Analysis of the logistical process at the bed cleaning department in a hospital

Master Thesis
Econometrics and Management Science
Operations Research and Quantitative Logistics

in cooperation with,

Erasmus Medical Center

Rotterdam

14-11-2013

Author:

Joran van Heyningen (333369)

Academic supervisor:

A.P.M. Wagelmans

Co-reader:

Wilco van den Heuvel

External supervisor:

J.M. van Oostrum

Abstract

The field of Operations Research is nowadays used a lot in university hospitals. In this master thesis different methods of Operations Research are used to analyze data of the Erasmus Medical Center in Rotterdam. In the first part of this thesis forecasts are made for the number of admissions and discharges in the Erasmus Medical Center by using different regression models. These forecasts can be used for all kind of different decisions that have to be made in the future and are especially interesting now, since a new hospital is being built. For example, the number of beds needed in this new hospital depends on the forecasts.

These forecasts can also be used for a specific process in the hospital: the process of transporting clean and dirty beds through the hospital. The second part of this thesis is about this process and the crew that is responsible for this process. A simulation is used to test different options for the system that is used to assign tasks of picking up dirty beds and delivering clean beds to the wards. The results of this simulation will be used to make this process more efficient.

Keywords: healthcare, hospital, bed cleaning department, forecasts, simulation

Contents

1. Introduction

Worldwide the costs of healthcare expenditure have grown a lot in the past decades. Especially in the Netherlands, the region of this research, the healthcare expenditure has become high and is expected to rise in the upcoming years. The total expenditure on healthcare in 2010 was 12.0% of Gross Domestic Product (GDP) and it has grown with 5.6% per year from 2000 until 2010. Both the total expenditure as the trend upwards in it are quite high compared to other countries in Europe (CPB, 2013). The increase is too high in relation to national gross product, so measures must be taken (Machielsen, 2011).

Therefore processes have to be improved or optimized in the medical sector and the field of Operations Research (OR) is getting more important in this sector (Malhotra & Zodpey, 2010). A lot of the healthcare expenditure goes to hospitals and in this research different methods will be used from the field of OR to analyze and improve processes in the Erasmus Medical Center (MC) in Rotterdam. Another reason for making the processes in Dutch hospitals more efficient is the introduction of a new financing system in the Netherlands: the DOT system (see Appendix A). This system is created to make the costs that hospitals make more transparent and therefore hospitals will compete more with each other now (Machielsen, 2011). The prices of the treatments are also becoming more negotiable and this is also a reason for an increasing competition for hospitals to obtain contracts with payers. This means that to make profits a hospital has to be as efficiently as possible now.

Erasmus MC is one of the eight academic hospitals in the Netherlands and it is the largest and most versatile university medical center in the Netherlands. Being a university medical center means that the hospital has two more core tasks than only the patient care: research and education. Erasmus MC aims to be one of the best medical research institutes in the world and it aims to be one of the favorite medical training institutes in the Netherlands.

To reach these goals a construction of a new hospital started in 2009. The new hospital should be finished in 2017 and the first part is already delivered in 2013. On a strategic level it is important to know how many beds are needed in the new hospital. Therefore an estimation of the number of admissions and discharges of patients in the upcoming years needs to be made. This information is useful for the project of the newly built hospital, but also for different processes on an operational level in the hospital. The focus of this thesis will be on one of these processes: the logistical process behind the cleaning of the beds in the hospital. And more specifically the focus will be on the crew that is responsible for moving these beds through the hospital. We want to analyze the admissions and discharges of patients to make the crew schedule depend upon this information. We also want to analyze the way the tasks are assigned to the crew to see if this can be done more efficiently. With more efficiently is meant that fewer tasks are late and the lateness should be minimal. The main question will then be:

How can the logistical process of movements of clean and dirty beds be improved?

1.1 Erasmus MC

Erasmus MC contains three main sections: Centrum Location, Oncology Centrum Daniel den Hoed and the children's hospital Sophia. Currently these three sections are located separately, but they will be integrated in the new hospital. Each of these sections has their own organization in cleaning the beds. In this research only the Centrum Location will be analyzed, which is the largest section of the hospital. The psychiatry is also excluded from this analysis, because they have their own beds which are cleaned at the psychiatry.

1.2 Bed cleaning department

The bed cleaning department is responsible for picking up dirty beds at the wards, cleaning them and delivering cleaned beds to the wards. Only the team of workers that transport the beds will be analyzed in this research, because in the new hospital will be cleaning machines that clean the beds automatically. Currently a separate team of workers is present at the hospital which only cleans beds.

1.3 I-transport

I-transport is a real-time communication and information system which is developed by the company dir/Active bv. This system has already been used for years in Erasmus MC for patient transport. Since March 2013 it is also used for transportation of beds. In this system the requests are made for a clean bed or a dirty bed that needs to be picked up and cleaned. The system creates a task for each request and assigns the tasks to the workers of the bed cleaning department. This happens with use of a mobile phone. The phone is used by a worker to tell the system a task is started or finished which means a worker is busy with a task or ready for a new task.

1.4 Motivation

Almost every action in the hospital is registered, which generates a lot of data. However, a lot of this information is not used in making decisions on a tactical and operational level. The motivation for this thesis comes from the bed cleaning department of Erasmus MC. The bed cleaning department has to make a crew schedule for each week, but this is done without using any historical data. This results in the same size of the workforce for each day. By analyzing the historical data we want to give the bed cleaning department useful information which they can use in making these schedules.

1.5 Structure of this paper

The first part of this thesis is mainly about analyzing the data of admissions and discharges of patients. This part starts with finding certain patterns in the data, for example that Monday is the busiest day of the week. The next step is to make forecasts of the admissions and discharges for 2013. Different models are tested and the best model is used for making these forecasts. The forecasts are then translated into tasks, which need to be fulfilled by the bed cleaning department. The last step is to link these tasks to the number of workers needed for the bed cleaning department.

The second part of this thesis is about simulating the process of transporting the beds through the hospital. We want to analyze if the way the workers are assigned to tasks can be done more efficiently. Therefore a simulation model needs to be build which can simulate the whole process of transporting the beds. With this simulation the effect of changing different options in the process will be tested, for example using a different way of assigning tasks to workers, but also the effect of making it possible to let a worker transport two beds in one task.

2. Literature Overview

On a strategic level the focus of Erasmus MC is on investigating how many beds are needed in the hospital and a lot of articles have been written about this subject. One of these articles is about building a model to estimate the total number of hospital beds needed in a hospital in New York (Green, 2002). This model uses a M/M/s queuing model to model the flow of patients, which means that patients arrive according to a time-homogeneous Poisson process and the length of stay has an exponential distribution.

Another popular method in the field of Operations Research (OR) which is used in subjects related to admissions and discharges of patients is Integer Linear Programming (ILP). This is used for example to solve the problem of planning admissions in the hospital over different specialisms (Adan & Vissers, 2002).

To forecast the number of admissions and discharges regression models are often used. Linear regression, neural networks regression and regression trees can all be used to model the admissions of patients (Garcia, 2009). Regressions are also often used in the financial sector and different financial articles and books are written which broadly explain how to handle outliers when using regression models (Wilson, 1997 and Heij et al., 2004).

