bi-objective approach captured some of the complexity, while it also allowed for easier interpretation of the results.

With respect to the model constructed in Chapter 2, the constraints included were relatively simple. This was based on the previously made assumptions, but again, other aspects could have been included to make the model more realistic and representative of the actual situation. Specifically, the inclusion of staffing and equipment could be used to make the stations heterogeneous, which is more realistic since not all the current fire stations are exactly the same. However, this would have been more relevant if all of Northern Jutland had been considered, and since the motivation from the fire department was to create one more full-time station, it was deemed a fair assumption that the stations would be homogeneous. Still, it might have been interesting to explore the possibility of only opening a part-time fire station instead. In terms of equipment, it would require a deeper analysis of where the different types of equipment are needed the most, which could be interesting and relevant to include, but was outside the scope of this project.

In terms of demand and response times, these were assumed to be deterministic. An assumption that could be argued to be too simple and unrealistic. Statistically, there were small but still significant changes in demand at different periods, which could have been included through uncertainty or the use of a dynamic model. On the other hand, the data analysis showed very little variation in the response times, indicating that this was not a major problem for the department. Still, stochastic demand might have been more realistic, and in such a case, a simulation of the situation could have been beneficial and given a better indication of the performance. The robustness of the solution could also have been investigated through simulation, making the results more trustworthy. However, based on the assumptions, simulation was deemed unnecessary in this project.

Moving on to the implementation of the model, a major simplification was made with respect to the distance calculation. In this case, the Manhattan distance was used as an approximation of the actual road distances. This simplification also ignored the influence of the Limfjord. In fact, when adding a north/south indicator as a factor variable to the regression model in Figure 1.8, the coefficient indicates that if an emergency is north of the fjord, the expected response time should be increased by 0.76 minutes. Thus, it might have been more realistic to include this factor in the distance calculation. Despite these arguments, the use of the Manhattan distance made the modeling possible and could be argued to be still more representative than other measures, such as the Euclidean distance. In this connection, it is also noteworthy that some of the initial data analysis was made using actual road distances from the Google API, which unfortunately was not possible for the implementation. However, implementation-wise, the distances are kept in a distance matrix, which is easily changed if actual road distances could be calculated.

A last aspect of the model construction is the fact that Aalborg was analyzed and optimized in isolation. Even though the chosen area accounted for about 92% of the station's demand points, the model ignored the possibility of support from other surrounding fire stations, which, as explained by the fire department, is sometimes used in case of large emergencies. Thus, if a larger area were to be optimized, cooperative coverage might be beneficial to include and make the model more realistic as well. In relation to the area and grid construction, it was chosen to weigh the demand points by population and incidents. However, this might not be the best representation, since it ignores other aspects such as industrial zones and infrastructure, which might make a demand point more important. At the same time, the population and number of incidents seem to correlate, which could argue that it is unnecessary to include both. Thus, a more representative weighing of the demand points might have been beneficial.

The last part of the discussion is dedicated to the solution method. Here, NSGA-II was used based on its ability to incorporate multiple objectives and its successful use in the literature. At the same time, the use was also motivated by the computation time being too long for analytical methods. However, since this is a decision made on a strategic level, computation time is not as important as on the tactical or operational levels. Nevertheless, the time constraint for this project made it a necessary inclusion, which also allowed the analysis of the cases where more stations were placed. This would have been even more computationally demanding for analytical or brute force methods, but barely affected the NSGA-II.

With respect to the NSGA-II, the choice of crossover and mutation operators could also be points for discussion. The ones used in the project were relatively simple, and it might have been beneficial to use other and possibly problem-specific operators. However, this might increase computation time, and the current implementation still gives relatively good approximations of the Pareto front.

# 8 Conclusion

As mentioned in the problem statement, the purpose of this project was to formulate and solve the fire station location problem in Aalborg as a multi-objective optimization problem incorporating both efficiency and equity. This was based on the initial analysis of GIS data from the last eight years, where it was found that a large part of the emergencies are not responded to within the desired service goal, and that some of the population experienced longer response times than others. To solve the problem, it was first stated mathematically as a mixed integer linear programming problem, which incorporated the objectives from the maximum coverage location problem and the $p$-center problem.

To implement this formulation in a practical setting, the Aalborg and the surrounding area were divided into a grid of $1 \times 1$ km squares. These were used as both possible locations for fire stations and demand points, with the service loads determined by the number of emergencies and population in each square. The facility location problem was then solved by finding the Pareto optimal solutions with the Non-dominated Sorting Genetic Algorithm (NSGA-II). The algorithm was tuned with respect to hypervolume in order to find the optimal setting for the parameters. Here, it was found that more exploration of the solution space in general led to better solutions in terms of hypervolume.

Two scenarios were solved and compared to the current situation. One placing an extra station, while keeping the current one, and another placing two stations freely. Here, it was concluded that both cases led to a general improvement in the two objectives when compared to the current situation, though the second case led to the most optimal solutions overall with respect to the objectives. Indeed, the largest observed improvement in the coverage objective was 11.58%, while it was 23.45% for the equity objective. Additionally, it was found that the best solution with respect to keeping the service goal led to a decrease in the failure rate of 23.3 percentage points, according to the model.

