Optimizing Healthcare: Reducing Wait Times with Smart Facility Placement

Undergraduate Challenge: Operational Research A&E Modeling

Abstract

This paper addresses the optimization of patient total waiting times (TT) - sum of wait time and driving time - and healthcare facility locations. Using predictive modeling and simulation techniques, it aims to reduce delays and improve resource allocation within healthcare systems. This research employs hybrid methodologies and applies them to the given dataset, integrating patient characteristics and geographical patterns. The analysis employs classification models to predict patient-site preferences, regression models to forecast patient attendance, and optimization algorithms to determine the best locations for new healthcare facilities. The effectiveness of these methods is evaluated using both heuristic and empirical approaches to minimize total wait times across multiple sites. Additionally, the study compares the impact of different optimization techniques on reducing patient inflow to overburdened facilities. Results indicate that strategic placement of new facilities can significantly alleviate patient wait times, with reductions up to 22.8%.

Camilla Andreozzi & Glauco Rampone
University of Glasgow

1 Introduction

Healthcare systems worldwide face significant challenges with long wait times for medical services, often detrimental to patient health and satisfaction (Willcox et al., 2007; Siciliani & Hurst, 2004). Morales et al. (2024) introduced a hybrid strategy combining predictive modeling and simulation to identify bottlenecks and optimize resource allocation. Their methodology integrates regression models with simulation to enhance decision-making and operational efficiency. Predictive tools, including multiple linear regression (MLR) and artificial neural networks (ANN), provide critical insights into system performance and enable proactive resource management. These models, particularly when combined with simulation frameworks, have proven effective in reducing delays without compromising care quality (Morales et al., 2024). This enables decision-makers to evaluate multiple scenarios and implement solutions dynamically. Other than studying the mechanisms behind patient wait time, locating healthcare facilities optimally is crucial to ensure accessibility and efficiency. Salami et al. (2023) used genetic algorithms in a two-stage optimization approach to solve healthcare facility location-allocation problems. This method evaluates criteria such as population density, facility capacity, and service type to determine the best location for maximizing utility.
Monitoring injuries is essential for recognizing areas where focused prevention efforts are needed. The geographic patterns of injuries helps align prevention strategies with the populations that would benefit the most. However, since precise data on where injuries occur is often unavailable, the patient's residence is often the only known location. Haas et al. (2015) states that there is a significant connection between injury site and residential area of provenience, therefore making the proxy between the two appropriate. This allows us to correctly state mostly of the assumptions used in this research. This paper is focused on the combination of both techniques: modeling and understanding patterns in patient wait time and finding the optimal location for a new site to reduce the total time overall. To abbreviate, total wait time - the interest of this paper - will be referred to interchangeably with "TT".

1.1 Objectives

This paper aims to address the causes of patient TT and healthcare facility optimization through the following objectives:

1. **Analyze Patient Wait Times**: Use predictive modeling techniques to model TT and statistical resources to identify discrepancies in Sites and Patients in healthcare systems delay.
2. **Model the preferences of Patient-Site**: Use Classification modelling to predict, based on the characteristics of a patient, which Site they will flow into.
3. **Model the Number of Attendances by Site**: Use Regression well-established models to predict the number of attendances in each site, assuming relevance in the distance of the patient.
4. **Optimize the Best New Site (Model Assumptions)**: Apply optimization algorithms to determine the optimal locations for new healthcare facilities, aiming at maximising the relief in other sites, using the regression models built for the attendances.
5. **Optimize the Best New Site (Heuristic Approach)**: Through building the assumptions and the framework in a empirical manner, use optimization techniques and Voronoi mapping to allocate the best new site via simulation.
6. **Simulate Impact of New Facility Locations**: Use a combination of regression models and spatial optimization techniques to predict the effects of introducing new healthcare sites on existing facilities' patient volumes and overall wait times.
7. **State the Best Solution and its Advantages**: Comparing the results between the two types of optimisation applied, find the best solution that will alleviate the most wait time in other sites.

2 Data and its Manipulation

The original dataset included information about patient groups, with each group characterized by shared attributes and an associated count representing the number of individuals in that group. To analyze the data at an individual level, we expanded the dataset by replicating each group entry according to the associated count. We transformed categorical variables representing time intervals—driving time and wait time—into numerical values. This transformation was achieved by generating random values uniformly distributed within the ranges of each category. It is relevant to point-out the misclassification of Wait Time due to the missing category between 329 and 360 minutes of wait.

