Dynamic Ambulance Redeployment with uncertain driving times

M.S.A.A. Hofmeijer 361171

July 8, 2016

ERASMUS UNIVERSITY ROTTERDAM
Bachelor thesis Econometrie en Operationele Research

Supervisor: R.B.O. Kerkkamp MSc.
Second assessor: Dr. W. van den Heuvel.

Abstract

For emergency services it is crucial to have an ambulance nearby and ready for almost all accidents. The scheduling this requires is what this thesis is about. There have been many approaches to scheduling ambulances efficiently based on coverage concepts. This thesis aims to modify some previous approaches to include uncertainty in travel time and to be more realistic.

Contents

1 Introduction

In the Netherlands responsibility for ambulance care is divided between a number of regions which cooperate but are in principle independent of each other, these regions are called Regionale Ambulance Voorzieningen (RAVs). In these RAVs ambulances need to be scheduled so that ambulances have a good starting point to start their next journey to a new accident after they have finished with their presently assigned accident.

There are different types of accidents that have different priorities. The highest priority accident is A1, for which the maximum allowed reaction time is 15 minutes. This will be the type of accidents this thesis is interested in. There are however two other types of accidents namely A2 and B, which are less urgent than A1. It is important to note that 15 minutes is not the maximum driving time, it is the maximum time between the call and the arrival of the ambulance. This means the driving time can be at most about 12 minutes because approximately 3 minutes are spent taking and processing the call to the emergency service .

2 Problem description

The main problem is to schedule ambulances efficiently, specifically ambulances that need to be on the scene of an accident in little time, preferably in 15 minutes from the time that the call came in. The ambulances start at a base location, drive to an accident, and then choose to either bring the patient(s) to the hospital or not. After they are done with this task, the ambulances need to be assigned to their next base location. This problem is hard in the sense that time is continuous; this means that there is an uncountable number of feasible solutions. Furthermore it matters how many discrete locations are chosen to represent the real world and how these are chosen from a very large number of possible locations. The problem that we are facing is that ambulances need to react quickly to events, but then have to decide where to go from there, without knowing what will happen next. Another issue is that scheduling ambulances preferably needs to happen in real time, so that the emergency services can react to developing situations. This means that the solution method we choose for this problem needs to not only find a good solution, it also has to do it in manageable time.

3 Methodology

Our general approach to solve this problem is to formulate a policy to send an ambulance to a base location after it has become idle. We then evaluate our policy by calculating a service level. We calculate this service level by running a Discrete Event Simulation (DES) where we register how many ambulances arrive too late at an accident as a fraction of the total number of accidents.

3.1 Mathematical formulation of the problem

To model this problem mathematically, we define A as the set of ambulances, and V as the set of the locations where demand for ambulances can occur. Let $W \subseteq V$ be the set of base locations and let $H \subseteq V$ be the set of hospitals. Let $d_i \in \mathbb{R}_{\geq 0}$ be the fraction of the total demand that occurs at $i \in V$. Note that this means that the demand area distribution is represented as a limited number of discrete points. An ambulance is idle if it returns to a base location or is already at one. Furthermore, an ambulance can only be sent to an accident when it is idle. If and when an accident occurs, we require the idle ambulance that is closest in terms of time to respond to the accident. Note that accidents that occur when no ambulances are idle are put in a first-come first-serve queue. The accidents in the queue will have an ambulance assigned to them as soon as an ambulance is available. Location travel times $t_{ij} \in \mathbb{R}_{\geq 0}, i, j \in V$ are deterministic. When an ambulance arrives at the scene of the accident it is determined how long treatment at the accident $t_{accident} \in \mathbb{R}_{\geq 0}$ will last. After this the need for transportation of the patient to a hospital is determined. When this decision has been made, the ambulance takes the patient to the nearest hospital and waits at the hospital for a certain time $t_{hospital} \in \mathbb{R}_{\geq 0}$. Given this, we minimize the number of accidents that are not reached before a certain threshold time $T \in \mathbb{R}_{\geq 0}$.

3.2 The MEXCLP model

The Maximum Expected Coverage Location Problem (MEXCLP) is a Mixed Integer Problem (MIP) and models how a number of facilities at a number of locations can cover a certain area. We can use this to model our ambulances, their bases and the accidents as is done in [11]. However, the MEXCLP formulation does not incorporate continuous time, so the optimal solution of this problem will only be used as a benchmark for other methods. Nevertheless it is a valuable resource in that capacity. There have been numerous variations of this problem with notable examples being [2, 14, 9]. There have also been various definitions of the coverage concept [4, 16, 15, 3]. We use the following formulation as our main static model which is similar to the formulation in [4] and is the same formulation used in [11]:

$$\max \sum_{i \in V} \sum_{k=1}^{p} d_i (1-q) q^{k-1} y_{ik}$$

$$\begin{aligned}
\sum_{j \in W} x_j &\geq \sum_{k=1}^{p} y_{ik}, && i \in V, \\
\sum_{j \in W} x_j &\leq |A|, && i \in V, \\
x_j &\in \mathbb{N}, && j \in W, \\
y_{ik} &\in [0,1], && j \in W, k = 1, \ldots, p.
\end{aligned}$$

With $q \in [0,1]$ being a constant busy fraction, meaning the proportion of time an ambulance is busy. The variable $y_{ik} \in \mathbb{B}$ is one if point i is in range of at least k vehicles and zero otherwise. The variable $x_j \in \mathbb{N}$ is the number of vehicles that are at base j. Notice that this model assumes that if an ambulance covers a demand location it either covers it or it does not: there is no partial coverage.

3.3 The Dynamic MEXCLP model

Our other model for the assignment problem is dynamic, but uses the same coverage concept as MEXCLP. A solution is found for this model every time an ambulance becomes idle. This solution is the best base location for that ambulance to go to. It can be considered as a MEXCLP where we ignore all busy ambulances and fix all already idle ambulances in their place. There are also other approaches to modeling of this problem, namely Approximate Dynamic Programming (ADP) models such as in [12]. However these lead to drivers not knowing why they should move. Knowing that is one of the benefits of the Dynamic MEXCLP model. Other alternatives are tabu search algorithms as used in [6], which could be used in combination with some other methods, but are not used by us because they are out of the scope of this thesis. We use Algorithm 1 which is the same algorithm as in [11] to determine the best base to station an ambulance using Dynamic MEXCLP.

The algorithm we use to solve the Dynamic MEXCLP is a somewhat brute force based algorithm, but it is nevertheless fast enough to run every time an ambulance becomes idle. There are other methods possible in this framework namely lookup tables [1] where the problem is solved for many different states so that when an answer is required, all that is needed is looking up the answer. The drawback however is that for our problem there are many states and exact optima would be hard to find, so that it will be difficult to calculate all the answers needed within a reasonable time.