Further analysis was made to explore the situation, where more fire stations could be placed. This led to the expected conclusion that more fire stations improved the performance, but at the same time, it was seen that a saturation point was achieved at eight stations, in which case, the whole area was covered. A last addition also explored the influence of the chosen objectives by changing them individually. Two new objectives were tested, namely the Gini coefficient and the use of backup coverage. It was concluded that optimizing the Gini coefficient tended to push the stations to the boundaries, resulting in equally bad service to all demand points, whereas optimizing the gradual coverage gave the opposite result and dragged both stations to the center, thereby covering the most important demand points twice.

In conclusion, it was possible to formulate the fire station location problem in Aalborg as a multi-objective optimization problem optimizing both efficiency and equity, while solving the problem led to an improvement in both objectives. In this regard, the use of NSGA-II made for a fast and adaptable solution approach, which allowed easy changes in the parameters and variables of the model.

## 8.1 Future Work

As mentioned in the discussion, the approach used in this project had some drawbacks and assumptions that made certain aspects less realistic. Thus, future work should focus on improving these aspects. First, a more realistic and representative distance matrix should be used to make the results more justifiable. This could be done by calculating the actual road distances among the demand points instead of using the Manhattan distance. At the same time, the Limfjord connections should be taken into account in the new distances as well.

Another possibility for further work would be to improve the weighing of the demand points. It could be made more realistic by including other aspects such as industrial zones, cultural sights, or other socio-economic factors. Knowledge and experience from the emergency department could also be incorporated to make the weighing more representative of the real situation.

Lastly, it might be interesting to expand the model to incorporate more or all of Northern Jutland, instead of only optimizing Aalborg in isolation. In this case, it might be beneficial to allow for heterogeneous fire stations in terms of staffing, equipment, and coverage radius, to better represent the situation.

# 9 | Bibliography

[1] Analyse og statistik om redningsberedskabet [Article]. www.brs.dk/da/borger/om-redningsberedskabet-i-danmark/analyse-og-statistik-om-redningsberedskabet/: Beredskabsstyrelsen; 2025 [cited 05-05-2025].

[2] Kiran KC JC. Modelling residential fire incident response times: A spatial analytic approach. Applied Geography. 2017;84:64–74.

[3] Savas ES. On equity in providing public services. Management Science. 1978;24(8):800–808.

[4] Postnummerinddeling [Data]. https://dataforsyningen.dk/data/4841: Dataforsyningen; 2025 [cited 05-05-2025].

[5] Aleisa E. The fire station location problem: A literature survey. International Journal of Emergency Management. 2018 01;14:291.

[6] Routes API [Documentation]. https://developers.google.com/maps/documentation/routes: Google Maps Platform; 2025 [cited 05-05-2025].

[7] Cepiku D, Mastrodascio M. Equity in Public Services: A Systematic Literature Review. Public Administration Review. 2021;81(6):1019–1032.

[8] Hassler J, Ceccato V. Socio-spatial disparities in access to emergency health care—A Scandinavian case study. PLoS One. 2021;16(12):e0261319.

[9] Chung S, Smith M, Jin A, Hogewood L, Kitsak M, Cegan J, et al. Access to Emergency Services: A New York City Case Study. Transportation Research Interdisciplinary Perspectives. 2024;25:101111.

[10] Kenny A Hendrickson AMC Kula A Francis. Public Trust, Community Resilience, and Disaster Response in the U.S. Virgin Islands. Natural Hazards Center. 2022;.

[11] NHS facing ‘crisis of public trust’ as most people fear being failed by A&E services [Article]. https://www.theguardian.com/society/2025/feb/22/three-four-people-uk-fear-failed-ae-services-nhs: The Guardian; 2025 [cited 22-05-2025].

[12] Farahani RZ, Abedian M, Sharahi S. Dynamic facility location problem. In: Facility location: Concepts, models, algorithms and case studies. Springer; 2009. p. 347–372.

[13] Hajipour V, Fattahi P, Bagheri H, Babaei Morad S. Dynamic maximal covering location problem for fire stations under uncertainty: Soft-computing approaches. International Journal of System Assurance Engineering and Management. 2022;13(1):90–112.

[14] Ming J, Richard JP, Qin R, Zhu J. Distributionally robust optimization for fire station location under uncertainties. Scientific Reports. 2022 03;12. https://doi.org/10.1038/s41598-022-08887-6.

[15] Hakimi SL. Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. Operations Research. 1964;12(3):450–459.

[16] Alp O, Erkut E, Drezner Z. An efficient genetic algorithm for the p-median problem. Annals of Operations research. 2003;122:21–42.

[17] Resende MG, Werneck RF. A hybrid heuristic for the p-median problem. Journal of heuristics. 2004;10:59–88.

[18] Toregas C. A covering formulation for the location of public facilities [Ph.D. Thesis]. Cornell University Ithaca, NY, USA; 1970.