3 Exploratory Analysis: Difference in TT depending on Patient Characteristics

The data reveals several key insights about waiting times (as summarised by the variable TT: Total Time) based on patient characteristics. Older patients (80+) experience the longest waiting times, with an average of 203 minutes, likely due to the increased complexity and co-morbidities often associated with this age group. In contrast, younger patients (20-39) tend to wait the least, averaging 161 minutes, as they may have less complex cases. Unplanned visits, especially emergency cases, take longer, averaging 177 minutes, in contrast with the urgent nature of these visits, whereas planned return visits average only 131 minutes. Variations in waiting times are also observed across patient locations, with certain areas like Site Code 3 having longer waits (184 minutes) and others, like Site Code 13, showing shorter waits (149 minutes). Site type also plays a role, with EDs requiring more time to treat severe cases (184 minutes) compared to Minor Injury Units (MIU/OTHER), which have much faster turnaround times (82.8 minutes). The variation in waiting times based on Pat_Loc_GPs can be explained by the number of General Practitioners (GPs) in a patient's location. For example, locations with more GPs (13) experience shorter average wait times (149 minutes), possibly due less attendances because of a more efficient presence of GPs. In contrast, locations with fewer doctors (3) have longer average wait times (184 minutes), likely due to limited resources to handle patient volume, which can lead to delays. The profile of patients who wait the longest includes older individuals, unplanned visits, and those from specific locations with higher acuity, while those who wait the least are younger, have planned visits, or come from sites handling less urgent cases. Operationally, improvements should focus on streamlining processes for older patients, enhancing triage for unplanned visits and addressing infrastructure or staffing issues at high-wait locations. The results are summarised in Table 1.

Least Advantaged Group	**Average TT**	**More Advantaged Group**	**Average TT**	**Variable**
80+	203	20-39	161	Age_Group
New - unplanned	177	Return - planned	131	Attendance_Type
3	184	13	149	Pat_Loc_GPs
28257	214	38411	143	Pat_X
112966	214	109297	143	Pat_Y
80	190	30	155	Site_Loc_GPs
ED	184	MIU/OTHER	82.8	Site_Type
330000	189	130000	155	Site_Loc_GP_List
2	184	7	167	Month
2	180	4	174	Year
11	190	3	82.9	Site_Code

Table 1: Average TT by Variable, where only the most and least advantages categories for each factors are displayed

4 Wait Time Modelling

The goal of this task was to predict the wait time for a new patient using the features provided in the dataset. The models employed were Linear Regression, Gradient Boosting Regressor and Neural Networks. The performance of these models was evaluated using two key regression metrics:

- **Mean Absolute Error (MAE)**: This metric represents the average of the absolute errors between the predicted and true values. A lower MAE indicates a better fit, as it shows the average magnitude of errors without considering their direction.
- **Root Mean Squared Error (RMSE)**: RMSE gives the square root of the average of the squared differences between predicted and actual values. It penalizes larger errors more than MAE, which makes it more sensitive to outliers.

4.1 Results

As shown in Table (2), the performance of all three models is suboptimal, with only minor differences in their error metrics. The neural network model, while slightly outperforming the other two in both MAE and RMSE, still produces errors that are too large to be considered accurate. The gradient boosting regressor follows closely behind, while linear regression exhibits the highest error values.
These results suggest that predicting wait times with the available features from the dataset is not feasible. A perfect model would ideally produce predictions that exactly match the true values, resulting in a straight line when comparing predicted values to actual values. The scatter plots for each model display the degree of deviation from the ideal line, where the spread of points indicates the extent of errors. As seen in Figure 1, the models show significant deviations. In conclusion, modeling the waiting time with the current data set is not possible using the selected approaches. The results suggest that additional features or entirely different methodologies are needed. These alternative approaches will be discussed in the following sections.

Model	MAE	RMSE
Linear Regression	74.07	93.30
Gradient Boosting Regressor	73.21	92.38
Neural Network	72.21	91.57

Table 2: Performance metrics for Wait Time prediction models.

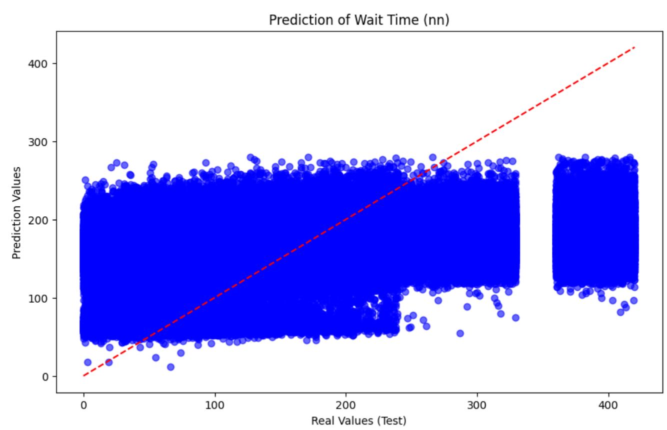

Figure 1: Predicted vs Actual Wait Time for different models.