Algorithm 1 Dynamic MEXCLP

```
function Dynamic MEXCLP(ambulances, eventlist, nextevent, instance)
    initialize BestImprovement, BestLocation
    check which ambulances are in idleambulances
    for j ∈ W do
        CoverageImprovement=0
        for i ∈ V do
            k=0
            if Traveltime_{ji} ≤ maxresponsetime then
                k=k+1
                for a ∈ idleambulances do
                    if Traveltime_{dest(a)i} ≤ maxresponsetime then
                        k=k+1
                    end if
                end for
                CoverageImprovement+ = d_i(1 − q)q^{k−1}
            end if
        end for
        if CoverageImprovement > BestImprovement then
            BestLocation=j
            BestImprovement=CoverageImprovement
        end if
    end for
return BestLocation
end function
```

3.4 The simulation model

We use a simulation model to evaluate our policies. The simulation model as previously mentioned is a DES, which means that the time is only updated when an event happens. This simulation is very similar to the simulation in [11]. There are four types of events. Firstly, there is an event for a new accident. Secondly, there is an event for an ambulance arriving at an accident. Thirdly, there is an event for an ambulance arriving at a hospital. Lastly, there is an event for making an ambulance idle and send it to its base location. After each event occurs it is checked if one of the ambulances is idle and if there is an accident that has no ambulance assigned to it. If this is the case an ambulance is dispatched to the first accident in the queue. These checks are not always necessary but does not take a long time and avoids errors. Pseudo code for the main simulation is given in Algorithm 2 and some of the secondary functions are given in Algorithms 5 to 9 which can be found in the algorithms appendix .

Algorithm 2 Main simulation algorithm

```
procedure MAIN SIMULATION(a complete instance)
    initialize ambulances, eventlist, accidents, queue,
    latearrivals, arrivals, idleambulances=true
    plan first accident
    while t ≤ maximumrunningtime do
        get nextevent
        if nextevent is of type new accident then
            run generate a new accident()
            update state of simulation
        else if nextevent is of type arrival of an ambulance at an accident then
            run arrival of an ambulance at an accident()
            update state of simulation
        else if nextevent is of type arrival of an ambulance at a hospital then
            run arrival of an ambulance at a hospital()
            update state of simulation
        else if nextevent is of type make ambulance idle then
            run make ambulance idle()
            update state of simulation (idleambulances =true)
        end if
        if size of queue ≥ 1 AND idleambulances =true then
            run dispatch ambulances()
            update state of simulation
        end if
        set idleambulances to false if appropriate
    end while
end procedure
```

Our simulation works by having an outer loop where time and which event is next are handled and two inner loops that handle dispatching ambulances and events, of which there are four types. The outer-loop is a while loop which regulates the inner-loops and makes sure they happen at the right time and in the right order. The first inner loop is an if-loop that handles all events namely:

- Firstly, there is the event to make a new accident (Algorithm 5). This event makes a new accident and stores this in the queue. It then plans the next accident.
- Secondly, there is the event to make an ambulance arrive at an accident (Algorithm 6). This event updates the eventlist, queue and ambulances. Then it determines if this was a late arrival or if the arrival was on time. After that the algorithm determines if the patient should be taken to the hospital or not and then makes an event according to its decision.
- Thirdly, an event for ambulances that arrive at the hospital(Algorithm 7) is used. This event plans a new event to make the ambulance idle.
- Lastly, Algorithm 8 is an event to make ambulances idle, it runs either Dynamic MEXCLP or gets the base location of this ambulance and sends it to the selected base, it of course also makes the ambulance idle. It should be noted that MEXCLP is only run once at the beginning of the simulation. In this run each ambulance gets assigned a home base that it returns to every time it becomes idle. However it is possible to have a simulation where MEXCLP (with some adjustments) is run every time an ambulance becomes idle, this is however very time consuming and very similar to running Dynamic MEXCLP if one does not want to move every idle ambulance.

The last part of the simulation is the part where first it is determined, if there are accidents in the queue and if there are idle ambulances. If these conditions are true it then runs an algorithm, Algorithm 9, that sends the closest idle ambulance to the accident that happened earliest.

The way we used to calculate distances in algorithm 9 for ambulances that are between bases is not completely realistic, however we have no data regarding these travel times, and had to make an approximation.

We start by calculating the Euclidean distance between the gps-coordinates of the relevant areas. Then we calculate how many percent of the time that is needed to travel from the origin to the current destination has past. We determined the coordinates of the ambulance as if it was traveling in a straight line. After this we calculate the distance between the ambulance and the accident. Lastly, we divide this last distance by the distance between the current destination and the accident to get the relative distance and then multiply it by the time it takes to get from the current destination to the accident, which gives us its new travel time. This last calculation is shown in the following formula:

$$t_{travel} = \left(\frac{d_{loc}}{d_{dest}}\right)(t_{dest}). \tag{1}$$

With t_{travel} being the travel time between the current location of the ambulance and the accident, d_{loc} being the distance between the current location and the accident, d_{dest} being the distance between the current destination and the accident, and t_{dest} being the time that is needed to travel between the current destination and the accident.

This way of calculating the travel time assumes Euclidean space. This assumption is false, but it is realistic enough for large distances. For small distances the inaccuracy is large compared to the distance but in total terms is still small.

3.5 Modification to the MEXCLP coverage concept

A major problem with our assumptions is that in real life situations travel times, are not deterministic but stochastic. Another problem is that our definition of demand locations is discrete instead of continuous. This means that not all accidents that occur in the same postal code actually have the same distance to the same base location. Both these problems can be solved to a certain degree by introducing partial coverage, meaning that instead of an ambulance covering a location in full or not at all, it can cover that location to a certain degree. The concept being that some parts of the location can be reached in time, but others cannot. Partial coverage can also simulate the chance that a location can be reached in time for the patient to have a good outcome.

We do this by adding a random number to the average travel times. The random number that is added will be normally distributed with mean zero and standard deviation $\sigma_{constant} + \sigma_{variable}$. We do not have any reliable data about the distribution of this number, but it is reasonable to assume that there is a constant part to the standard deviation caused by the imprecision of where an accident happened in a neighborhood say $\sigma_{constant}$, and a part that is dependent on the mean travel time say $\sigma_{variable}$ that represents differences in travel times (caused by open bridges for example). In most cases $\sigma_{constant}$ will be relatively small, and is less significant but still relevant. Unfortunately this can result in negative travel times. To solve this we round negative numbers up to zero, this means that for very small average travel times the average travel time will significantly differ from their mean. However because the constant variance is small this should not make too much difference.

Of course this means that y_{ik} is no longer well defined since all ambulances now cover all bases with a certain chance. This gives rather large problems for the MEXCLP formulation because we could not find a new objective function that is linear. However for the Dynamic MEXCLP we will provide a new objective function. It is interesting to note that this can also be used to represent the mortality rate instead of the fraction of late arrivals as is done in [5]. After our second modification we introduce a objective function for the whole modified model which includes both modifications. Because of this second modification we do not use the formulation in [14].

3.6 Modification to the busy fraction concept

Another possible limitation to the current model is that it uses the same busy fraction for all bases. This of course is not necessarily correct, therefore we want to look at how the fraction of time that ambulances are busy is represented in our model. We identified two ways to approach this problem. First, there is a data based method that looks at historical busy fractions and applies this data to the current situation; with this method busy fractions for a certain base location at a given time of day are always the same. Second, there is a moving average approach, which updates predictions of future busy fractions using the data of a period in the past, for example the last few hours, by calculating a moving average. We choose for a predetermined busy fraction because the calculation of busy fractions in real time tends to be biased. It is biased in the sense that the busy fractions react to random fluctuations in how many accidents there are in a certain region, which does not have any predictive value. A way to solve this problem is to take the moving average over a long period, however this is very similar to using predetermined busy fractions. This means that for our research predetermined busy fractions are sufficient, however if one is in the position that there is no reliable way of predetermining busy fractions one should use moving averages.