Simulation is another method which is often used when OR is used in the medical sector. For example, a simulation model is constructed to follow unique patients from pre-admission until discharge. This model was used to analyze the hospital bed utilization and test the influence of different bed allocations policies while maintaining the same service level (Dumas, 1985). Another similar simulation model is used to determine the optimal bed allocation policy. This computer simulation model appears to perform in a manner consistent with the “real world” (Goldman et al., 1968).

Also an article is written which gives an overview of different simulation models used in health care. “This article describes a multi-dimensional approach to the classification of the research literature on simulation and modelling in health care.” (Brailsford et al., 2009). The main conclusion of this article is that although the literature on simulation in health care is expanding rapidly, a main framework is missing. The different articles lack standards and consistency which makes it hard to compare them.

3. Problem Description

In the next sections the problem will be explained in details.

3.1 Circulation of the beds

Whenever a patient enters the hospital a clean bed is waiting for him/her at a specific ward. Usually the patient stays in this bed for the whole length of his stay. During his stay he may switch wards, but this will mostly happen in the same bed. Once the patient leaves the hospital the bed has to be cleaned. A bed that needs to be cleaned is registered by a nurse in the system I-transport and then gets picked up by a worker of the bed cleaning department. In the bed cleaning department the bed is cleaned and after that the cleaned bed will be delivered again to a ward that ordered a bed with the use of I-transport.

The interesting movements of the beds will thus be the movements from the clinic to the bed cleaning department and back again. A bed will be at the clinic for the length of stay (LOS) of a patient. Then a transportation time will be needed to deliver a bed to the bed cleaning department and this transportation time includes the waiting time for example in front of an elevator. At the bed cleaning department the bed is cleaned which will be done within five minutes; especially with the new cleaning robots. Then a transportation time will be needed to deliver the bed to a ward again. The circulation of the beds is shown in figure 1.

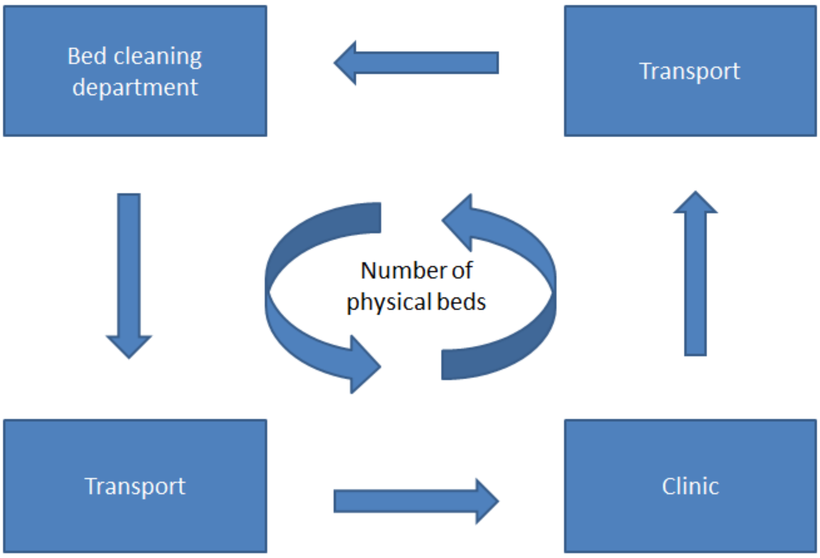

Figure 1: circulation of beds

The definition of physical beds is used, because in the hospital two different terms for beds are used: physical beds and operational beds. An operational bed is a management term which relates the number of staff to the beds. However, in this research only the term physical beds is used, which means the real physical beds that are present in the hospital.

Currently the storage of the beds takes place everywhere in the chain. Dirty beds are for example put on the hallways of the hospital while not even a signal is given to the bed cleaning department that it can be picked up. This happens because the system of I-transport has been introduced recently in the hospital and therefore the system does not work yet as it is supposed to.

After some time the teething troubles will be overcome and the system will work as it is supposed to be. The circulation of beds will then be a push system. This means that once a bed needs to be cleaned it is immediately pushed through the system from the clinic, so a signal is given immediately to the bed cleaning department, with the use of I-transport, that it can be picked up. When a bed is delivered at the bed cleaning department it is cleaned as fast as possible. After the bed is cleaned it is ready to be delivered to a ward again.

There are many different wards at which a bed can be delivered and we do not want to estimate a different transportation time for each ward. Therefore wards have to be aggregated in such a way that less transportation times have to be estimated and the same transportation time can be used for wards located close to each other.

<u>Research Questions:</u>
How can we estimate the total time needed to transport a bed to (or from) a specific ward from (or to) the bed cleaning department?
How can we aggregate certain wards in the hospital?

3.2 Admissions and discharges

We want to focus on the crew planning for the future and therefore the historical data of admissions and discharges needs to be analyzed. The autonomy of the surgeons plays an important role in the planning of the operating theatres (van Oostrum, 2009) and this is causing the peaks in admissions at certain days. We want to find these patterns in the data for different days, months or years. This analysis can be used for other processes as well, for example for making decisions for the new hospital.

The next step is to make forecasts on the number of admissions and discharges. In making the forecasts we have to take into account that the market is more like a supply-driven market than a demand-driven market. In other words; there are enough patients that need a treatment, but the production of the hospital is the restricting factor. The hospital has a maximal number of patients that they can treat in a year, because of the limited number of specialists that work in the hospital. Another limiting factor is the space in the hospital, for example the number of operating theatres restricts the number of operations that can be done during a year.

Next to the influence of the specialists the time dimension plays an important role. For example, the number of admissions in the weekends or on holidays will be much lower than on weekdays. Also the month might have a lot of influence on the number of admissions. For example, in December there might be a lot of patients who want to have a certain treatment that is not reimbursed by their health insurance anymore in the next year. In forecasting the number of admissions and discharges for the upcoming year we will build different regression models.

Not every admission will need a clean bed which means these forecasts cannot be translated directly into tasks for the bed cleaning department. Which fraction of the admissions does need a bed is unknown and therefore the last step is to find a way to link the forecasts to task for the bed cleaning department.

Research Questions:

Are there any (seasonal) patterns in the data?
Which models can be used to forecast the number of admissions and discharges?
How can we translate the forecasted values into a number of tasks for the bed cleaning department?

3.3 Assigning the tasks

In Erasmus MC the system of I-transport is used to assign a task to a worker. Each task contains the following information: the sort of task (picking up a dirty bed or delivering a clean bed), the location it comes from or goes to and the estimated time it will be finished. A registered task will have two hours before it has to be finished, so whenever a task is registered by a nurse in I-transport it gets a planned arrival time of two hours later. However, the bed cleaning department aims at picking up, cleaning and delivering a new bed within one hour.

Every task has an estimated time which it takes to finish the task. This is the transportation time to go from start location to end location. When the system was first implemented a matrix of transportation times was created, the start matrix, from each location to another location. To go from location A to B or from location B to A will have the same time in this matrix, so basically it is a triangular matrix. However, these are rough estimates based upon the number of floors the task has to go up and down and the distance it has to cover. In reality these times will differ and therefore the real transportation time of every task that is performed is stored in a history table. Whenever a new task is created between location A and location B the estimated transportation time will be taken in the following order:

1. from history table, from A to B
2. from history table, from B to A
3. from the start matrix

After the system has been implemented for a few months and the history table is filled with transportation times for all combinations of locations, only option 1: the history table from A to B will be used. The start matrix will still be used for the travel times that a worker has to make without a bed from an end location of a completed task to a start location of a new task.