5 Statistical Comparison of Patient Postcodes and Site Frequencies

The analysis focuses on understanding patient TT and its variability cause by variation in Postcodes and Site Codes. The goal is to explore the significant differences in total time across different locations and identify patterns that could help optimize patient wait times and overall service delivery. The methodology was to apply two different ANOVA tests on two different hypotheses as shown in Table 3.

	Postcode Analysis	**Site Analysis**
H_0	$\mu_{\text{Postcode 1}} = \mu_{\text{Postcode 2}} = \cdots = \mu_{\text{Postcode 90}}$	$\mu_{\text{Site 1}} = \mu_{\text{Site 2}} = \cdots = \mu_{\text{Site 11}}$
H_1	At least one of the Postcodes has a different mean TT	At least one of the Sites has a different mean TT

Table 3: Hypotheses for Postcode and Site Analysis

5.1 Results

The result of the ANOVA test shows that the factor Postcode has a very significant impact on TT. The analysis provides a p-value ¡2e-16, which suggests a highly significant result. This indicates that different postcodes are likely

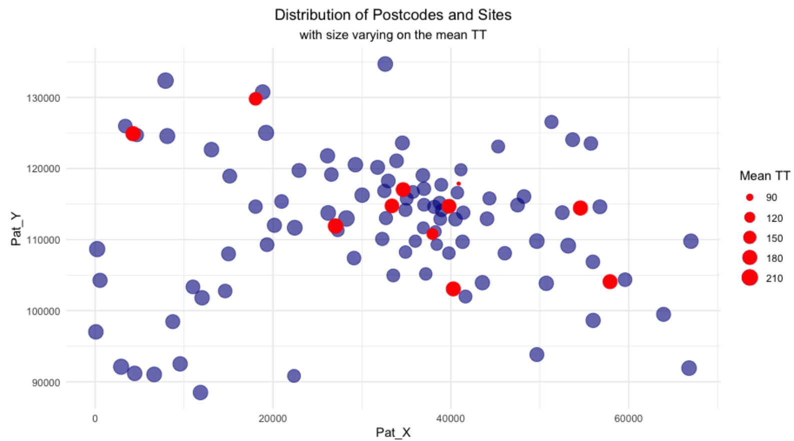

Figure 2: Spatial display of all patient and site locations recorded, with size of the point depending on the mean TT for patient postcode or site.

associated with different driving and wait times, potentially due to factors like distance, accessibility, or local service efficiency. The analysis also identifies the top 5 and bottom 5 postcodes based on their average wait time. The top postcodes have significantly higher average wait times, with values ranging from 205 minutes to 180 minutes, while the bottom postcodes have much lower average wait times, ranging from 131 minutes to 138 minutes.
The ANOVA summary for the Site_Code variable tests whether the mean TT differs significantly across the different Site_Codes. The P-value for the ANOVA test is reported as ¡2e-16, which is extremely small, this allows us to reject the null hypothesis. Therefore, there is significant evidence to suggest that the mean total times differ across the different site codes. We can identify the top 5 and bottom 5 sites based on their average total times. The top sites tend to have higher average total times, with site codes 11, 10, and 1 showing averages between 187 and 190 minutes. On the other hand, the bottom sites have much lower average total times, with site codes 3, 4, and 9 averaging between 83 and 160 minutes. The results from this analysis are summarised in table 1, 2, 3 & 4.

Site Code	Avg. TT
11	190
10	187
1	187

Table 4: Top Sites by Average TT

Site Code	Avg. TT
3	83.0
4	134
9	160

Table 5: Worst Sites by Average TT

Pat_X	Avg. TT
28257	205
26202	188
22445	187

Table 6: Top Pat_X by Average TT

Pat_X	Avg. TT
38411	131
22362	132
41130	137

Table 7: Worst Pat_X by Average TT

6 Distributional Approach to Modeling Site Preferences and Attendances

For this part of the research, the focus was to model the preferences when choosing a Site. This will not only explain the disparities between number of attendances depending on Site, but also help figure out the magnitude of the reduction in attendances in other sites, once a new site is added. The broader goal is split into two smaller objectives: predicting the likelihood of a patient visiting a specific site (using Site_Code as response variable), predicting the number of attendances in a specific Site, Month and Year for a patient from a given location (number of attendances as response variable). The second objective goes further, as we want to exploit the newly-found predictive power to find a new optimal location and assess its impact. The different use of response variables requires two different pivoted versions of the dataset.