3.7 The modified model

Before the new objective function is presented we need to introduce some notation. Particularly let $j \in W_i$ be ordered so that $j = 1$ is the base in W_i which has the highest chance of serving i well, and $j = 2$ is the base with the second highest chance and so forth. Let $p_{i,j}$ be the chance that an ambulance from base $j \in W_i$ serves a demand location $i \in V$ adequately. Let k_j be the number of ambulances at base $j \in W_i$ and let q_j be the busy fraction of base $j \in W_i$. Then the expected coverage for one demand point i is:

$$\sum_{j \in W_i} d_i p_{i,j} (1 - (q_j)^{k_j}) \prod_{m=1}^{m<j} (q_m)^{k_m} \tag{2}$$

This means the expected coverage for the entire model can be denoted with:

$$\sum_{i \in V} \sum_{j \in W_i} d_i p_{i,j} (1 - (q_j)^{k_j}) \prod_{m=1}^{m<j} (q_m)^{k_m} \tag{3}$$

While we do not use it, it might be useful to give the restrictions that complete this model which are:

$$\begin{aligned} &\sum_{j \in W} k_j \leq |A|, && \\ &k_j \in \mathbb{N}, && j \in W. \end{aligned}$$

It should be noted that this modification is, although independently developed, very similar to what is done in [8] and is actually a special case of the extended model presented in this paper, if all bases are open. It is also mathematically identical to the model in [10] with slightly different interpretations. The following dynamic version of this model can therefore be considered to be a dynamic version of that model as well. Our method for solving our modified model is very similar to what was done for the Dynamic MEXCLP in the sense that it also solves the problem by calculating all possibilities and selects the best. This program does need more calculations and so one possible drawback is that it might for some data cases not be able to run in real time.

Algorithm 3 Adjusted Dynamic MEXCLP

function ADJUSTED DYNAMIC MEXCLP(ambulances, eventlist, queue, accidents, instance)
initialize best location (BL)
initialize current coverage (CC);
    run `Adjusted Dynamic MEXCLP current coverage()`
    initialize coverage improvement (CI)
    for $z \in W$ **do**
        initialize adjusted coverage (AC)
        for $i \in V$ **do**
            initialize specific coverage (SC);
            initialize the multiplication busy fractions (MQ)
            get W_i
            for $j \in W_i$ **do**
                get number of idle ambulances at base j ($NOIA$);
                if (**then**$z = j$)
                    $NOIA = NOIA + 1$
                end if
                get the busy fraction for this base (BF)
                initialize base specific cover (BSC)
                get chance of sufficient service for this j and i ($COSS$)
                if j is the highest priority base for i **then**
                    $BSC = COSS(1 - BF^{NOIA})$
                    $MQ = BF^{NOIA}$
                else
                    $BSC = COSS(1 - BF^{NOIA}) * MQ$
                    $MQ = MQ(BF^{NOIA})$
                end if
                $SC = SC + BSC$
            end for
            $AC = AC + SC$
        end for
        if $CC - AC > CI$ **then**
            $CI = CC - AC$
            $BL = z$
        end if
    end for
return BL
end function

Algorithm 4 Adjusted Dynamic MEXCLP current coverage

function ADJUSTED DYNAMIC MEXCLP CURRENT COVERAGE(ambulances, eventlist, queue, accidents, instance)
initialize current coverage (CC);
  for $i \in V$ **do**
    initialize specific coverage (SC);
    initialize the multiplication busy fractions (MQ)
    get W_i
    for $j \in W_i$ **do**
      get number of idle ambulances at base j ($NOIA$);
      get the busy fraction for this base (BF)
      initialize base specific cover (BSC)
      get chance of sufficient service for this j and i ($COSS$)
      if j is the highest priority base for i **then**
        $BSC = COSS(1 - BF^{NOIA})$
        $MQ = BF^{NOIA}$
      else
        $BSC = COSS(1 - BF^{NOIA})MQ$
        $MQ = MQ(BF^{NOIA})$
      end if
      $SC = SC + BSC$
    end for
    $CC = CC + SC$
  end for
return current coverage
end function

4 Results

4.1 Data

The data we gained from RIVM[7] consists of the average driving time of ambulance between locations (represented as 4-digit postal code areas) in 2008, the population of each postal code area (used as an indication for demand in that postal code), and the gps coordinates for all postal code areas used. We also have all the possible base locations for the ambulances. In addition we need the fraction of accidents that require ambulance transport, we got this information from [13], unfortunately this is not for the same year as our other data but it can be assumed to be reasonably constant at around 70%. Lastly we also need to know the location of all hospitals. This was slightly harder since this changes over time. We also see that our data set does differ from the data set in [11] because our case has less base locations but more hospitals, than that case. We use the arrival rate used in [11] which is 9.5 accidents per hour on average, and inter arrival rates which are exponentially distributed with $\lambda = 0.0026388$. We use a constant busy fraction of 0.3. We use a hospitalization chance of 0.701. We use a Weibull distribution with $\alpha = 1.5$ and $\lambda = 0.0009259$ to generate hospital treatment times. We use an exponential distribution with $\lambda = 0.0013888$ to generate treatment times at the accident. There are 19 ambulances in our model. We have 16 hospitals, 11 base locations and 217 demand locations, of which the hospitals and base locations are given in the data appendix.

4.2 Results from MEXCLP and Dynamic MEXCLP

Our first results concern our base models, MEXCLP and Dynamic MEXCLP. We summarize the results of our simulations using these two models in the next figure. To get these results we ran both algorithms twenty times. We used the IBM CPLEX Optimizer to solve the MEXCLP and Algorithm 1 to solve Dynamic MEXCLP. Figure 1 gives the mean fraction of late arrivals (as a red line) and the spread of the fraction of late arrivals. These results where obtained with the use of 20 seeds. The Random Number Generator (RNG) used is a Mersenne Twister RNG.

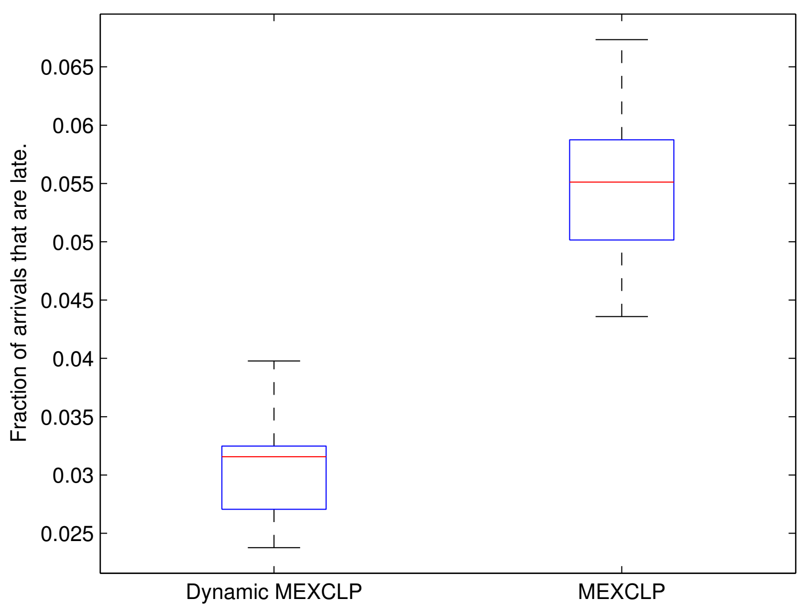

Figure 1: Fraction of arrivals that are late for Dynamic MEXCLP and MEXCLP respectively in a simulation without uncertainty in the driving times and without varying busy fractions.