By subtracting the transportation times from the planned arrival time the latest start time of a task is calculated. This creates a list of tasks which have a latest start time. Based upon these times and the location of the workers the tasks can be assigned to a worker. There are different ways of assigning the tasks and different options in I-transport which influence the efficiency of the system. With efficiency is meant that tasks will be finished on time and if tasks are finished late the lateness should be minimized.

<u>Research questions:</u>

Which way of assigning the tasks is the most efficient?
What is the influence of the different options in I-transport on the efficiency?

4. Data

A lot of data is available at Erasmus MC, because everything that happens in the hospital is registered. This is all stored in databases and it goes back until 1993. However, the more recent data will be more reliable than data of 10 years ago. Also the way of registration is improved continuously, because advisors educate the medical staff that good registration is important. For example, some divisions of the hospital might register every patient that leaves at any moment in the morning at a discharge time of 8 A.M. while actually they leave somewhere during the morning. The registration of these divisions can be improved by making the discharge times more precise. Another important reason for using only recent data is the ageing of the Dutch population. Older people have a higher probability of needing service from a hospital and therefore the demand will be higher. Because of these reasons only data of the last three years is used: 2010, 2011 and 2012.

In these datasets is data available about the admissions and discharges of patients. This data contains the exact date, hours and minutes of admissions and discharges of patients. Patients can stay at a lot of different divisions during one treatment and every transfer to another division is stored as well. The registrations for this data are very reliable when taking the days as dimension, but get less reliable when you want to use the exact hours or even minutes as dimension. Some divisions in the hospital might only register whole hours as admission or discharge time or they might use a discharge time which is always the same. Therefore the number of admissions and discharges will not be further drilled down than on a daily level.

The system I-transport also contains relevant data. Data is available about finished tasks which contains the time a task is planned to be finished, the actual time it is finished, the transportation time, the start location and end location. Also data is available about the crew of the bed cleaning department. This data contains the number of task a worker finishes in a day, the part of a working day the worker is actually busy and more. However, this data is less reliable since I-transport is used only recently for the bed cleaning department and in the first months the system was not working as it was supposed to be. Therefore only a few months of data is available out of I-transport.

5. Data Analysis

In this chapter the data of admissions and discharges of 2010, 2011 and 2012 will be analyzed. The first step will be to remove outliers from the data. After the outliers have been deleted the data will be further analyzed.

5.1 Outliers

Outliers are often very influential and that is why they need to be dealt with. Especially when making forecasts the outliers have to be analyzed. In general it does not mean that every outlier is deleted from the data automatically: some outliers can be important for the process you want to model.

Although a hospital will never be entirely closed it will most likely have less admissions and discharges on Dutch Holidays. To get a good view on the average number of admissions and discharges for each day, those Dutch Holidays can be seen as outliers.

Before deleting the Dutch Holidays we want to use a simple and fast method to detect these outliers. Therefore boxplots, also called box-and-whisker plots, are used. Boxplots are developed by John W. Tukey (Tukey, 1977) and are often used to identify outliers. The results of these boxplots are shown in figure 2. The boxplot are shown per day, because it can be easily seen that the number of admissions and discharges heavily depends on the day of the week. For all boxplots holds that the bottom and the top of the box are the first and third quartiles ($\boldsymbol{q_1}$ and $\boldsymbol{q_3}$) and the median is shown as a band inside the box. However, the maximum length of the whiskers ($\boldsymbol{w}$) may differ for different type of boxplots (Benjamini, 1988). This means that points are drawn as outliers if they are larger than $\boldsymbol{q_3 + w(q_3 - q_1)}$ or smaller than $\boldsymbol{q_1 + w(q_3 - q_1)}$. The maximum length of the whiskers is defined as 1.5, which is the most common length.

The discharges of 2012 are shown as an example in figure 2, but the same analyses is done for both the admissions and discharges of 2010, 2011 and 2012.

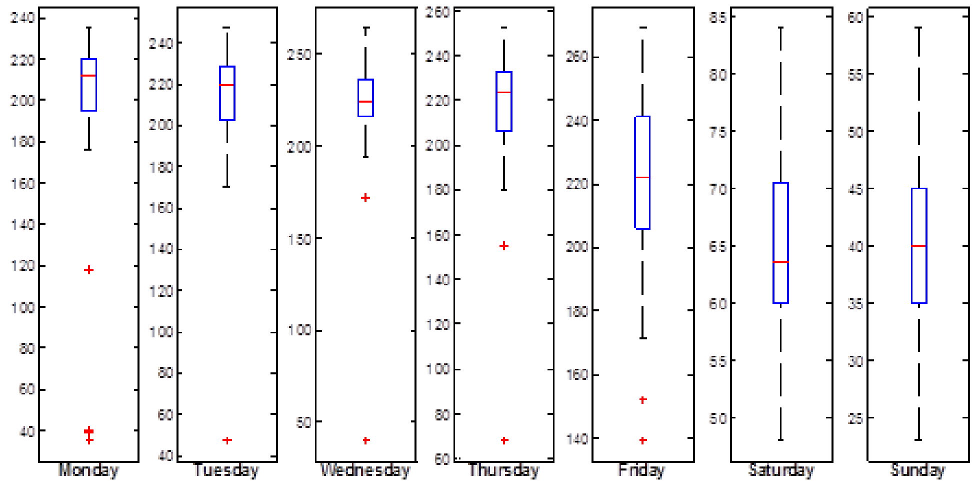

Figure 2: boxplots of historical data of discharges of 2012

By analyzing the data we found out that the following Dutch Holidays can be considered as outliers: New Year's Day, Easter (First and Second), Queen's Day, the Ascension Day, Christmas (First and Second), Whit Sunday and Monday. These are also the days for which everybody in the Netherlands, who does not work in a special sector like the medical sector, is free. However, there are more days which can be seen as an outlier. For example, the Friday after Ascension Day, Christmas Evening and the days between Christmas Evening and New Year's Day (24th December – 1st January). Although these days might not be holidays themselves, the cause of them being outliers are the Dutch Holidays. In figure 3 shows the result after deleting all these days out of our dataset.