6.1 Methodology

The methodology for this stage was to group the dataset depending on: Site_Code, Site_Type, Patient Location, month and year, to account for site variability but also any possible temporal variability. The number of attendances depending on these factors was recorded as a simplified version with respect to the one provided by the original dataset, that was further ramified depending on age group and type of attendance. For the purpose of this section, we are not interested in age differences or the nature of the attendance. A first exploratory analysis was conducted by fitting a primitive linear mixed models with site code as fixed factors and month and year as random effects. Once the results on the predictors' impact were recorded, the two objectives were pursued.

6.1.1 Objective 1: Predicting the Likelihood of a patient visiting using a Random Forest

The dataset used in this phase is the extended dataset (see Section 2). The latter was first divided in training and test set with a 70/30 split. The train set was used to train a classification random forest model with Site_Code as the factor response variable and patient location, patient age group, attendance type, site type, month and year. Accuracy metrics were recorded on the model.

$$\text{Site\_Code} \sim \text{Pat\_X} + \text{Pat\_Y} + \text{Site\_Type} + \text{Month} + \text{Year} + \text{Age\_Group} + \text{Attendance\_Type} \tag{1}$$

6.1.2 Objective 2: Predicting the Number of Patients Visiting a Site with a Regression Model

For this stage, a further manipulation was applied. As distance between the patient postcode position and site position could be relevant for the analysis, a new predictor "Distance" was computed, using euclidean distance in the two-dimensional framework. Three optimization processes were conducted based on the regression models: random forest, generalized additive model (GAM), and generalized linear model (GLM). Each optimization sought to determine the optimal coordinates for a new site to minimize patient inflows to other overburdened facilities. Predictive performance metrics were visualised for the three models.

$$n \sim \text{Distance} + \text{Site\_Type} + \text{Month} + \text{Year} \tag{2}$$

Lastly, objective functions were built and optimised, aiming at maximising the reduction impact on other sites depending on the coordinates for a new site. The optimisation processes were three, each one built on one of the predictive models. The final optimal coordinates were visualised, along with their impact in the re-distribution of attendances.

The model-fitting and optimizations were conducted using the R Programming language. The optimization techniques used the predictions of the models to compute the impact of adding a new sites at new coordinates as follows: subtracting the sum of all patients flowing in the new sites from the total of patients in the simulation, using this new total and re-distributing it according to the previous proportion of distribution among the "old" sites (from 1 to 11).

6.2 Results

6.2.1 Objective 1

The random forest classification model effectively captured the factors influencing a patient's choice of hospital, with predictors including the patient's location coordinates (Pat_X and Pat_Y), hospital type (Site_Type), and temporal features such as Month and Year. The model achieved an overall accuracy of 75.09%, demonstrating its ability to predict hospital choices with reasonable precision. The model revealed strong associations between a patient's location and their chosen hospital, highlighting the importance of geographical accessibility in hospital choice. This results guarantees the importance of fitting a new Site (Objective 2) using "Distance" as a predictor. Temporal trends further emphasized that hospital attendance patterns are not static but vary depending on the time of year and potentially other external factors. Some hospitals, such as **Site Code 11**, served a disproportionately large number of patients, indicating their role as central hubs within the network. In contrast, other facilities, such as **Site Code 9**, was less frequently chosen, likely reflecting its Major Injury Unit nature (as opposite to Emergency Department). Despite its overall strong performance, the model struggled with under-represented classes. For example, Site Code 6 exhibited lower sensitivity, indicating that the model had difficulty identifying patients choosing these less frequented hospitals. Misclassification patterns also pointed to geographical and operational similarities between certain hospitals, such as those in Site Codes 4 and 8, which were frequently confused with one another. This suggests that hospitals in close proximity or with overlapping services might be perceived as interchangeable by patients, further supporting the vitality of spatial proximity. The model can be employed to predict patient inflow to specific hospitals for emergency simulation purposes, enabling healthcare administrators to allocate resources such as staff and equipment more effectively. Furthermore, the insights can guide strategic planning for the establishment of new hospitals, particularly in underserved areas, to minimize patient travel distances and alleviate overcrowding at existing sites. Additionally, the model could serve as a valuable tool for assessing the impact of policy changes, such as improvements in transportation infrastructure or expansions in hospital services, on attendance patterns.