[19] Church R, ReVelle C. The Maximal Covering Location Problem. Papers in Regional Science. 1974;32(1):101–118.

[20] Aktas E, Ozaydin O, Bozkaya B, Ulengin F, Önsel S. Optimizing Fire Station Locations for the Istanbul Metropolitan Municipality. Interfaces. 2013 02;43.

[21] Idayani D, Puspitasari Y, Sari LDK. Maximal Covering Problem Model for Determination of Fire Station Location. In: International Conference on Innovation in Science and Technology (ICIST 2020). Atlantis Press; 2021. p. 20–23.

[22] Murray AT. Optimising the spatial location of urban fire stations. Fire Safety Journal. 2013;62:64–71. Special Issue on Spatial Analytical Approaches in Urban Fire Management.

[23] Berman O, Krass D. The generalized maximal covering location problem. Computers & Operations Research. 2002;29(6):563–581. Location Analysis.

[24] Jabalameli M, Tabrizi B, Javadi M. A Simulated Annealing method to solve a generalized maximal covering location problem. International Journal of Industrial Engineering Computations. 2011 04;2.

[25] Karatas M, Razi N, Tozan H. A Comparison of p-median and Maximal Coverage Location Models with Q–coverage Requirement. Procedia Engineering. 2016;149:169–176. International Conference on Manufacturing Engineering and Materials, ICMEM 2016, 6-10 June 2016, Nový Smokovec, Slovakia.

[26] Hogan K, Revelle C. Concepts and Applications of Backup Coverage. Management Science. 1986;32(11):1434–1444.

[27] Marsh MT, Schilling DA. Equity measurement in facility location analysis: A review and framework. European Journal of Operational Research. 1994;74(1):1–17.

[28] Drezner Z. The p-Centre Problem-Heuristic and Optimal Algorithms. The Journal of the Operational Research Society. 1984;35(8):741–748.

[29] Abubakar Z, Ejima O, Mustafa A, Ibrahim A. p-CENTER PROBLEM OF FIRE STATIONS IN SOKOTO METROPOLIS. Asian Journal of Mathematics and Computer Research. 2017 06;p. 143–151.

[30] Drezner T, Drezner Z, Guyse J. Equitable service by a facility: Minimizing the Gini coefficient. Computers & Operations Research. 2009;36(12):3240–3246. New developments on hub location.

[31] Gunnarsson B, Björnsdóttir KM, Dúason S, Ingólfsson Á. Locating helicopter ambulance bases in Iceland: efficient and fair solutions. Scandinavian journal of trauma, resuscitation and emergency medicine. 2023;31(1):70.

[32] Untadi A, Li LD, Li M, Dodd R. Incorporating Socio-Economic Factors in Maximizing Two-Dimensional Demand Coverage and Minimizing Distance to Uncovered Demand: A Dual-Objective MCLP Approach for Fire Station Location Selection. Axioms. 2024;13(1).

[33] Badri MA, Mortagy AK, Alsayed CA. A multi-objective model for locating fire stations. European Journal of Operational Research. 1998;110(2):243–260. EURO Best Applied Paper Competition.

[34] Yao J, Zhang X, Murray AT. Location optimization of urban fire stations: Access and service coverage. Computers, Environment and Urban Systems. 2019;73:184–190.

[35] Yu W, Guan M, Chen Y. Fire stations siting with multiple objectives and geospatial big data. Earth Science Informatics. 2021;14:141–160.

[36] Moche L. A Multi-Objective Optimisation Approach for the Location of a Public Primary Health Facility [Ph.d. Thesis]. University of the Witwatersrand; 2018. Available from: `https://www.proquest.com/dissertations-theses/multi-objective-optimisation-approach-location/docview/3160883543/se-2`.

[37] Gendreau M, Laporte G, Semet F. The maximal expected coverage relocation problem for emergency vehicles. Journal of The Operational Research Society - J OPER RES SOC. 2006 05;57:22–28.

[38] Macit I. SOLVING FIRE DEPARTMENT STATION LOCATION PROBLEM USING MODIFIED BINARY GENETIC ALGORITHM. European Scientific Journal. 2015 05;11.

[39] Xia L, Xie M, Xu W, Shao J, Yin W, Dong J. An empirical comparison of five efficient heuristics for Maximal Covering Location Problems. In: 2009 IEEE/INFORMS International Conference on Service Operations, Logistics and Informatics; 2009. p. 747–753.

[40] López Jaimes A, Zapotecas-Martínez S, Coello C. In: An Introduction to Multiobjective Optimization Techniques. Nova Science; 2011. p. 29–57.

[41] Wang SC, Chen TC. Using NSGA-II to solve multi-objective competitive location problem with cooperative coverage for distance-based attractiveness. Journal of Intelligent & Fuzzy Systems. 2021 01;40:1–12.

[42] Hu Z, Wang L, Qin J, Lev B, Gan L. Optimization of facility location and size problem based on bi-level multi-objective programming. Computers & Operations Research. 2022;145:105860.