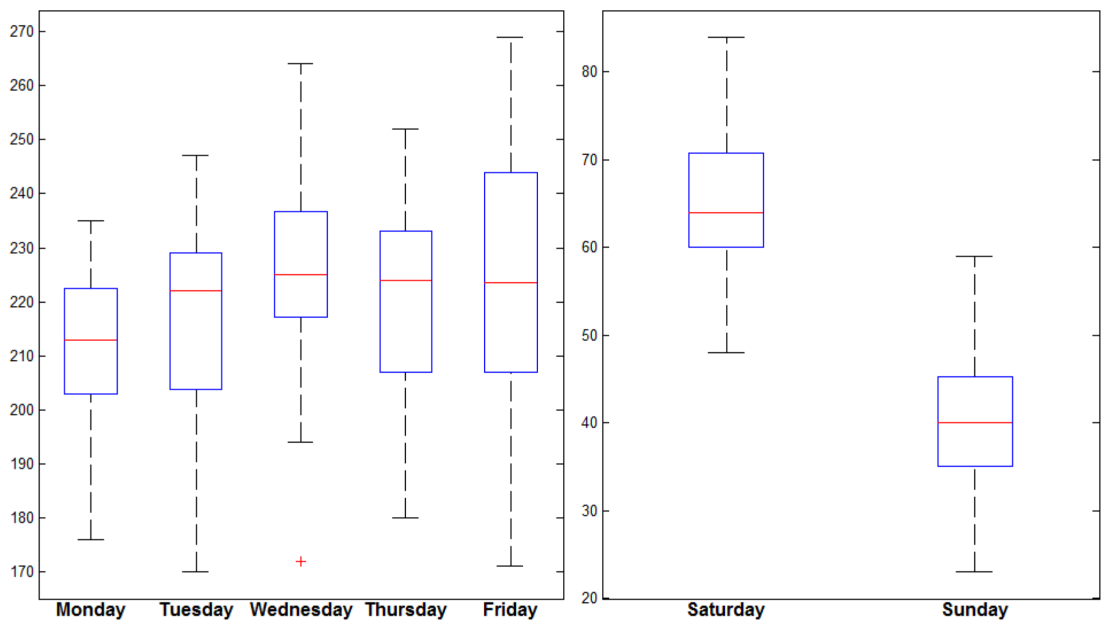

Figure 3: boxplots of historical data of discharges of 2012 after deleting outliers

Almost all outliers out of figure 2 are deleted from the data now. For the discharges of 2012 only one outlier, which is detected by the boxplot method, is left over. This detected outlier is on Wednesday the 1st of April. There seems to be no reason to delete this day from the data. It is not related to a Dutch Holiday and it does not seem likely that the admissions and discharges have been wrongly registered for this day. Therefore this Wednesday is kept in the data and could be an important observation of our data.

The Friday seems to have the widest spread in number of discharges. Figure 3 also shows that in weekends the spread and the means are significantly lower than on weekdays. Further analysis of the data will be done in the next section.

5.2 Features

The next step is to see if certain features in the data can be found; for example a trend or a seasonality pattern. We would expect that the number of admissions and discharges heavily depends on the day or month. In the following two tables the mean (μ), standard deviation (σ), minimum (min) and maximum (max) are shown for each day.

	2010				2011				2012			
	μ	σ	min	max	μ	σ	min	max	μ	σ	min	max
Monday	234.2	19.3	174	268	246.5	16.0	201	288	255.5	15.8	223	281
Tuesday	204.2	20.2	160	251	227.2	17.9	186	276	229.5	15.3	194	256
Wednesday	197.1	15.9	158	237	215.5	14.1	188	258	228.0	14.3	199	259
Thursday	201.6	20.1	156	235	207.7	15.8	162	247	222.2	19.1	177	261
Friday	156.8	16.1	115	189	169.5	16.2	137	198	184.0	16.9	150	221
Saturday	36.9	6.3	26	49	39.5	7.2	28	59	41.8	7.3	27	56
Sunday	56.1	7.1	41	73	58.9	9.8	41	88	55.2	7.8	36	74

Table 1: statistics of historical data of admissions

	2010				2011				2012			
	μ	σ	min	max	μ	σ	min	max	μ	σ	min	max
Monday	181.6	17.2	128	211	196.8	14.0	154	223	211.4	13.6	176	235
Tuesday	199.6	17.5	171	233	220.7	17.4	190	266	217.0	17.0	170	247
Wednesday	194.1	21.0	148	241	210.5	17.1	178	255	224.8	16.8	172	264
Thursday	198.6	20.3	140	253	202.6	17.4	166	241	221.2	17.5	180	252
Friday	203.4	21.6	150	248	217.6	18.8	178	253	223.9	22.5	171	269
Saturday	59.6	9.5	35	78	65.4	8.3	50	84	65.4	8.1	48	84
Sunday	36.6	6.3	16	50	39.2	7.4	28	63	40.4	7.6	23	59

Table 2: statistics of historical data of discharges

The first thing that is noticed is that the Saturdays and Sundays have a significant lower amount of admissions and discharges compared to the weekdays. These two days are different themselves as well, because there are more admissions on a Sunday and more discharges on a Saturday.

For the discharges the statistics of the weekdays are quite similar, especially in 2012. Notice that there are still somewhat less discharges on the Monday. For the admissions the differences are larger: on Monday there are significant more admissions while on Tuesday, Wednesday and Thursday they are quite similar. However, on Friday there are a lot less admissions. The last point we mention is that there seems to be a trend upward for both the admissions and the discharges. This trend is especially recognizable for the weekdays, on Sunday there is almost no trend up or even a downward trend. See Appendix B for bar charts in which these patterns can also be recognized easily.

The difference in the minimum and maximum value for each day might be explained by the month in which these days are. The bar chart in figure 4 is created to recognize patterns in the months.

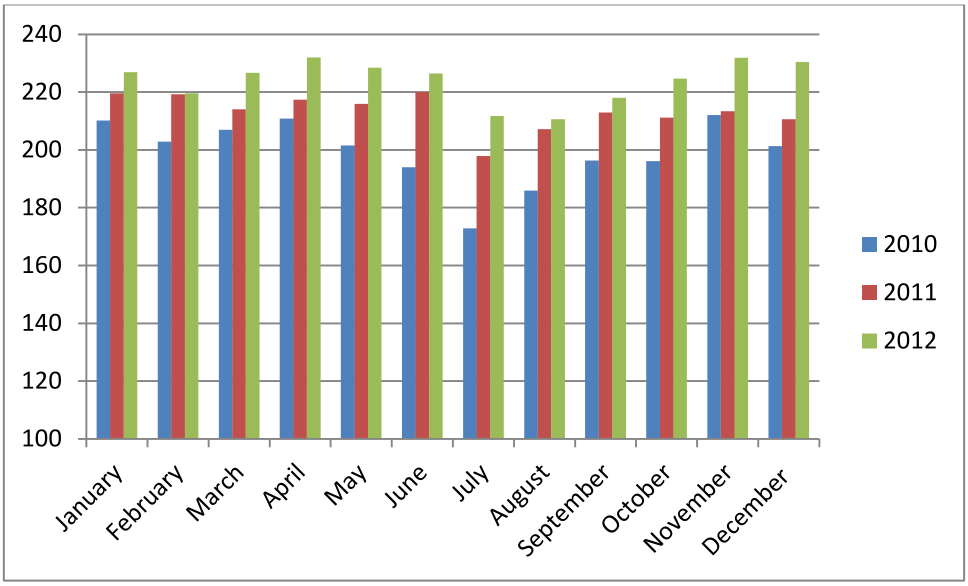

Figure 4: average number of admissions on a weekday for each month

This figure only shows the admissions, the pattern for discharges is shown in Appendix B and is quite similar. The admissions are shown as average number of admissions on a day instead of showing the admissions for the whole month. This is because the Dutch holidays are deleted from the data and this would lower the number of admissions a lot in months with a lot of holidays. To keep the numbers meaningful also the weekends are left out and the averages are thus calculated only over the weekdays. Another reason for showing the average number of admissions on a weekday is that some Dutch Holidays can be in a different month each year. Notice that the average number of admissions on a weekday in December is a bit distorted, because 24 December until 31 December is deleted from the data.