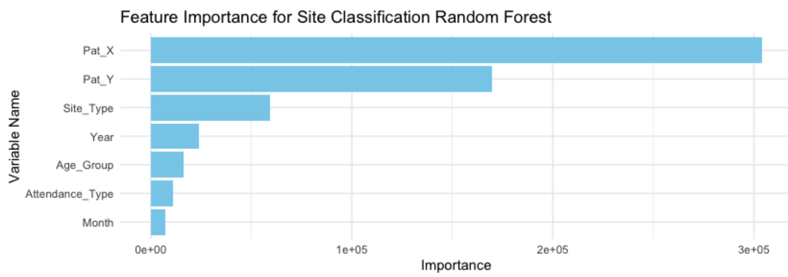

Figure 3: Variable Importance for Objective 1 Random Forest model

Statistic	Value
Accuracy	0.7642
95% CI	(0.7498, 0.7521)
No Information Rate	0.2479
P-Value [Accuracy less NIR]	2.2e-16
Kappa	0.7119

Table 8: Overall Statistics

6.2.2 Objective 2

The three optimization successfully managed to determine the optimal coordinates for a new site to minimize patient inflows to other overburdened facilities. The random forest-based optimization identified locations that effectively redistributed patient attendance but was limited by the model's non-parametric nature. The GAM-based optimization provided smoother and more interpretable spatial predictions, enabling a more balanced redistribution of patient volumes across sites. However, its reliance on smoother relationships limited its ability to capture abrupt changes in patient behavior. The GLM-based optimization leveraged its Poisson framework to model count data effectively, identifying locations that maximized reductions in patient volumes at overcrowded facilities while maintaining realistic constraints on patient redistribution. Overall, the three optimizations agreed on the same strategic site placement (X-coordinate = 33537; Y-coordinate = 111593), however, they differentiated in quantifying the impact of these new coordinates as shown in table 9. The new site placement can be visualised in Figure 4.

Model	Optimal Impact
GAM	0.1666
GLM	0.1694
Random Forest	0.1451

Table 9: Optimal impacts for different models.

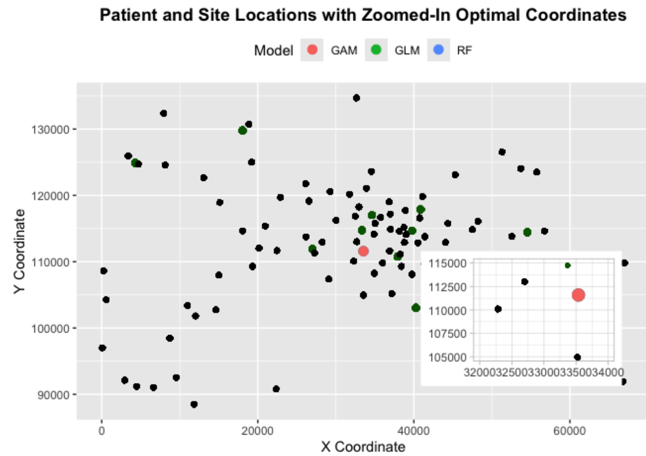

Figure 4: The optimal new site location identified via modelling

7 A new approach: empirical simulation design and optimization

7.1 Methodology

7.1.1 Data and its Manipulation

To identify the most optimal location for placing a new department, we aimed to consider both the driving time and wait time for each patient. Our goal was to minimize the overall time spent by patients from the moment they left home until they were assisted at the department. A patient with a very low TT did not require further assistance, whereas patients with very high TTs were prioritized. However, we also wanted to account for the relative number of patients in need. A high number of patients, each with a medium TT, concentrated in a specific region had to be prioritized over just one patient with a high TT, even if their TT was higher than the individual TTs of others. For this reason, for each postcode area, we sought a measure of the total need of the patients in the area. This could be represented by their number and their individual TT. To achieve this, we aggregated the individual totals based on geographic regions defined by the patients' locations. For each unique location, we computed the TT for all associated patients, producing a summary that highlighted the spatial distribution of time metrics. As a result, each postcode was associated with a priority level based on the aggregated TTs of patients within its area.

7.1.2 Voronoi Mapping

We encountered a limitation in the dataset: while we had information about the patients' locations, this data only referred to the central point of each postcode, not the entire area. For privacy reasons, we were only provided with point data for each postcode location. To model the spatial distribution more effectively, we decided to spread the patients uniformly across the entire postcode area. To do this, we constructed a Voronoi map, which divided the space into regions, with each region corresponding to the area closest to a specific postcode location according to Euclidean distance. The boundaries of these regions were clipped using a bounding box to ensure the map remained finite and meaningful within the area of interest. Each region in the Voronoi map was represented as a polygon, as showed in the left-hand image in Figure 5.