[43] Lopez A, Sánchez-Oro Calvo J, Laguna M. A New Scatter Search Design for Multiobjective Combinatorial Optimization with an Application to Facility Location. INFORMS Journal on Computing. 2020 09;.

[44] Miettinen K. Nonlinear Multiobjective Optimization. 1st ed. International Series in Operations Research & Management Science, 12. New York, NY: Springer US; 1998.

[45] Emmerich M, Deutz A. A tutorial on multiobjective optimization: fundamentals and evolutionary methods. Natural Computing. 2018 09;17.

[46] Siarry P, Siarry P. Metaheuristics. 1st ed. Cham: Springer International Publishing AG; 2017.

[47] Karl F, Pielok T, Moosbauer J, Pfisterer F, Coors S, Binder M, et al. Multi-objective hyperparameter optimization in machine learning—An overview. ACM Transactions on Evolutionary Learning and Optimization. 2023;3(4):1–50.

[48] Li M, Yao X. Quality Evaluation of Solution Sets in Multiobjective Optimisation: A Survey. ACM Computing Surveys. 2019 03;52:1–38.

[49] Grids [Data]. https://ec.europa.eu/eurostat/web/gisco/geodata/grids: Eurostat; 2025 [cited 08-05-2025].

[50] Alzahrani A, Hanbali A. Maximum Coverage Location Model for Fire Stations with Top Corporate Risk Locations: MCLP for Fire Stations with Top Corporate Risk Locations. International Journal of Industrial Engineering and Operations Management. 2021 12;03:58–74.

[51] Schumacher D.: ompr: Model and Solve Mixed Integer Linear Programs. R package version 1.0.4. Available from: https://CRAN.R-project.org/package=ompr.

[52] Chan KC, Tansri H. A study of genetic crossover operations on the facilities layout problem. Computers & Industrial Engineering. 1994;26(3):537–550.

[53] Bhattacharya R, Bandyopadhyay S. Solving conflicting bi-objective facility location problem by NSGA II evolutionary algorithm. The International Journal of Advanced Manufacturing Technology. 2010;51:397–414.

[54] Marler R, Arora J. Survey of Multi-Objective Optimization Methods for Engineering. Structural and Multidisciplinary Optimization. 2004 04;26:369–395.

[55] Madsen H, Thyregod P. Introduction to general and generalized linear models. CRC press; 2010.

# A Supplementary Data Exploration

This appendix includes additional plots related to the data exploration. These primarily concern the linear regression models used in Section 1.1.

## A.1 Model diagnostics

When fitting linear regression models, certain assumptions are made about the data [55]. This section is dedicated to checking these assumptions, which include linearity of the data, normality of the residuals, and homoscedasticity. The diagnostics will be based on graphical interpretations, and the three regression models introduced in Section 1.1 will be investigated individually.

### Monthly Demand Regression

From Figure A.1 it is seen that a linear regression model might not be the best fit to the monthly demand data. The residuals seem to be fairly well scattered around zero, but with a slight tendency towards the negative side. Even though the normality assumption seems to be satisfied, the errors do seem to display some slight Heteroscedasticity. Though the shortcomings might be explained by the lack of data points.

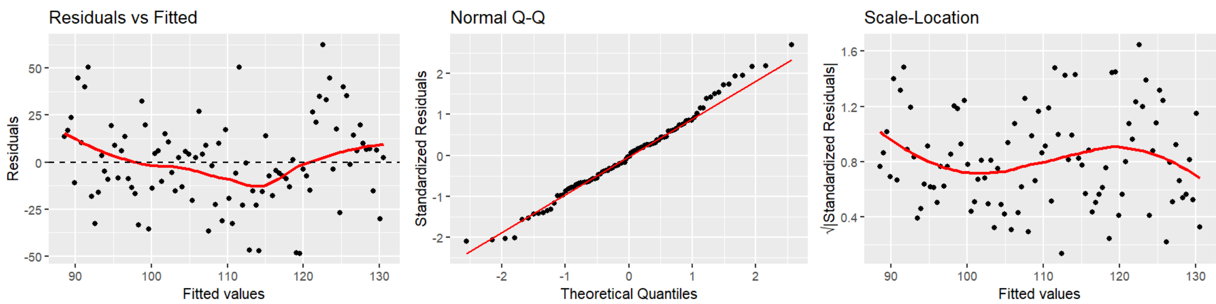



Figure A.1: Model diagnostics for the linear regression model of monthly demand.

### Weekly Demand Regression

The weekly data seems to be a better fit for a linear regression model, since it satisfies both the linearity assumption and has homoscedastic residuals. The residuals do not exhibit quite as good a normal distribution as the monthly model, but still seem reasonably close.