The mean fraction of late arrivals is 0.0309 for Dynamic MEXCLP and 0.0544 for MEXCLP. The standard deviation is 0.0041 for Dynamic MEXCLP and 0.0064 for MEXCLP. In all cases Dynamic MEXCLP performed better, and it is clear that Dynamic MEXCLP performs better overall because of the large differences between the models results. We can see that the results are somewhat similar to the results in [11], even though the difference between Dynamic MEXCLP and MEXCLP is relatively large and service levels are higher (or the fraction of arrivals that are too late is lower). This could have various reasons, but is probably due to the different data cases. These results are however in line with the performance of this RAV which is given in [13], in this report the fraction of late arrivals for RAV Utrecht is given as 4.6% which is similar to our results.

4.3 Results from the modified model

To evaluate our modified model we use the Dynamic MEXCLP as a benchmark. We have several cases, firstly we have the default case were there is only one value for the busy fraction and no uncertainty in the travel times. Secondly, there is a case where there is uncertainty in the travel times in the model but one value for the busy fraction. Thirdly, there is a case with no uncertainty in the travel times and different busy fractions. Lastly there is a case where there is uncertainty in the travel times and different busy fractions. We only give the results for the first and the last case to show the similarities between the Dynamic MEXCLP and the modified model in a situation for which both should work approximately the same, and the improvements that can be made by the modified model. The other cases can be found in the results appendix. It should be noted that Dynamic MEXCLP does not incorporate varying busy fractions, and so when an instance has varying busy fractions, Dynamic MEXCLP uses the mean of the varying busy fractions. We will test the significance of the difference between each pair of results using non parametric tests. Using the fact that the random generators used have the same seeds we can do a one-sided paired sign test to test if one model has a significantly higher median than the other model. It is also very useful that this test makes very little assumptions, and does not assume symmetry since symmetry in this case would probably not hold. We will report the null hypothesis and the alternative hypothesis, the way we calculated the end results and finally the chance that the null hypothesis holds. All tests in the following sections will be one-sided paired sign tests.

The random travel times are generated by using a mersenne twister RNG.

The results in Figure 2 and 3 are generated by setting $\sigma_{constant}$ to 30 seconds and a $\sigma_{variable}$ of 15% of the mean travel time. The busy fractions we used are generated by running the Dynamic MEXCLP with a constant busy fraction and then calculating the averages for each base location. The used busy fractions are reported in full in the data appendix. The running time of the modified model was somewhat longer but still within acceptable ranges, since one execution of the modified algorithm took less than 2 seconds. In each case we ran both algorithms 20 times.

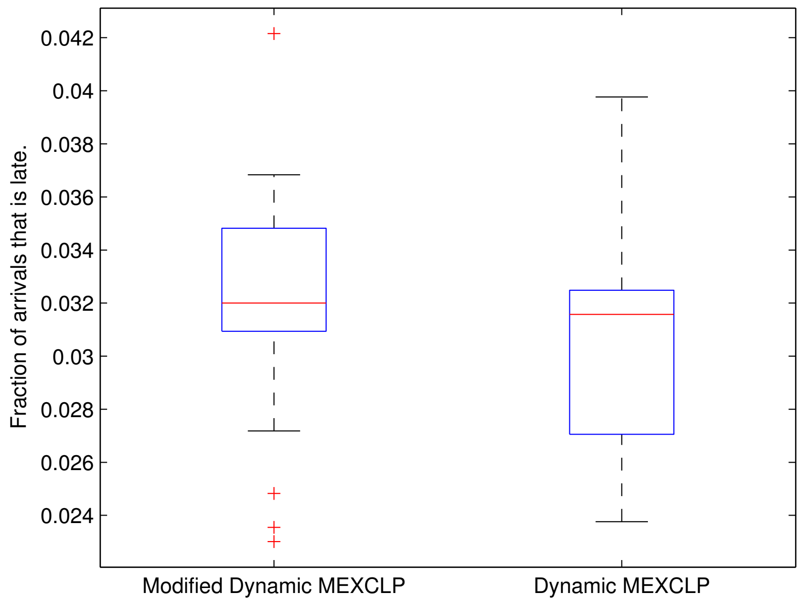

Figure 2: Fraction of arrivals that are late for Dynamic MEXCLP and the modified model respectively in a simulation without uncertainty in the driving times and without varying busy fractions.

As can be seen in Figure 2, the modified model performs similarly to Dynamic MEXCLP if the busy fraction is not entirely correct and there is no uncertainty in driving times. The mean fraction of late arrivals is 0.0309 for Dynamic MEXCLP and 0.0320 for the modified model. The standard deviation is 0.0041 for Dynamic MEXCLP and 0.0046 for the modified model. Dynamic MEXCLP had a lower fraction of too late arrivals in 12 out of 20 cases. The H_0 is that the modified model has a lower or equal median fraction of arrivals that are late than Dynamic MEXCLP. The H_a is that the modified model has a higher median fraction of arrivals that are late than Dynamic MEXCLP. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 8 successes and a success chance of 50%. This gives a chance of $p = 0.2517$ and we do not reject H_0. This means that Dynamic MEXCLP does not perform significantly better than the modified model. This is not surprising since the modified model is in essence an extended version of the Dynamic MEXCLP model.

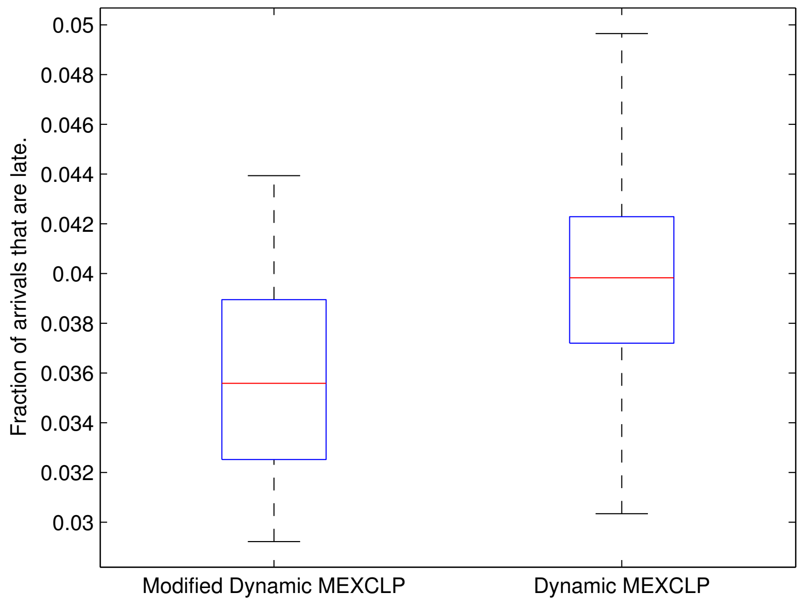

Figure 3: Fraction of arrivals that are late for Dynamic MEXCLP and the modified model respectively in a simulation with uncertainty in the driving times and with varying busy fractions.

In Figure 3 we can see that the modified model performs better than the original in a simulation with uncertainty in the driving times and with data about the busy fractions of different locations. The mean fraction of late arrivals was 0.0359 for the modified method and 0.0404 for Dynamic MEXCLP. The standard deviations were 0.0042 and 0.0046 respectively. The modified model had lower fractions of late arrivals in 16 out of 20 simulations. The H_0 is that Dynamic MEXCLP has a lower or equal median fraction of arrivals that are late than the modified model. The H_a is that Dynamic MEXCLP has a higher median fraction of arrivals that are late than the modified model. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 4 successes and a success chance of 50%. This gives a chance of $p = 0.0059$ and we reject H_0. This means that the modified model does perform significantly better than Dynamic MEXCLP.