7.2 Simulating the Demand for Assistance

To identify the optimal location for a new department, we used a simulation approach to represent the demand for assistance across different geographic regions. The simulation began by quantifying the relative need for support in each region. We simulated 10,000 points and assigned them to different postcode areas approximated by the Voronoi map, based on the aggregated TT associated with each postcode. Regions with higher TTs were assigned a greater number of points to reflect their higher levels of unmet need. For each geographic area, a function generated random points uniformly distributed within the boundaries of the corresponding Voronoi polygon (see the right-hand image in Figure 5). Before generating these points, we addressed situations where certain regions, due to their spatial arrangement, fell outside the predefined Voronoi polygons. These regions were clipped to stay within a reasonable extent and not extend to infinity. Their attributes were stored separately. For these regions, we generated points directly at their central locations (as provided in the dataset), and these points were later incorporated into the simulation dataset to ensure their demand was adequately represented. By focusing on areas with a higher density of points, we were able to identify potential locations for new departments that effectively addressed unmet needs and reduced waiting times in underserved regions. Higher density indicated higher need.

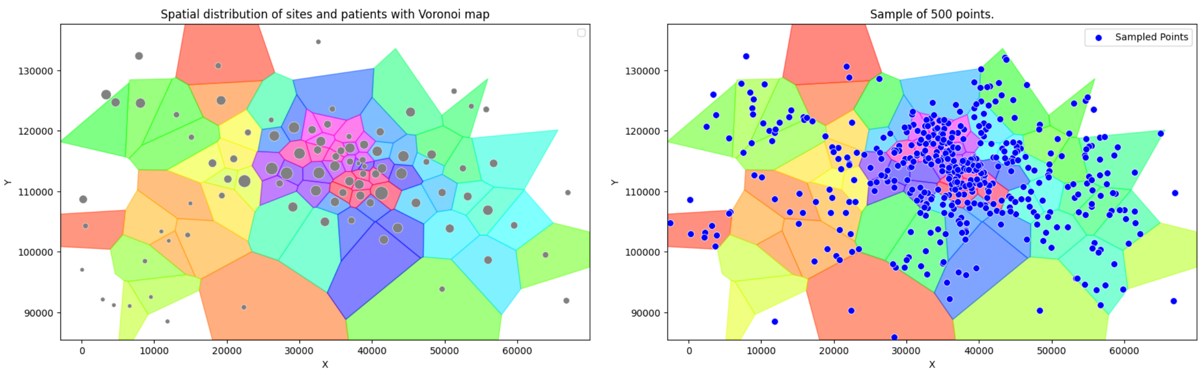

Figure 5: On the left the plot of the Voronoi Map. On the right a sample (n=500) of the generated points.

7.3 Addressing for the presence of the existing departments

A significant concentration of points appeared in the center of the area, as expected. Since most existing departments were located in the center, we needed to account for the presence of these existing service sites. We worked under the assumption that each patient always chose the nearest department. This assumption implied that a new department would have no effect if it were located too close to an existing one. Therefore, our task was to find the optimal location that minimized the TT, assuming only the closest patients would choose the new location. To tackle this, we introduced an optimization process to evaluate and minimize the total Euclidean distance between the generated demand points and the existing service sites. The analysis began by calculating the distances between each demand point and all the unique service locations, assigning each point to its nearest site. The cumulative distance for each site was then calculated by adding the distances of all assigned points and the total total distance was determined at all sites. The goal was to minimize this total distance.

7.4 Simulating the Introduction of a New Service Site

We simulated the introduction of a hypothetical new location. We recalculated the total distance, considering the proximity of the demand points to all existing sites as well as the new site. By defining a grid of 40,000 candidate coordinates, the algorithm iterated over each point, calculating the total Euclidean distance between generated demand points and all service sites, including the candidate site. This exhaustive search ensured that the location minimizing the overall distance was identified as the optimal new site, with coordinates X = 33537 and Y = 111593. Figure 6 shows the new best site location after the simulation.

8 Comparison and Discussion

For the comparison phase, plausible attendances for the new sites coordinates were recorded for all the three models and the simulation framework and used to predict the impact the new site would have on the overall distribution of attendances under the different assumptions. The best new site computed in Section 6 (with coordinates X = 33537, Y = 111593), accounts for a reduction in attendances in the other Sites of 306926 patients that would prefer attending the new site (Site_Code : 12) during the time accounted by the dataset. This reduces the average TT in other sites of almost 30 minutes (reduction of 16.9%). The metrics for this redistribution are computed assuming that the new site will ease other sites of attendances in a proportionate manner. Meaning that the proportion of reduction in attendances is the same as the distribution of attendances before the addition of the new-site. Then, a new average TT is computed on the old Sites but with the new total attendances and compared to the old average TT. Viceversa, the best new site computed in Section 7 (with coordinates X = 10640, Y = 105691) would reduce attendances in other Sites of 96079 patients, for the months/years accounted by the dataset. This amounts to an