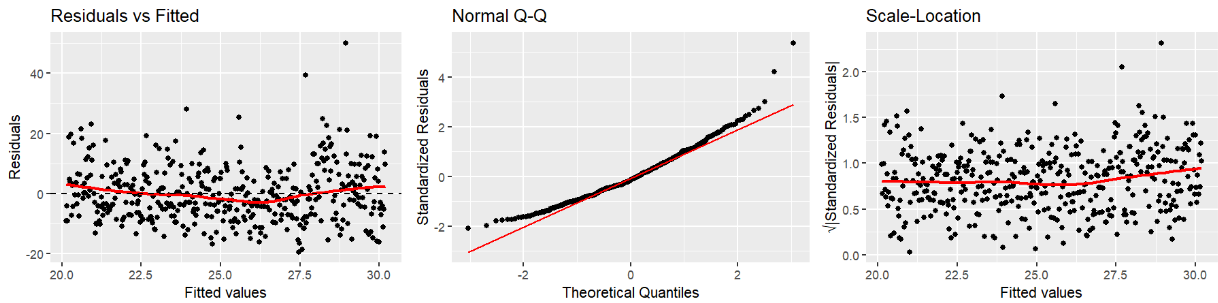



Figure A.2: Model diagnostics for the linear regression model of weekly demand.

## Response Time Regression

This model does not seem to be the right fit to the data, since neither of the three assumptions holds, according to the plots in Figure A.3. However, due to some very influential points, it is difficult to determine from the plots alone. These influential points might even be considered as outliers and could be removed to get a more representative model.

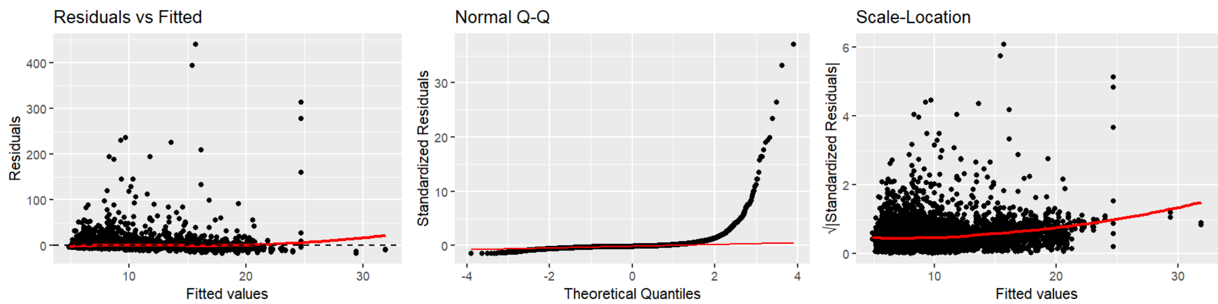



Figure A.3: Model diagnostics for the linear regression of the response times with road distance as explanatory variable.

## A.2 Recent Trend Analysis

This section is motivated by the seemingly higher demand in 2017 and 2018, seen in Figure 1.3. Thus, a new regression model is constructed without data from these years, and the result can be seen in Figure A.4.

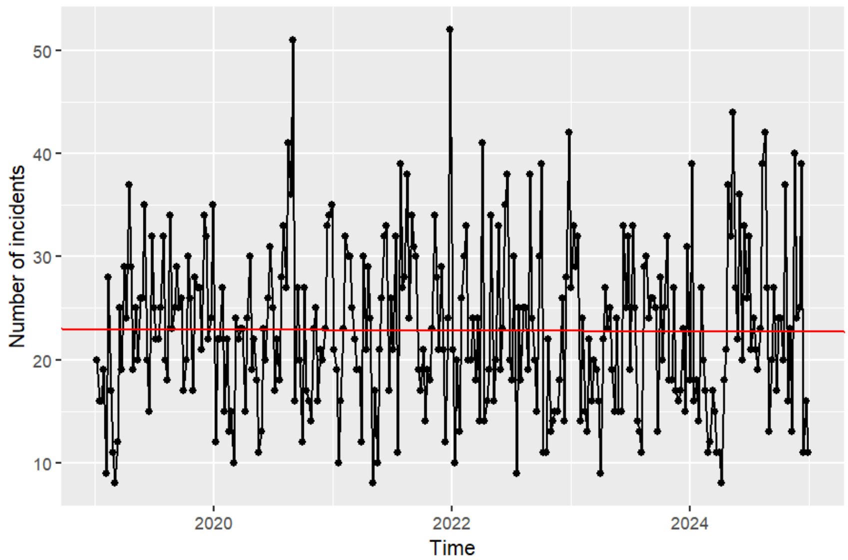



Figure A.4: Weekly trend for the number of incidents between 2019 and 2024.

In this case, the trend coefficient is only $-0.00084$, while it is also statistically insignificant and can be assumed to be 0, according to the $t$-test. The model diagnostics for this model can be seen in Figure A.5 and are very similar to those in the previous weekly demand regression models, except that the residuals seem to be slightly more normal distributed.