The fact that the improvement is not more significant is caused by various reasons, but mostly because these busy fractions are based on the Dynamic MEXCLP outcomes, and are thus more accurate for that method than for the modified method.

4.4 Sensitivity analysis

We also wanted to investigate how the Modified Dynamic MEXCLP and Dynamic MEXCLP react to small changes in parameters especially in the uncertainty in driving times, but also in busy fractions. We are most interested in the modified model, but since we expect that performance will change we will compare it to Dynamic MEXCLP to see how both models react. Our first variation is a simulation with larger variance in both ways namely a $\sigma_{constant}$ of 40 seconds and a $\sigma_{variable}$ of 20% of the mean travel time.

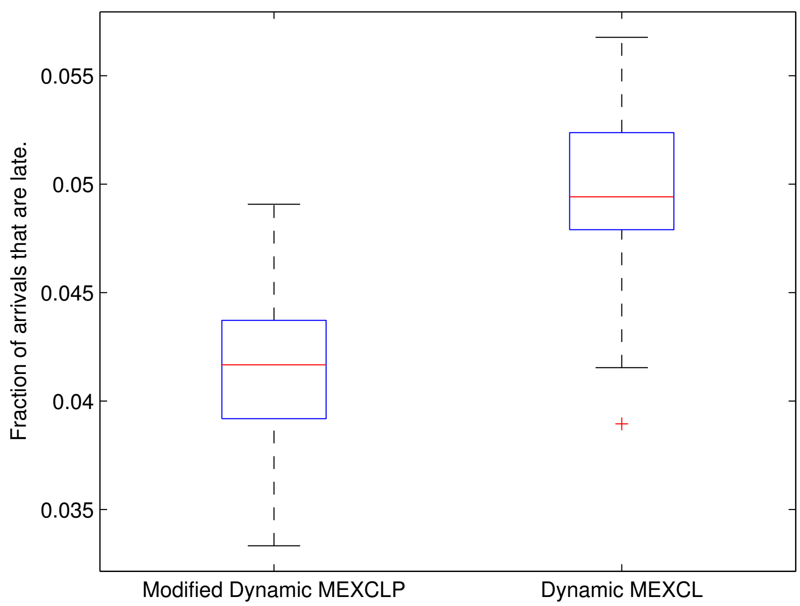

Figure 4: Fraction of arrivals that are late for Dynamic MEXCLP and MEXCLP respectively in a simulation with extra uncertainty in the driving times and with varying busy fractions.

In Figure 4 we can see that the modified model performs much better than Dynamic MEXCLP with high uncertainty in driving times and varying busy fractions. The mean fraction of late arrivals was 0.0418 for the modified method and 0.0492 for Dynamic MEXCLP. The standard deviations were 0.0041 and 0.0046 respectively. The modified model had lower fractions of late arrivals in 19 out of 20 simulations. The H_0 is that Dynamic MEXCLP has a lower or equal median fraction of arrivals that are late than the modified model. The H_a is that Dynamic MEXCLP has a higher median fraction of arrivals that are late than the modified model. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 1 success and a success chance of 50%. This gives a chance of $p = 0.0000$ and we reject H_0. This means that the modified model does perform significantly better than Dynamic MEXCLP.

Our second variation is a simulation with smaller variance in both ways namely a $\sigma_{constant}$ of 20 seconds and a $\sigma_{variable}$ of 10% of the mean travel time.

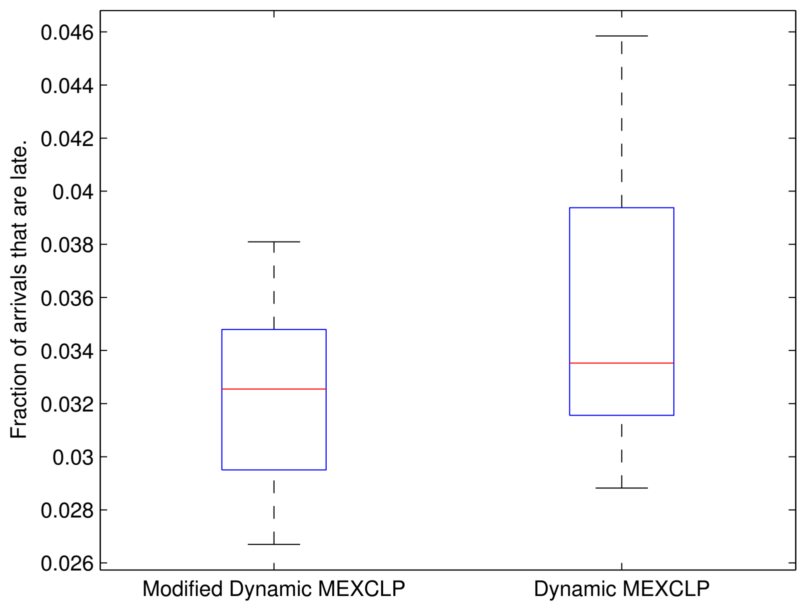

Figure 5: Fraction of arrivals that are late for Dynamic MEXCLP and MEXCLP respectively in a simulation with less uncertainty in the driving times and with varying busy fractions.

In Figure 5 we can see that the modified model performs only slightly better than Dynamic MEXCLP with low uncertainty in driving times and varying busy fractions. The mean fraction of late arrivals was 0.0322 for the modified method and 0.0353 for Dynamic MEXCLP. The standard deviations were 0.0033 and 0.0048 respectively. The modified model had lower fractions of late arrivals in 14 out of 20 simulations. The H_0 is that Dynamic MEXCLP has a lower or equal median fraction of arrivals that are late than the modified model. The H_a is that Dynamic MEXCLP has a higher median fraction of arrivals that are late than the modified model. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 6 successes and a success chance of 50%. This gives a chance of $p = 0.0577$ and we do not reject H_0. This means that the modified model does not perform significantly better than Dynamic MEXCLP. We can see that the modified model performs worse with more uncertainty in driving times but in relation to Dynamic MEXCLP it does perform better. This is what we would expect since the modified model incorporates this uncertainty, but still suffers from it, while Dynamic MEXCLP does not incorporate it at all.

Besides the influence of uncertainty on performance we also wanted to investigate the influence of higher and lower busy fractions. These were generated by respectively adding 0.1 to or subtracting 0.1 from all busy fractions and rounding when this results in a negative number or a number above 1.