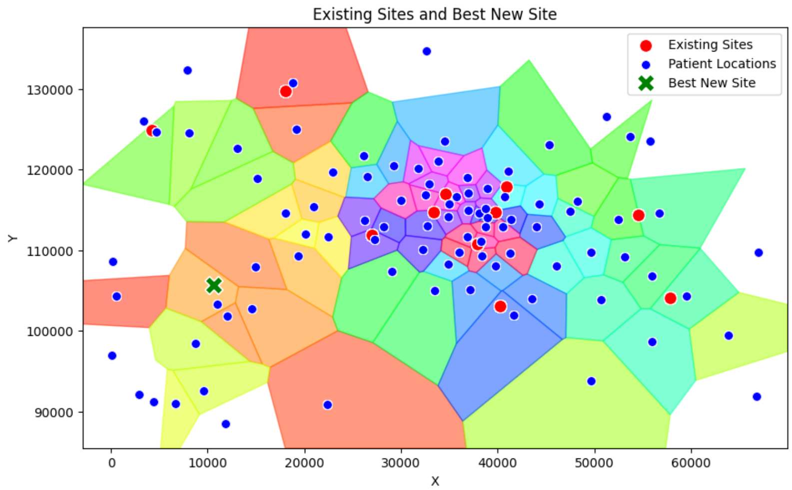

Figure 6: The optimal new site location identified through the simulation.

impact in reducing average TT of 9.35 minutes (5.3%). To further evaluate the efficacy of these new sites, their performance was tested using the advanced model described in Section 7. Under this framework, the peripheral site at X = 10640 and Y = 105691, demonstrated a more pronounced effect. Specifically, the new site reduced the average TT by 13 minutes (7.12%), alleviating attendance pressures significantly across all departments in the surrounding regions. We also explored the possibility of combining the power of both new sites. Under the framework defined in Section 7, the computed new average TT in the 11 Sites with both new sites present would be of 136 minutes, a reduction of ~ 40 minutes from the original average TT. The impact of the strategy would be a percentage reduction of TT in Sites from 1 to 11 of 22.8%.

8.1 Limitations

Limitations in this study lay in the assumptions holding both of the simulated New Site frameworks. First, the spatial features of the models are only approximated due to the residence postcode being the only location information on the patient. However, studies, such as Haas et al. (2015), argue that this approximation is enough close to the location of injury, therefore the origin of the patient's visit to hospital sites. Moreover, some uncertainty in the results is added by the data manipulation using the uniform distribution assumption to allow for numerical variables. Second, the predictive models used in Section 6 are held by assumptions such as independence between observations which may be inaccurate. Also, the redistribution of patient after adding the new sites assumes that the patients will proportionally abandon the old sites to choose the new site. This may oversimplify complex patient decision-making behaviors influenced by factors like hospital reputation, service quality, or specialized care availability. Third, the approach detailed in Section 7 makes several assumptions. The model assumes that each patient always travels to the closest department. However, when this assumption is tested on the dataset, it does not consistently hold true. Additionally, the simulation of patient locations presumes a uniform distribution of individuals across the entire region. Another limitation arises from the use of Voronoi mapping to delineate regions. The regions generated through this method do not correspond to real postcode boundaries. Achieving more accurate results would require detailed knowledge of the specific territory under study. Lastly, it exclusively considers the distance and wait time for each patient while ignoring other potentially significant factors.

9 Conclusions

The research conducted provided a comprehensive understanding of the data and an hypothetical framework for real-word results on the total time until service of patients. The exploratory analysis focused on the differences in travel time (TT) across healthcare facilities and patient groups. Patients living closer to densely populated urban centers generally experienced shorter travel times due to the proximity of multiple healthcare sites, whereas those in rural or remote areas faced considerably longer journeys. This discrepancy highlighted the uneven accessibility of healthcare services, with rural patients often relying on fewer and more distant facilities. The analysis also identified differences in travel time across patient demographics, such as age groups. The statistical comparison of patient postcodes and site frequencies revealed that the distribution of healthcare visits was highly concentrated around urban centers, where site densities were higher and travel times shorter. Statistical tests confirmed significant differences in visit frequencies across postcodes. These findings emphasize the importance of understanding the interplay between geographic factors, forming the foundation for the targeted intervention proposed, aimed at addressing travel-related inequities. Therefore, this justifies the optimization of a new healthcare site placement. The predictive model constructed successfully to model the preferences of a patient, based on one's characteristics, in choosing a site to be assisted, revealed the importance of the proximity between the patient's postcode and the distance with the Site. This laid the base for the first optimization. After the number of attendances were modelled between sites with three different models for comparability, accounting for temporal variability and distance of the patients, a new site was found maximising the total number of patients flowing into the new coordinates. A new approach was also proposed, where the emergency attending framework was better designed according to reasonable assumptions. This heuristic model highlighted areas (creating a Voronoi map of the postcodes) where accessibility challenges are most pronounced, specifically aiming at easing the overall distribution of patients, not only the crowded urban area. This was done simulating the demand for assistance proportionally to the average TT computed on each region. Finally, accounting for the over-representation of the central area of the dataset, the new site was found assuming that patients will prefer the closest location available, by minimizing the total distance between patient and new site. The final proposed solution is that both departments are to be implemented. This would balance-out the challenges deriving from each department. The crowded urban central area would be relieved by the addition of the Section 6 department, while the Section 7 department would still account for the other high-demand peripheral areas. Both sites take into account the importance of spatial proximity to patients and promote a resilient healthcare system where the issue of waiting times is largely mitigated (by 22.8% with the combined solution).