Again the trend upwards is easily recognized. Also the summer months (June, July and August) seem to have a significant lower amount of admissions in 2010, but this pattern is less recognizable in 2012. It seems that especially in the summer months the production has increased a lot over the years. That is why for 2012 the differences in the months have become a lot smaller.

To conclude; especially the day of the week and somewhat less the month of the year seem to have an influence on the number of admissions and discharges. That is why these variables will be used in the next chapter when making forecasts.

6. Forecasting

In the previous chapter historical data has been analyzed. However, the main goal is to get information for creating a crew schedule for the bed cleaning department for the future. In this chapter try different methods of forecasting the number of admissions and discharges will be tested and the results will be compared. To be able to compare the result of the different methods the data will be split in two parts. The first part, data of 2010 and 2011, will be used as training data set and forecasts will be made for the second part, data of 2012. The data of 2012 will be used to compare the results of the different models. The best model will then be used to forecast the number of admissions and discharges for 2013 by using the data of 2010, 2011 and 2012.

6.1 The models

In total five different models will be tested: three linear regression models, a non-linear regression model, and an autoregressive(AR) model. In explaining the models we will focus on the admissions, but the same analysis is done for the discharges.

6.1.1 Linear regression model

In the data analysis we concluded that the day of the week seems to play an important role in number of admissions. Therefore the first and most simple model which will be used is a linear regression model in which the admissions are taken as dependent variable and the days of the week as independent variables. Also a trend is taken as independent variable, because there was an obvious trend noticed in the data analysis. Since this trend differs a lot per day, a separate trend is added for each day. This gives us the following regression model:

$$y = \sum_{i=1}^{7}(\alpha_i Day_i + \beta_i Trend_i) + \varepsilon \qquad (6.1)$$

In which y is a vector of admissions, Day_i is the dummy variable for day i of the week and $Trend_i$ is the trend vector for day i of the week. This trend vector can be created by multiplying the dummy variables Day_i with a number from 1 to 154 depending on what number of week it is in the two years of data. The vectors α and β will both contain seven estimated parameters, one for each day. The vector of errors (ε) are white noise errors, which means that the elements (ε_t) of this vector have the three following properties:

- the mean of each ε_t is zero
- the variance of each ε_t is the same
- there is no correlation between any past, current and future ε_t.

Another conclusion of the data analysis is that admissions could also depend upon the month. This gives us the following second model:

$$y = \sum_{i=1}^{7}(\alpha_i Day_i + \beta_i Trend_i) + \sum_{j=1}^{12}(\gamma_j Month_j) + \varepsilon \tag{6.2}$$

where $Month_j$ is the dummy variable for the month of the year and γ_j the corresponding estimated parameter.

We even took it one step further in details by adding dummy variables for the weeks in the last regression:

$$y = \sum_{i=1}^{7}(\alpha_i Day_i + \beta_i Trend_i) + \sum_{j=1}^{12}(\gamma_j Month_j) + \sum_{k=1}^{53}(\delta_k Week_k) + \varepsilon \tag{6.3}$$

Using variables in a model that do not have any influence on the outcome, the parameter value is zero, causes a loss of efficiency of the relevant variables. That is why for each parameter it is tested if it differs significantly from zero by using a t-test. This test gives a P-value which is the probability of obtaining the estimated parameter value or a more extreme value while assuming that the parameter is equal to zero. If there are variables which have a P-value larger than 0.10 the following rule of deletion is used: the variable with the largest P-value is deleted from the model and the regression is done again and a variable is deleted until no variables with a P–value larger than 0.10 are left over.

6.1.2 Non-Linear regression model

A non-linear regression model might give better results than only linear regression models. To create a non-linear regression model an important variable can be taken as squared in the model. Therefore also a quadratic trend is added to the first model:

$$y = \sum_{i=1}^{7}(\alpha_i Day_i + \beta_{1,i} Trend_i + \beta_{2,i} {Trend_i}^2) + \varepsilon \tag{6.4}$$

In the other models the quadratic trend is not significant, so adding the quadratic trend does not make a change since it is deleted again by the deletion rule.

6.1.3 AR-model

In an autoregressive model the independent variable is modeled as a random process in which the variable depends upon its own historical values and the random error term ε_t. An AR-model where the value only (directly) depends upon the last two observations is often formulated as follows:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t \tag{6.5}$$

where c is a constant, φ_1, φ_2 are the parameters of the model and ε_t are again the white noise error terms.

A shock in the process has influence on all observations that come after that shock. However, we assume it is a stationary process, $\varphi_1 < 1$, which means the effect will extinguish with time. To find out which previous observations, which i's in y_{t-i} , should be taken in the model the autocorrelations are used. The autocorrelation, ρ_k, can be calculated for each k-th observation back in time, also called k-th lag. After fitting an AR-model the residuals (e) should contain no significant autocorrelation any more. This means a good AR-model adequately captures all the autocorrelation properties of the data.

In the AR-model the outlier dates need to be included again, because if no values are known for these dates they will disturb the pattern. However, including the actual values of the outliers will have a big negative effect on the results. That is why a method called Winsorizing (Wilson, 1997) is used, which means the outliers are replaced by the most extreme value in the sample for that day of the week in that month. By using this method the outliers are less influential, but not completely discarded.

By analyzing the autocorrelation of the admissions we see that significant autocorrelation is present in all lags. Since the admissions will mostly depend upon the admissions of that same day a week ago, the y_{t-7} is included in the model next to a constant and trend:

$$y_t = \alpha + \beta t + \varphi_7 y_{t-7} + \varepsilon_t \qquad (6.6)$$

The main goal is to eliminate the autocorrelation from the residuals, and therefore the autocorrelations of these residuals are calculated. Only the autocorrelation of the 7-th lag is still significant now. By adding y_{t-14} the 7-th lag is also not significant anymore. This means that observations seem to depend mostly on observations on the same day of the week. That is why the following model is taken:

$$y_t = \alpha + \beta t + \sum_{i=1}^{k} \varphi_{7i} y_{t-7i} + \varepsilon_t \qquad (6.7)$$

The last step is to determine k. Therefore the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC) are used. These criteria evaluate the number of in-sample fit while taking the number of estimated parameters as a negative aspect. In table 3 the results of the criteria are shown for different values of k.

Criter. \ k	**0**	**1**	**2**	**3**	**4**	**5**	**6**
AIC	11.512	8.654	8.431	8.380	8.301	8.301	8.305
SIC	11.524	8.673	8.457	8.412	8.341	8.347	8.358

Table 3: AIC and SIC of different AR-models

The k that minimizes these criteria should be taken. For both criteria this means taking $k = 4$. Therefore the AR-model that will be used when making forecasts is the following model:

$$y_t = \alpha + \beta t + \varphi_7 y_{t-7} + \varphi_{14} y_{t-14} + \varphi_{21} y_{t-14} + \varphi_{28} y_{t-14} + \varepsilon_t$$

6.2 Results

The performance of the models can be evaluated on two different criteria:

- How good does the model fit the 2010 and 2011 data?
- How good does the model forecasts the 2012 data?