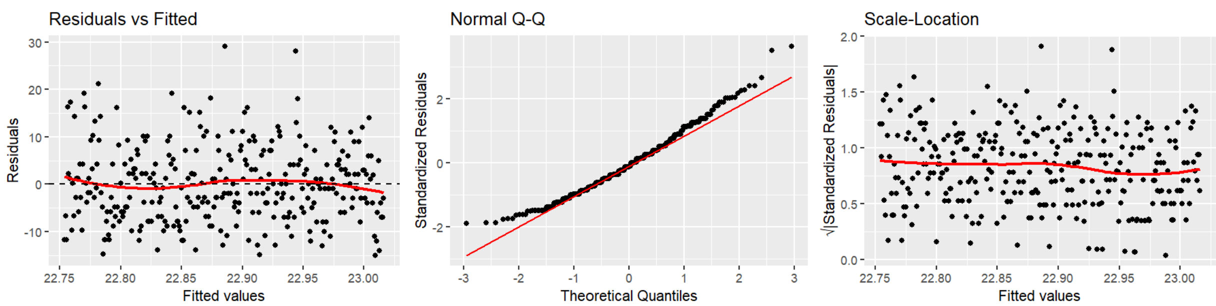



Figure A.5: Model diagnostics for the linear regression model of the weekly demand from 2019 to 2024.

# B Additional Plots for Parameter-Tuning

The following appendix includes additional plots, which are meant as a supplement for the parameter tuning in Section 4.3.

## B.1 Repetitions

The following plot shows the mean hypervolume obtained from different amounts of repetitions. The parameters used were a population size of 50, 100 generations, and a mutation rate of 0.5.

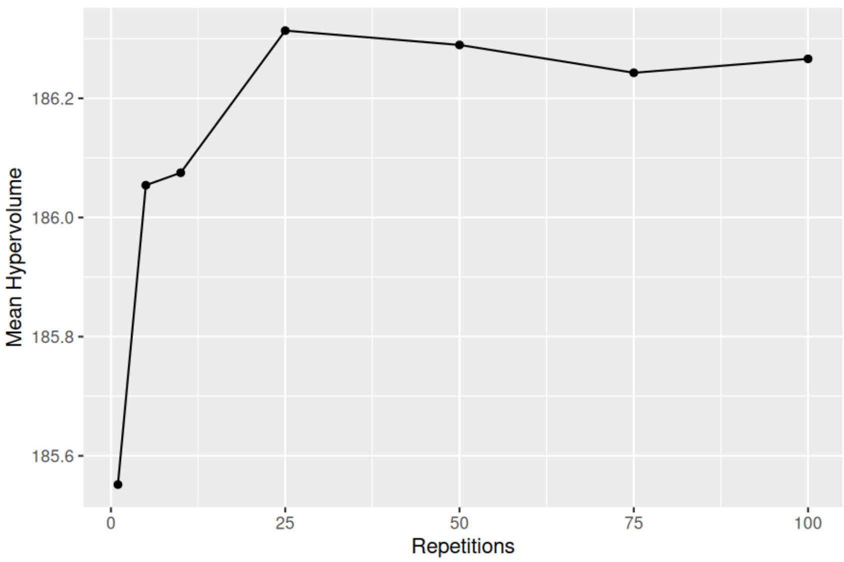



Figure B.1: Mean hypervolume for different numbers of repetitions.

Figure B.1 shows that the mean hypervolume does not change significantly above 25 repetitions. This number of repetitions is therefore used in the tuning process described earlier.

## B.2 True Pareto Front

The plot in Figure B.2 shows the true Pareto front of the model that optimizes weighted coverage and maximum distance. The solutions are obtained with brute force, and the objective values are shown for each possible Pareto optimal solution. The figure is mainly included for the sake of completeness and will not be commented further upon.