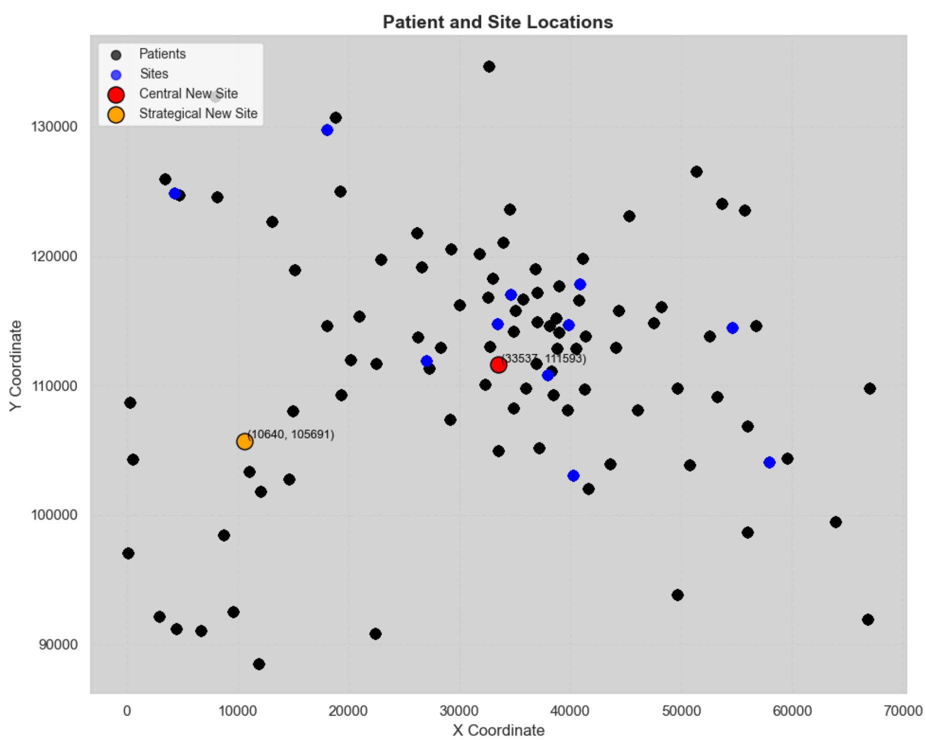

Figure 7: Plot of Patient and Site locations with the addition of the new sites proposed.

10 References

Haas B, Doumouras AG, Gomez D, de Mestral C, Boyes DM, Morrison L, Nathens AB (2015). *Close to home: an analysis of the relationship between location of residence and location of injury.* J Trauma Acute Care Surg, 78(4):860-5. PMID: 25807410; PMCID: PMC4375775.

Morales, J., **Silva-Aravena, F.**, & **Saez, P.** (2024). *Reducing Waiting Times to Improve Patient Satisfaction: A Hybrid Strategy for Decision Support Management.* Mathematics, 12(23), 3743. `https://doi.org/10.3390/math12233743`

Salami, A., **Afshar-Nadjafi, B.**, & **Amiri, M.** (2023). *A Two-Stage Optimization Approach for Healthcare Facility Location-Allocation Problems With Service Delivering Based on Genetic Algorithm.* International Journal of Public Health, 68, 1605015. PMCID: PMC10011119.

Siciliani L., **Hurst J.** (2004). *Explaining Waiting-Time Variations for Elective Surgery across OECD Countries.* OECD Economic Studies, 38(1):95–123.

Willcox S., **Seddon M.**, **Dunn S.**, **Tudor Edwards R.**, **Pearse J.**, **Tu J.V.** (2007). *Measuring and Reducing Wait Times: A Cross-National Comparison of Strategies.* Health Affairs (Millwood), 26(4):1078–87. doi: 10.1377/hlthaff.26.4.1078.