The first question can be answered by calculating the sum of squared residuals (SSR) and the total sum of squares (SST) and use these two values to calculated the coefficient of determination (R^2) of the model. These measurements are defined as follows:

$SSR = \sum_i (y_i - f_i)^2$ where f_i is the fitted model value (6.8)

$SST = \sum_i (y_i - \bar{y})^2$ where $\bar{y}$ is the sample mean (6.9)

$R^2 = 1 - \frac{SSR}{SST}$ (6.10)

In the following table these values are shown for the different models:

Model	SSR	SST	R^2
1. LR Days	179390.0	4069308	0.9559
2. LR Months	143669.0	4069308	0.9649
3. LR Weeks	116861.5	4069308	0.9713
4. NLR Days	178953.5	4069308	0.9560
5. AR	162856.8	4102186	0.9603

Table 4: evaluation of the in-sample fit

The values of the coefficient of determination are close to each other and all above 0.95, which means that more than 95% of the total variance of the data is explained with these models. However, a larger R^2 does not automatically mean a better model, since R^2 never decreases when adding variables to the model and will often increase. This is also the reason why the linear regression model with the week variables has the highest R^2. The main conclusion that can be made is that all models fit the data quite well.

The next step is to make forecasts with the models and evaluate them. In the AR-models two different methods of forecasting will be used: dynamic and static. By using static forecasts a one-step ahead forecast is calculated one by one while using the actual values rather than the forecasted values for the lagged dependent variables. This way of forecasting is only possible when the actual data is known, so not for a period for which the data of the dependent variable is not available. To make forecasts without data of the dependent variable is possible with the dynamic method of forecasting, because multi-steps forecasts are calculated. To calculate the forecasts in future periods in the dynamic method the forecasted values for the lagged dependent variables are used.

To evaluate the forecasts the following statistics are calculated: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the Theil Inequality Coefficient (TIC). The definitions of these statistics are as follows:

n: number of forecasted values

$$e_i = (y_i - f_i) \tag{6.11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_i}^2} \tag{6.12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |e_i| \tag{6.13}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} |e_i/\mathrm{y}_i| \tag{6.14}$$

$$TIC = \frac{MSE}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} {f_i}^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n} {y_i}^2}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_i}^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} {f_i}^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n} {y_i}^2}} \tag{6.15}$$

The *RMSE* and *MAE* depend upon the scale of the dependent variable while the *MAPE* and *TIC* do not depend upon that. Since in our models the scale of the dependent variable is the same, all four statistics can be used to compare the models. The $RMSE$ and MAE should be as close as possible to zero since this means that the error is small and the model forecasts accurately. A practical interpretation of the MAE is as follows: it shows how much on average the difference is between the forecasted number of admissions and the actual admissions, i.e. how many admissions there are on average too much or too few forecasted on a day. The $MAPE$ is a percentage and should as close as zero as possible as well. The TIC measures the degrees to which the forecasts differ from the actual values. The TIC is always between zero and one, where zero indicates a perfect fit.

Model\ Meas.	$RMSE$	MAE	$MAPE$	TIC
1. LR Days	16.90	14.14	11.97	0.0446
2. LR Months	16.52	13.80	14.99	0.0429
3. LR Weeks	16.27	12.80	9.14	0.0427
4. NLR Days	17.31	14.73	14.79	0.0451
5. AR static	14.76	11.26	8.50	0.0384
6. AR dynamic	20.36	16.49	12.33	0.0547

Table 5: evaluation of the out-of-sample fit

The AR model with static forecasts definitely outperforms the other models. However, by doing static forecasts the model also makes use of the data of 2012, which the other models do not. This is also the reason why the static method of forecasting cannot be used for forecasting the admissions and discharges for 2013.

The second best model is the LR weeks model, which has a better value for all measurements than the other models. The difference between the measurements are not that large for the models LR weeks and the best model, AR static, so the using the LR Weeks model to forecast the admissions for 2013 will satisfy the requirements. The same analysis is done for the discharges and in this case the LR Months model was slightly better for each measurement than the other models (except for AR static). In table 6 and 7 the forecasts of admissions and discharges are shown.

Notice that no forecasts are made for the days that were excluded out of the data. The main reason for this is that most of these days heavily depend upon the type of Dutch Holiday and on which day the holiday is that year. Even if it is on the same day every year the difference over the years can be big without any recognizable pattern in it.

Days\Months	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	------	201	198	------	236	46	259	208	59	246	204	59
2	245	54	51	254	221	59	238	177	267	247	50	268
3	232	68	65	254	189	264	239	26	246	230	67	248
4	198	264	268	240	40	243	223	47	247	200	276	250
5	53	242	246	207	58	243	191	254	230	46	255	231
6	66	241	245	59	262	228	41	233	199	58	257	202
7	268	227	231	66	241	196	48	234	47	267	239	47
8	246	194	199	270	241	46	254	218	55	246	209	59
9	245	48	51	248	------	64	233	187	263	248	55	269
10	232	61	61	248	------	270	234	35	242	230	58	249
11	199	271	265	234	45	248	218	53	243	200	267	251
12	53	249	243	201	59	249	186	260	226	46	247	232
13	67	248	242	53	264	233	35	239	196	52	249	203
14	273	234	228	68	242	202	42	240	43	260	231	48
15	251	201	196	272	242	52	249	223	55	239	201	55
16	250	54	48	250	227	59	227	193	263	241	46	265
17	237	68	67	250	195	265	228	41	242	223	62	245
18	203	261	271	235	46	243	212	55	243	193	271	247
19	57	239	249	203	------	244	181	263	226	40	251	228
20	71	239	249	55	------	228	30	241	196	58	253	199

21	269	225	235	69	256	197	40	243	43	267	234	44
22	246	192	202	274	257	46	247	226	55	246	205	70
23	246	45	54	252	241	59	226	195	263	248	50	280
24	232	58	68	252	210	265	227	43	242	230	54	260
25	199	251	272	237	60	243	210	61	244	200	264	------
26	53	229	250	205	79	244	179	268	226	47	243	------
27	67	228	249	56	285	228	28	247	196	53	245	------
28	269	214	235	63	263	197	31	248	43	262	227	------
29	247		203	267	263	46	238	231	55	241	197	------
30	246		54	------	248	59	217	201	263	243	43	------
31	233		------		216		218	48		225		------