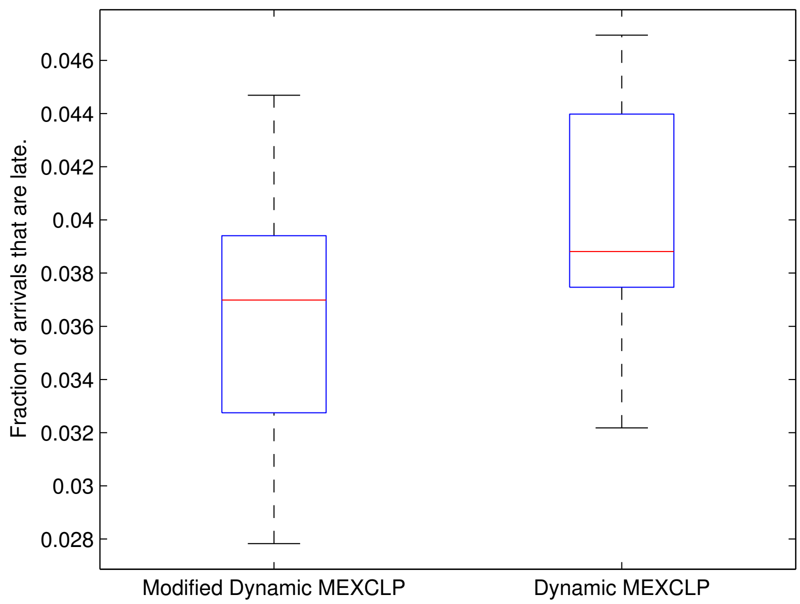

Figure 6: Fraction of arrivals that are late for Dynamic MEXCLP and MEXCLP respectively in a simulation with uncertainty in the driving times and with high varying busy fractions.

In Figure 6 we can see that the modified model does perform slightly better than Dynamic MEXCLP with uncertainty in driving times and high varying busy fractions. The mean fraction of late arrivals was 0.0364 for the modified method and 0.0401 for Dynamic MEXCLP. The standard deviations were 0.0042 and 0.0042 respectively. The modified model had lower fractions of late arrivals in 15 out of 20 simulations. The H_0 is that Dynamic MEXCLP has a lower or equal median fraction of arrivals that are late than the modified model. The H_a is that Dynamic MEXCLP has a higher median fraction of arrivals that are late than the modified model. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 5 successes and a success chance of 50%. This gives a chance of $p = 0.0207$ and we reject H_0. This means that the modified model does perform significantly better than Dynamic MEXCLP.

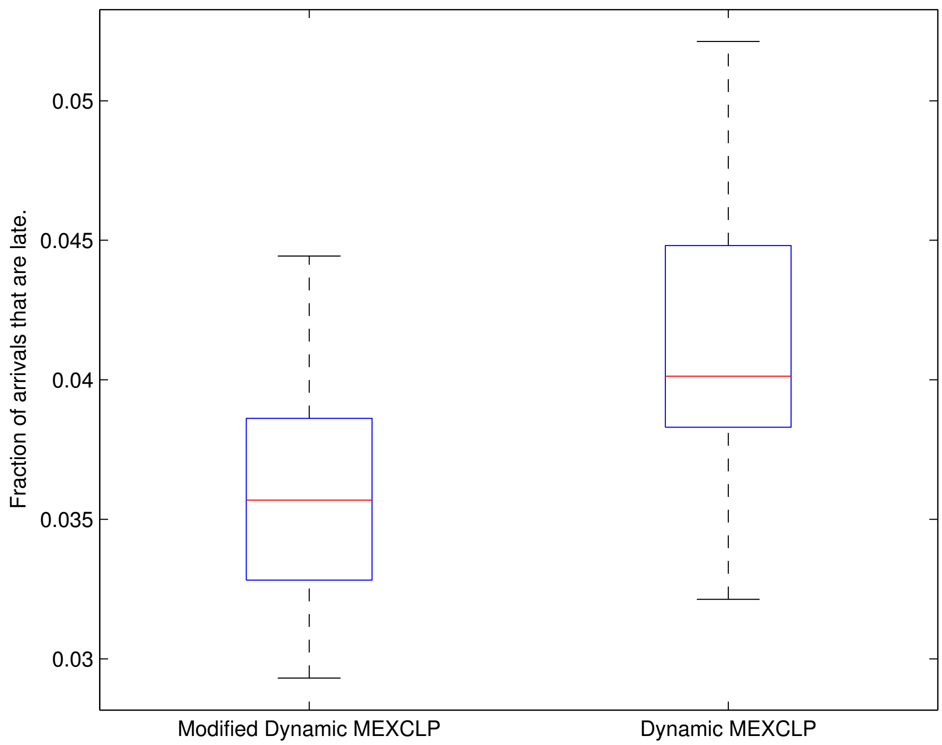

Figure 7: Fraction of arrivals that are late for Dynamic MEXCLP and MEXCLP respectively in a simulation with uncertainty in the driving times and with low varying busy fractions.

In Figure 7 we can see that the modified model performs better than Dynamic MEXCLP with uncertainty in driving times and low varying busy fractions. The mean fraction of late arrivals was 0.0359 for the modified method and 0.0412 for Dynamic MEXCLP. The standard deviations were 0.0040 and 0.0052 respectively. The modified model had lower fractions of late arrivals in

16 out of 20 simulations. The H_0 is that Dynamic MEXCLP has a lower or equal median fraction of arrivals that are late than the modified model. The H_a is that Dynamic MEXCLP has a higher median fraction of arrivals that are late than the modified model. We test with $p < 0.05$ significance. We calculate the chance of this happening under H_0 by calculating the CDF of the binomial distribution with 20 trials and 4 success and a success chance of 50%. This gives a chance of $p = 0.0059$ and we reject H_0. This means that the modified model does perform significantly better than Dynamic MEXCLP.

Note that both results are very similar to the original results. This was expected for Dynamic MEXCLP since in [11] there was established that Dynamic MEXCLP is relatively invariant with regards to the precise busy fraction. The modified model also seems to be relatively invariant with regards to the precise busy fractions.

5 Discussion and conclusion

We want to discuss some limitations and benefits to the methods and concepts in this thesis, and offer some suggestions to modify the models so that these problems are less severe. Then we will summarize our findings and conclude the thesis.

5.1 Discussion

There are various limitations to the algorithms presented in this thesis. The first limitation is that acquiring a reliable busy fraction is very hard, and that busy fractions should actually depend on how many ambulances there are at each base, and thus on the policy up until that point. Another limitation is that the current model only includes currently idle ambulances, while not taking into account the ambulances that are currently busy but could become idle in the near future. This omission has various drawbacks; it can cause sending ambulances from neighborhood to neighborhood without any necessity to do so and can be confusing for the staff. The last major limitation to these models is that these models do not incorporate that there are mass accidents, for example a large fire or pile-up car crash where tens or even hundreds of people are hurt at the same time. These events are very unlikely to occur in the simulation but they do occur in real life with effects that are not modeled, limiting the realism of these results.

The main strong points of the methods and models that are presented in this thesis, are that they are transparent and that they are able to run in real time in real situations. These benefits are preserved for the modified model while performance was enhanced.

We would also like to suggest some possible extensions that could be included into new research.

One possible modification addresses that the current algorithm does not include the information of the currently busy ambulances. To provide better real coverage, a system that uses information regarding the time the ambulances are dispatched and the spots they are dispatched to is probably useful. That information will enable the model to predict where these ambulances will be when they become idle and the moment when they become idle, thus enabling the algorithm to have a better estimate of future coverage.

A possible improvement is that busy fractions are calculated for a number of available ambulances at a base location, so that busy fractions actually vary depending on how many ambulances are present at a certain location. This would allow more realistic busy fractions and thus would probably give better results.

A possible improvement that could be made on the running time is restricting ambulances to only accept jobs that are within a certain distance of the base they are currently based at, since moving outside of that region takes so long that it is never (or almost never) worth it.