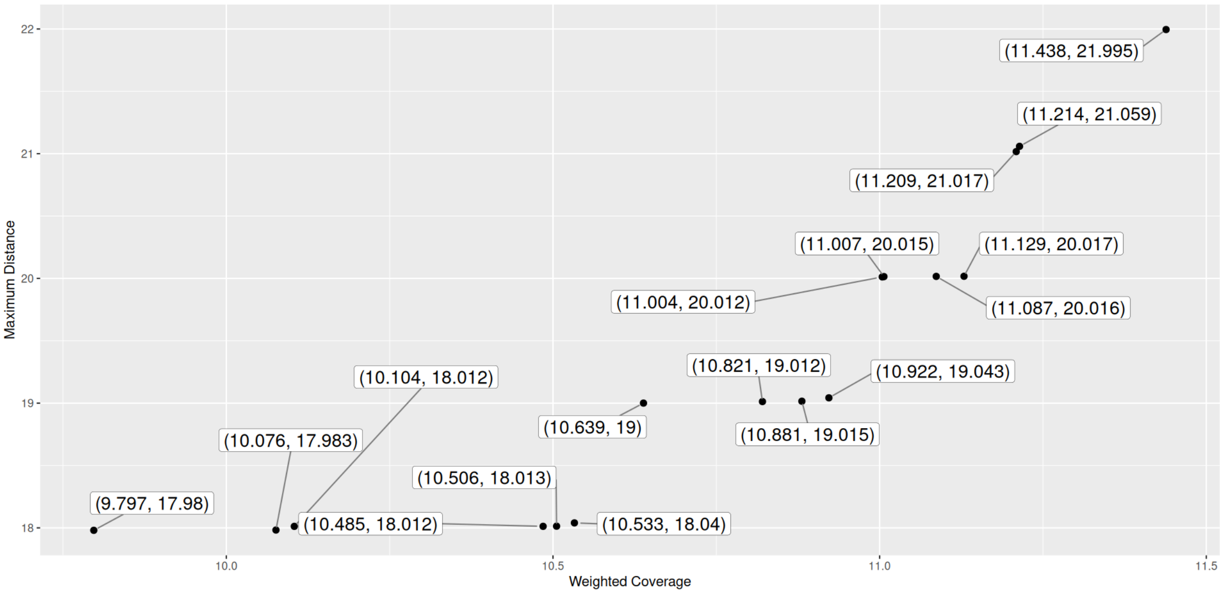



Figure B.2: True Pareto front from brute force evaluation of solution space.

# C Plots with More Stations

In this appendix, the result of adding more stations can be seen. It includes the cases where three and eight stations are to be placed. The plots are meant for comparison and will therefore not be explained or analyzed in depth.

## C.1 Three Stations

The following Figure C.1 shows the Pareto front of the model with coverage and $p$-center objectives, in the case where three stations have to be placed. An example of a solution from Figure C.1 can be seen geographically in Figure C.2.

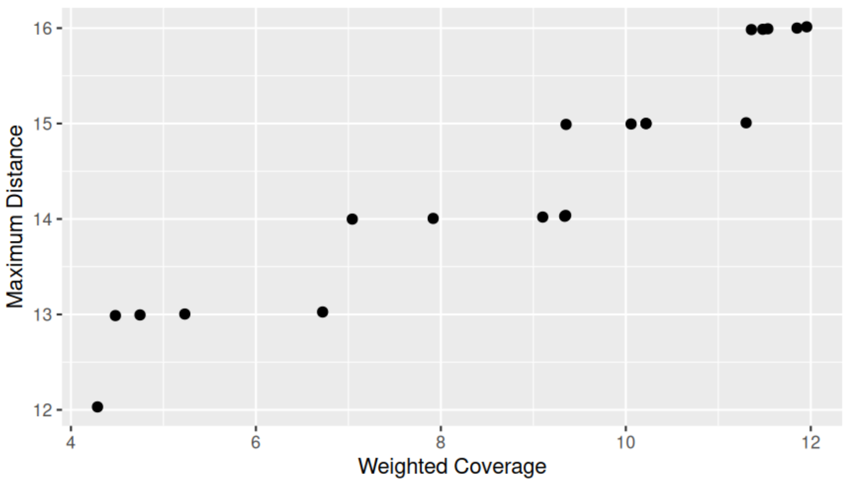



Figure C.1: Approximated Pareto front found by NSGA-II when placing three stations.

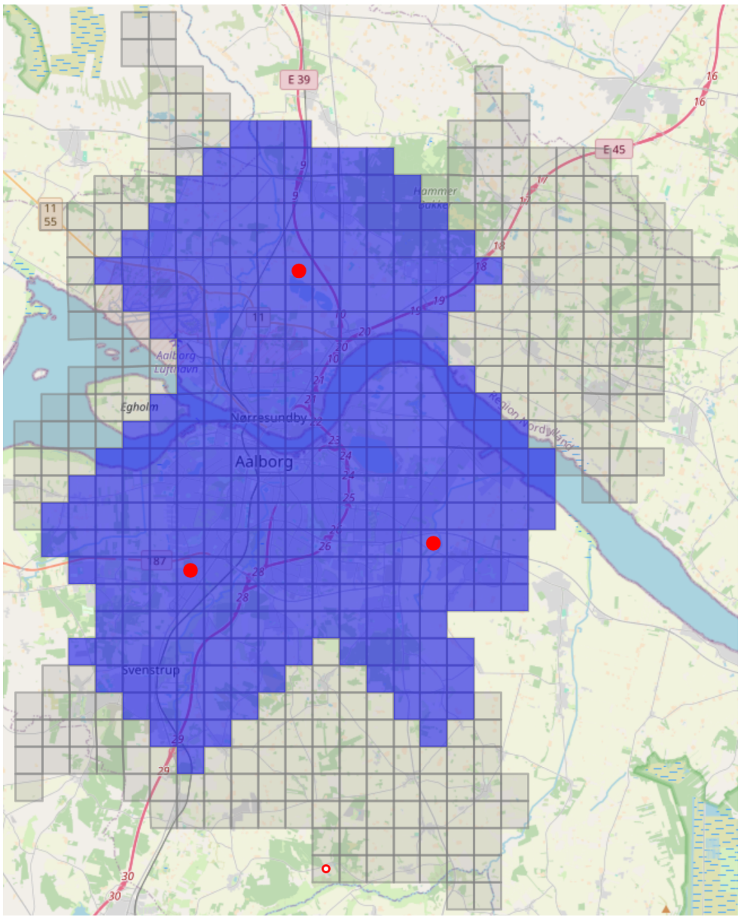



Figure C.2: Example of a solution with three stations placed.