Table 6: forecasted number of admissions for 2013

Days\Months	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	------	234	235	------	240	72	214	220	43	228	237	47
2	240	75	75	237	228	47	217	224	225	242	74	232
3	229	50	50	248	233	226	229	63	227	228	48	233
4	235	225	226	236	73	230	216	37	240	232	232	248
5	76	231	231	241	47	241	221	218	226	69	233	233
6	50	240	241	81	225	229	59	221	231	43	248	237
7	225	229	230	55	230	234	33	234	69	227	233	73
8	231	235	235	233	241	73	214	220	43	228	237	47
9	240	75	75	237	------	47	217	225	225	242	74	232
10	229	50	50	248	------	226	229	63	227	228	48	233
11	235	225	226	236	73	230	216	37	241	232	232	248
12	76	231	232	241	47	242	221	218	227	69	233	233
13	50	241	242	81	226	229	59	221	231	43	248	237
14	225	229	230	55	230	234	34	234	69	227	233	74
15	231	235	235	233	241	73	214	220	43	228	237	47
16	240	75	75	238	229	47	217	225	225	243	74	233
17	230	50	50	248	233	227	230	63	227	228	48	233
18	235	226	227	236	73	230	216	37	241	232	232	249
19	76	231	232	241	------	242	221	219	227	70	233	233
20	50	241	242	81	------	229	59	221	231	44	248	237
21	226	230	230	55	230	234	34	234	69	227	233	74
22	232	235	235	233	241	73	214	220	43	228	237	47
23	241	75	75	238	229	47	217	225	225	243	74	233
24	230	50	50	249	234	227	230	63	227	228	48	233
25	235	226	227	237	73	230	217	37	241	232	233	------
26	76	231	232	242	47	242	221	219	227	70	233	------
27	50	241	242	81	226	229	60	221	231	44	249	------
28	226	230	231	55	230	234	34	235	69	227	233	------
29	232		236	234	242	73	215	221	43	229	238	------
30	241		76	------	229	47	217	225	226	243	74	------
31	230		------		234		230	63		228		------

Table 7: forecasted number of discharges for 2013

We now have point forecasts but we also want to know the estimated range of values which is likely to include the actual value. Therefore 95%-intervals can be created by using the standard error of the forecasts. In the next table some statistics of the standard error of forecasts are shown:

	$Mean$	$Minimum$	$Maximum$
Admissions	14.9	14.5	15.7
Discharges	16.1	15.8	16.2

Table 8: statistics of standard error of forecasts

By rounding off these averages and assuming the standard error of forecasts are normally distributed, we can state that a estimation of the 95%-intervals will be created by adding and subtracting $2 \cdot 15 = 30$ and $2 \cdot 16 = 32$ for the admissions and discharges respectively. The forecasts for all admissions and discharges of patients in the hospital are constructed and in the next chapter they will be linked to the tasks for the bed cleaning department in I-transport and finally to a number of workers needed on a day.

7. Bed cleaning department

The bed cleaning department wants to know how many tasks they can expect in the upcoming period. These numbers can be used to create the crew schedule. Although we have constructed forecasts for all admissions and discharges, not every admission needs a bed. The reason is that some treatments do not need a bed, but also some beds are cleaned at the clinic by a nurse which means it does not have to be done by the bed cleaning department. These cases will not create a task in I-transport and need to be subtracted from the forecasts to get a good estimate of the number of tasks on a day in I-transport. However, it is unknown which part of the total admissions needs a bed and which do not. Most doctors in the hospital probably have a feeling of how many of their admissions need a bed, but these numbers are unknown at a hospital level. Therefore we want to create different possible scenarios for the number of tasks based upon the forecasts of admissions and discharges. The correct scenario can then be chosen when this missing data becomes available.

7.1 Scenarios

In creating the scenarios the opening hours of the bed cleaning department are relevant. The opening hours for the bed cleaning department are as follows:

- Monday until Friday: 7.30 – 16.30
- Saturday: 8.30 – 12.30
- Sunday: closed

The tasks for admissions and discharges on Sunday have to be done on a different day, since the bed cleaning department is closed. The tasks for admissions have to be done earlier, because a clean bed is needed for an admission. The tasks of discharges have to be done later, because a bed is ready to be cleaned after the discharge. However, it is unknown on which day these tasks will be done and therefore different options are possible.

One option is to add all the admissions to the Saturday and all the discharges to the Monday. However, the bed cleaning department is only open in the morning on Saturday and therefore it probably cannot handle all admissions for Saturday and Sunday. It is also known that clean beds for admissions in the weekend are often ordered on Friday. This gives us the idea of a better option for the admissions on Sunday: adding a fraction ρ of these admissions to Saturday and add the other part, 1-ρ, to Friday.

To find out which ρ best fits the data, the historical data of I-transport is used. In total two months of data is used: 53 days (no Sundays). For each day the number of tasks completed by the bed cleaning department is known. With this data the historical percentages of admissions and discharges that need a bed, thus actually create a task for the bed cleaning department, can be calculated. The following formula shows how to calculate the percentages:

$$p_t = \frac{\#tasks_t}{\#admission_t + \#discharges_t} \cdot 100 \qquad \text{where } t = 1, \dots, 53 \qquad (7.1)$$

The next table shows the average of the percentages for the different days without doing anything with the admissions and discharges of Sundays.

	Monday	***Tuesday***	***Wednesday***	***Thursday***	***Friday***	***Saturday***
Avg. percentage	56.08	51.47	49.96	48.84	52.71	67.23

Table 9: average percentages of tasks from total admissions and discharges for each day without data of Sunday

It seems likely that the percentages of forecasts that need a bed do not differ much over the days. Therefore the assumption is made that these percentages are the same for each day. The historical data supports this assumption, because the average percentages for Tuesday, Wednesday and Thursday differ at most $51.47 - 48.84 = 2.63$. This also explains why the percentages of Monday, Friday and Saturday differ more, since for these days the admissions and discharges for Sunday should still be added. In evaluating which ρ fits the data the best, the average of the average percentages of Tuesday, Wednesday and Thursday can be taken as a benchmark: $(51.47 + 49.96 + 48.84)/3 = 50.09$.

The next step is to find ρ which minimizes the sum of the absolute differences between the average percentages of Friday and Saturday compared with the benchmark value. This value is minimized to 0.263 by taking $\rho = 63.72$. This means that the following transformation of the data is done: all discharges are added to Monday, 63.72% of the admissions on Sunday is added to Saturday and 36.28% is added to Friday. In the next table the average percentages are shown after this data transformation on the historical data is performed.

	Monday	***Tuesday***	***Wednesday***	***Thursday***	***Friday***	***Saturday***
Avg. percentage	52.00	51.47	49.96	48.84	50.35	50.09

Table 10: average percentages of tasks from total admissions and discharges for each day with data of Sunday

Table 10 shows that the average percentages are all between 48.84% and 52.00%, which means the maximum difference is only 3.16%. Therefore this way of handling the admissions and forecasts of Sunday fits the historical data well.

Now the mean percentage and the standard deviation of all 51 days can be calculated, which are respectively 49.5% and 6.5%. A lot of calculations and estimations were needed to get to the 49.5%, which means there is a lot of uncertainty in this estimated value. This is one of the reasons to create different scenarios based upon the mean percentages and its corresponding standard deviation.

Another reason is to be able to translate the forecasts in a number of tasks to get a feeling of the influence of different scenarios on the number of forecasted tasks on a day.