5.2 Conclusion

We examined the algorithm in [11] and have found that it indeed generated higher service levels than MEXCLP. We then introduced modifications to incorporate uncertainty in travel times and different busy fractions between bases to this algorithm. Finally we analyzed the results and found that the modified model does perform significantly better with a decrease of arrivals that are too late of 11.1% on average. This occurs if there is uncertainty in the driving times and if there is knowledge of how busy bases are, but not in cases without both these characteristics. We also established that both models are relatively invariant for changes in busy fractions but not for chances in how large the uncertainty is. Finally we noted some weaknesses and strengths in the models. The weaknesses are mainly that it can be very hard to determine realistic busy fractions while the main strength of our modified model is that performs better and runs fast enough to be dynamically used in real situations.

6 Acknowledgments

First and foremost we want to thank the RIVM and the researchers from the REPRO project for providing us with most of the data we use. We would also like to thank our thesis coordinator R.B.O. Kerkkamp MSc for his valuable input and for arranging that we could use real data.

References

[1] Ramon Alanis, Armann Ingolfsson, and Bora Kolfal. “A Markov Chain Model for an EMS System with Repositioning”. In: *Production and Operations Management* 22.1 (2013), pp. 216–231. ISSN: 1937-5956. DOI: 10.1111/j.1937-5956.2012.01362.x. URL: http://dx.doi.org/10.1111/j.1937-5956.2012.01362.x.

[2] Pieter L. van den Berg and Karen Aardal. “Time-dependent MEXCLP with start-up and relocation cost”. In: *European Journal of Operational Research* 242.2 (2015), pp. 383 –389. ISSN: 0377-2217. DOI: http://dx.doi.org/10.1016/j.ejor.2014.10.013. URL: http://www.sciencedirect.com/science/article/pii/S0377221714008170.

[3] M S Daskin, K Hogan, and C ReVelle. “Integration of Multiple, Excess, Backup, and Expected Covering Models”. In: *Environment and Planning B: Planning and Design* 15.1 (1988), pp. 15–35. DOI: 10.1068/b150015. eprint: http://epb.sagepub.com/content/15/1/15.full.pdf+html. URL: http://epb.sagepub.com/content/15/1/15.abstract.

[4] Mark S. Daskin. “A Maximum Expected Covering Location Model: Formulation, Properties and Heuristic Solution”. In: *Transportation Science* 17.1 (1983), pp. 48–70. DOI: 10.1287/trsc.17.1.48. eprint: http://dx.doi.org/10.1287/trsc.17.1.48. URL: http://dx.doi.org/10.1287/trsc.17.1.48.

[5] Erhan Erkut, Armann Ingolfsson, and Gune Erdogan. “Ambulance location for maximum survival”. In: *Naval Research Logistics (NRL)* 55.1 (2008), pp. 42–58. ISSN: 1520-6750. DOI: 10.1002/nav.20267. URL: http://dx.doi.org/10.1002/nav.20267.

[6] Michel Gendreau, Gilbert Laporte, and Frederic Semet. “A dynamic model and parallel tabu search heuristic for real-time ambulance relocation”. In: *Parallel Computing* 27.12 (2001). Applications of parallel computing in transportation, pp. 1641 –1653. ISSN: 0167-8191. DOI: http://dx.doi.org/10.1016/S0167-8191(01)00103-X. URL: http://www.sciencedirect.com/science/article/pii/S016781910100103X.

[7] Kommer GJ and Zwakhals SLN. *Modellen referentiekader ambulancezorg 2008 : Documentatie rijtijden- en capaciteitsmode*. 2011. URL: http://www.rivm.nl/en/Documents_and_publications/Scientific/Reports/2011/augustus/Models_in_the_reference_framework_for_ambulance_care_2008_Technical_documentation_on_the_driving_time_model_and_the_capacity_model.

[8] Jeffrey Goldberg and Luis Paz. “Locating Emergency Vehicle Bases When Service Time Depends on Call Location”. In: *Transportation Science* 25.4 (1991), pp. 264–280. DOI: 10.1287/trsc.25.4.264. eprint: http://dx.doi.org/10.1287/trsc.25.4.264. URL: http://dx.doi.org/10.1287/trsc.25.4.264.

[9] Jeffrey Goldberg et al. "Validating and applying a model for locating emergency medical vehicles in Tuczon, AZ". In: *European Journal of Operational Research* 49.3 (1990), pp. 308 –324. ISSN: 0377-2217. DOI: http://dx.doi.org/10.1016/0377-2217(90)90402-W. URL: http://www.sciencedirect.com/science/article/pii/037722179090402W.

[10] Armann Ingolfsson, Susan Budge, and Erhan Erkut. "Optimal ambulance location with random delays and travel times". In: *Health Care Management Science* 11.3 (2008), pp. 262–274. ISSN: 1572-9389. DOI: 10.1007/s10729-007-9048-1. URL: http://dx.doi.org/10.1007/s10729-007-9048-1.

[11] C.J. Jagtenberg, S. Bhulai, and R.D. van der Mei. "An efficient heuristic for real-time ambulance redeployment". In: *Operations Research for Health Care* 4 (2015), pp. 27 –35. ISSN: 2211-6923. DOI: http://dx.doi.org/10.1016/j.orhc.2015.01.001. URL: http://www.sciencedirect.com/science/article/pii/S2211692314200075.

[12] Matthew S. Maxwell et al. "Approximate Dynamic Programming for Ambulance Redeployment". In: *INFORMS Journal on Computing* 22.2 (2010), pp. 266–281. DOI: 10.1287/ijoc.1090.0345. eprint: http://dx.doi.org/10.1287/ijoc.1090.0345. URL: http://dx.doi.org/10.1287/ijoc.1090.0345.

[13] Ambulancezorg Nederland. *"Ambulances in-zicht 2012 "*. Ambulancezorg Nederland, 2012.

[14] John F. Repede and John J. Bernardo. "Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky". In: *European Journal of Operational Research* 75.3 (1994), pp. 567 –581. ISSN: 0377-2217. DOI: http://dx.doi.org/10.1016/0377-2217(94)90297-6. URL: http://www.sciencedirect.com/science/article/pii/0377221794902976.

[15] Charles Revelle and Kathleen Hogan. "The maximum reliability location problem and α-reliablep-center problem: Derivatives of the probabilistic location set covering problem". In: *Annals of Operations Research* 18.1 (1989), pp. 155–173. ISSN: 1572-9338. DOI: 10.1007/BF02097801. URL: http://dx.doi.org/10.1007/BF02097801.

[16] Cem Saydam and Mark McKnew. "Applications and Implementation a separable programming approach to expected coverage: an application to ambulance location". In: *Decision Sciences* 16.4 (1985), pp. 381–398. ISSN: 1540-5915. DOI: 10.1111/j.1540-5915.1985.tb01490.x. URL: http://dx.doi.org/10.1111/j.1540-5915.1985.tb01490.x.