## C.2 Eight Stations

Figure C.3 shows the situation where eight fire stations have to be located in order to optimize the previously used objectives. It is the only Pareto optimal solution in this case, since it covers the entire area and it therefore reduces to a single objective problem. Indeed, it can be seen that all of the demand squares are colored blue, meaning they are all covered.

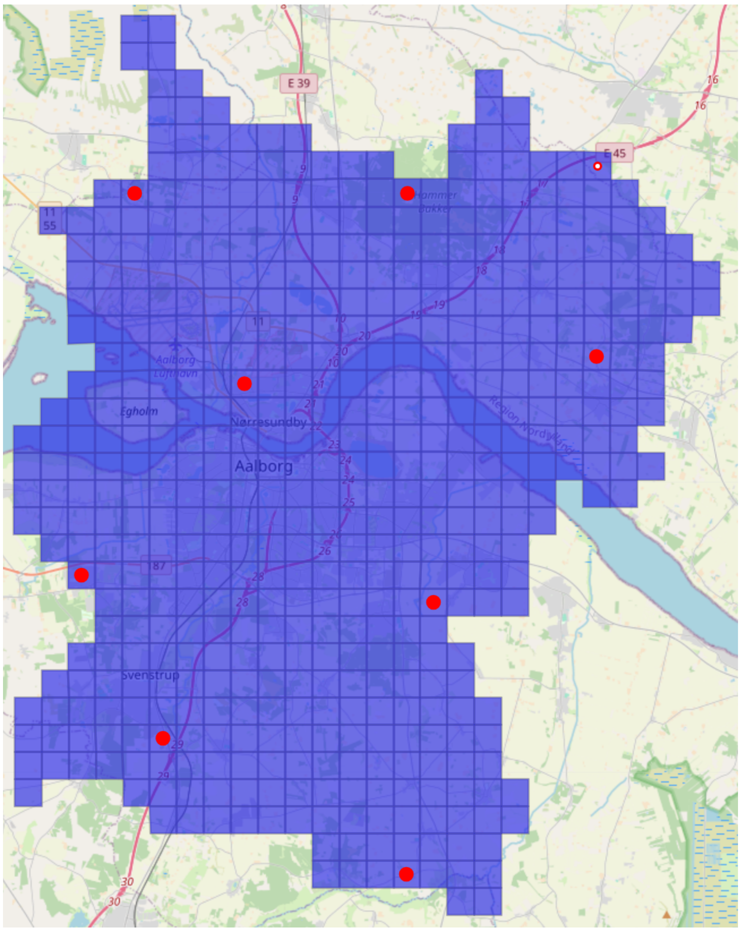



Figure C.3: Map of the solution when eight stations are placed.

# D Pareto Fronts for Other Objectives

This appendix contains graphs of the Pareto fronts for the new objectives introduced in Chapter 6. They are simply meant for comparison and will therefore not be commented further upon.

## D.1 Gini Coefficient

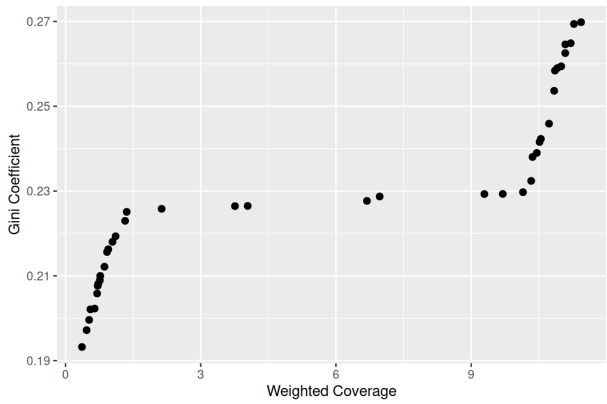



Figure D.1: Pareto front from NSGA-II optimizing weighted coverage and Gini-coefficient.

## D.2 Backup Coverage

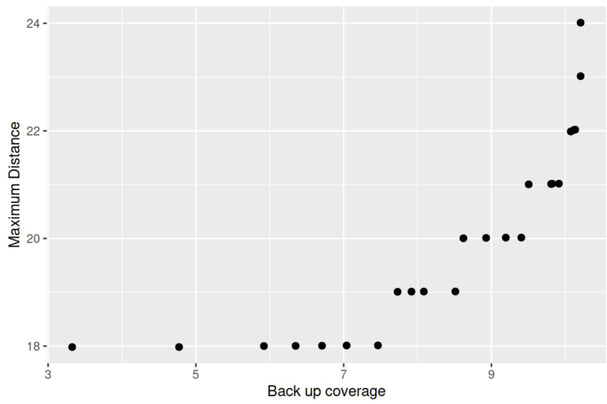



Figure D.2: Pareto frontier from NSGA-II optimizing backup coverage and p-center.