A Algorithms

Algorithm 5 Generate a new accident

```
function GENERATE A NEWACCIDENT(ambulances, eventlist, nextevent, instance, typeofalghorithm)
    initialize new accident Alpha
    according to the demand assign location to Alpha
    generate time for Alpha
    generate event Aleph of type "new accident"
    add Aleph to eventlist
    add new Alpha to accidents
    add Alpha to queue
    delete this event from eventlist
return queue, accidents, eventlist
end function
```

Algorithm 6 Arrival of an ambulance at an accident

```
function ARRIVAL OF AN AMBULANCE AT AN ACCIDENT(ambulances, eventlist, nextevent, instance, typeofalgorithm)
    arrivals=arrivals+1
    if currenttime-accidenttime > maxresponsetime then
        latearrivals=latearrivals+1
    end if
    delete relevant accident from accidents
    delete this event from eventlist
    decide if patient should be hospitalized
    if hospitalized=true then
        select closest hospital
        calculate travel time to that hospital
        determine service time at accident
        add service time to travel time to get new arrival time
        adjust ambulance to new hospital and new arrival time
        make new event Aleph of type arrival at hospital
        add Aleph to eventlist
    else
        determine service time at accident
        adjust ambulance to new service time
        make new event Aleph of type make ambulance idle
        add Aleph to eventlist
    end if
return accidents, eventlist, ambulances
end function
```

Algorithm 7 Arrival of an ambulance at a hospital

```
function ARRIVAL OF AN AMBULANCE AT AN HOSPITAL(ambulances, eventlist, nextevent, instance)
    determine service time at hospital
    delete this event from eventlist
    make new event Aleph of type "make ambulance idle"
    add Aleph to eventlist
return Eventlist
end function
```

Algorithm 8 Make ambulance idle

```
function MAKE AMBULANCE IDLE(ambulances, eventlist, nextevent, instance, typeofalghorithm)
    get which ambulance should be made idle
    if typeofalghorithm is MEXCLP then
        get predetermined base for selected ambulance
        set selected ambulance to idle
        update ambulances
    else
        set selected ambulance to idle
        identify all idle ambulances
        identify all destinations of idle ambulances
        run Dynamic MEXCLP()
        update ambulances
    end if

    delete this event from eventlist
return ambulances, eventlist
end function
```

Algorithm 9 Dispatch ambulances

function DISPATCH AMBULANCES(ambulances, eventlist,queue,accidents, instance)
    select the accident that happened earliest from queue
    delete selected accident from queue
    update this accident so that it indicates that an ambulance is on its way
    initialize t_{travel} for all ambulances
    for all ambulances **do**
        if ambulance has arrived at its destination **then**
            set t_{travel} to t_{dest}
        else
            calculate d_{loc}
            calculate d_{dest}
            get t_{dest}
            calculate $t_{travel} = \left(\frac{d_{loc}}{d_{dest}}\right)(t_{dest})$
        end if
    end for
    get idle ambulance with smallest distance(t_{dest})
    update selected ambulance
    update ambulances
    make new event of type arrival of an ambulance to an accident
    update eventlist
return ambulances, eventlist, accidents, queue
end function

B Results

we now give our results for some of the less important cases since they are not what either model was meant for.

Firstly we have the model with uncertainty in driving times but no varying busy fractions in Figure 8. The mean fraction of late arrivals was 0.0393 for the modified method and 0.0404 for Dynamic MEXCLP. The standard deviations were 0.0030 and 0.0029 respectively. The modified model had lower fractions of late arrivals in 13 out of 20 simulations.

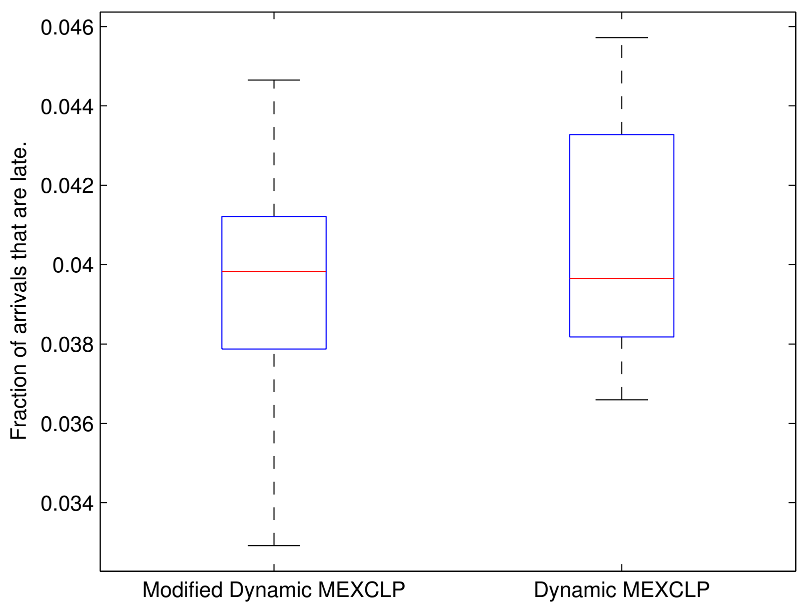

Figure 8: Fraction of arrivals that are late for Dynamic MEXCLP and the modified model respectively in a simulation with uncertainty in the driving times and without varying busy fractions.

Secondly we have the model without uncertainty in driving times but with varying busy fractions in Figure 9. The mean fraction of late arrivals was 0.0301 for the modified method and 0.0321 for Dynamic MEXCLP. The standard deviations were 0.0037 and 0.0037 respectively. The modified model had lower fractions of late arrivals in 12 out of 20 simulations.

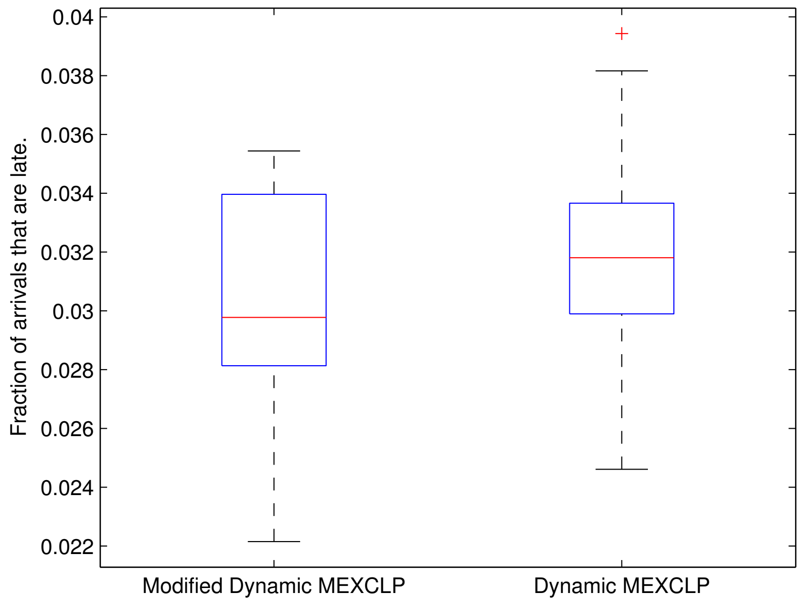

Figure 9: Fraction of arrivals that are late for Dynamic MEXCLP and the modified model respectively in a simulation without uncertainty in the driving times and with varying busy fractions.

C Data

Table 1: busy fractions per base

Base	high busy fraction	normal busy fraction	low busy fraction
3811	0.3572	0.2572	0.1572
3823	0.3434	0.2434	0.1434
3941	0.2334	0.1334	0.0334
3608	0.3056	0.2056	0.1056
3436	0.3724	0.2724	0.1724
3911	0.2563	0.1563	0.0563
3582	0.4867	0.3867	0.2867
3561	0.4100	0.3100	0.2100
3645	0.1921	0.0921	0.0000
3447	0.2455	0.1455	0.0455
3707	0.1382	0.0382	0.0000
mean	0.3037	0.2037	0.1100

Table 2: Hospital locations as postal codes

Hospitals
3435
3903
3941
3813
3994
3762
3743
3707
3584
3582
3562
3453
3542
3452
3447
3641