Each scenario uses a different percentage (p) to translate the point forecasts made in chapter 6 into the number of tasks. In the first scenario the mean is used, in the second scenario the standard deviation is added to the mean and in the third scenario the standard deviation is subtracted from the mean:

- Scenario 1: $p = 49.5\%$
- Scenario 2: $p = 56.0\%$
- Scenario 3: $p = 43.0\%$

For each scenario the number of expected tasks for each day in 2013 can be calculated by taking the percentage p of the forecasts after adding up the forecasted admissions and discharges and then do the data transformation with $\rho = 63.72$. The following table shows the statistics for the forecasted tasks for each scenario over the different days in 2013:

	Scenario 1			***Scenario 2***			***Scenario 3***		
	μ	***min***	***max***	***μ***	***min***	***max***	***μ***	***min***	***max***
Monday	332	247	371	376	280	419	288	215	322
Tuesday	234	215	249	265	243	281	204	187	216
Wednesday	241	222	289	273	251	327	209	192	251
Thursday	227	211	236	256	239	267	196	194	206
Friday	256	210	288	290	237	325	223	182	250
Saturday	102	56	129	115	63	146	88	48	112

Table 11: statistics of forecasted number of tasks for the different scenarios

The Monday seems to be the busiest day for the bed cleaning department, which could be expected since the bed cleaning department is closed on Sunday. All the dirty beds from discharges on Sunday will be cleaned on Monday. The minimum number of tasks on Monday is higher than the mean of all other days, except for Friday, which shows Friday is also a busy day. The difference in the means of scenario 2 and scenario 3 has a maximum of $376 - 288 = 88$. To get an idea of the impact of these differences for the crew schedule, we need to investigate the productivity of the workforce, i.e. how many tasks can a worker finish on a day?

7.2 Workforce

The workforce of the bed cleaning department is responsible for transporting clean and dirty beds from and to the bed cleaning department. From now on we will consistently refer to a male person when talking about a worker, because the majority of the crew are men. In our further research we exclude the Saturday and focus on Monday until Friday. A working day for a worker then contains eight working hours. In these eight hours two coffee breaks are included of 15 minutes. These coffee breaks are around 9.30 and 15.30, depending upon the time a worker finished his last task. A worker always finishes his current task before taking a break. Every worker also has a lunch break of 30 minutes which is around 12.15 and the lunch break is not included in the eight working hours.

Currently the schedule of the workers is constructed in such a way that 6 to 8 workers are planned on a day depending upon the availability of the workers. By analyzing the data of I-transport we see that on average 7.3 workers are working from opening until closing time, with a minimum number of 5 and a maximum of 9 workers.

In the eight hours of work the workers are not only busy with tasks. For example, after a worker delivered a bed (finished a task) he sometimes cleans up the room and put the bed on the right spot. Also before a worker accepts a task he sometimes makes the bed ready, which takes time which is not included in the time which is spent on tasks. By again analyzing the data of I-transport we see that workers are busy with tasks for 79.7% of the 8 hours. The average number of tasks by a worker completed in one hour of only doing tasks is 4.82, which means a task takes on average 12.45 minutes. Combining these numbers gives us the average number of task a worker will finish in a day: $0.797 \cdot 8 \cdot 4.82 = 30.7$ tasks. Table 12 shows all the statistics of the historical data for the crew of the bed cleaning department:

$Average\ \#\ of\ workers/day$	$Minimum\ \#\ of\ workers/day$	$Maximum\ \#\ of\ workers/day$	$\%\ time\ busy\ with\ tasks$	$Average\ \#\ of\ tasks/hour$	$Average\ time\ on\ one\ task$	$Average\ \#\ of\ tasks/day$
7.3	5	9	79.7	4.82	12.45 min	30.7

Table 12: statistics of historical data for crew of the bed cleaning department

The 95% forecast intervals for the admissions and discharges were created by adding and subtracting 30 and 32 respectively to and from the point forecasts, which means approximately one worker more or less on a day if every admission or discharge creates a task. However, not every admission or discharge creates a task, so the exact number of workers needed if the boundaries of the 95% interval are taken, depends upon on which scenario is picked. This shows that when the actual number of admissions on a day is the boundary of the forecast interval instead of the value of the point forecast, only one worker more or less could be needed. Out of this we can conclude that the forecasts are useful in planning the number of workers on a day, because even in one of the worst case scenarios (number of admissions is the boundary of the forecast interval instead of the point forecast) it is a difference of only one worker.

Next to the absolute difference also the relative difference in a worst case scenario can be interesting. Rounded down there are on average 7 workers working on a day. If one worker more or less is needed this will be a relative difference of $1/7 \cdot 100\% = 14.3\%$. This shows that when one worker too few is planned for a working day, all tasks could still be finished if all workers together can compensate for this 14.3%.

The average number of tasks finished by a worker on a day, 30.7, can be used to calculate the number of workers needed in a day. The statistics calculated for each day in section 7.1, see table 11, are divided by 30.7 to get the number of workers needed. This gives us the following statistics for the number of workers and for each scenario:

	Scenario 1			***Scenario 2***			***Scenario 3***		
	µ	***min***	***max***	***µ***	***min***	***max***	***µ***	***min***	***max***
Monday	11.1	8.3	12.4	12.2	9.1	13.6	9.9	7.4	11.1
Tuesday	7.8	7.2	8.3	8.6	7.9	9.2	7.0	6.4	7.4
Wednesday	8.0	7.4	9.6	8.9	8.2	10.7	7.2	6.6	8.6
Thursday	7.6	7.0	7.9	8.3	7.8	8.7	6.7	6.3	7.1
Friday	8.5	7.0	9.6	9.4	7.7	10.6	7.7	6.3	8.6

Table 13: statistics of historical data for the number of workers needed based upon forecasted number of tasks

These statistics show that on a Monday at least three more workers are needed compared to the other days. The difference in the mean number of workers on a day between the different scenarios can be as large as $12.2 - 9.9 = 2.3$, which shows that choosing the right scenario is crucial when using these statistics for making crew schedules. Currently on average 7.3 workers are scheduled on a day, which fits scenario 3 the best. However, currently a lot of tasks are not completed on time and this would imply that there is under capacity and more workers should be scheduled on a day, which makes scenario 1 or maybe even scenario 2 more likely. After the right scenario has been chosen, the estimates for the number of workers needed for every day of the year can be calculated and used in making the crew schedules.

Until now we have focused on the whole year 2013 by making forecasts for the admissions and discharges and translated these forecasts in tasks and finally in a number of workers needed. In the next chapter we will take a closer look at the tasks on a daily level and focus on the way the tasks are assigned to workers.

8. Simulation: the model

This chapter is about the simulation to analyze the way of assigning tasks to workers and different options which can be configured in I-Transport. The simulation model is built to simulate and analyze a single day of assigning tasks to workers. Example days from the historical data of I-transport will be used for the simulations. This data will contain all information about the tasks which are finished on these example days.

The duration of a task depends upon two parameters: the start/end location of the task and the sort of task. In total around 1200 specific addresses are located in the Centrum location of the hospital at which beds can be delivered. Based upon their location in the hospital these addresses can be divided in 38 different zones. These zones are used to get the transportation time of a task from one address to another address. Each task starts or ends at the bed cleaning department, which means only the other zone (z) than the bed cleaning department is relevant for estimating the transportation time. Using these zones as destinations answers the research question of how to aggregate the wards in the hospital.

A task can be to deliver a clean bed or to pick up a dirty bed. A worker can also handle two beds, which means two tasks can be combined when two beds have the same start and end zone. In that case the planned transportation time will be longer, because only one bed at a time can be transported. So if a worker has a grouped task of two tasks, the two beds can for example go together in the elevator, but have to be transported separately through the hall to the end point. The last and fifth task sort is cleaning beds at another location: the Thorax clinic. These beds have to be cleaned at the Thorax clinic instead of at the bed cleaning department. This task takes half an hour and has to be done once in the morning and once in the afternoon by two workers.

The estimated duration of a task depends upon the sort of task (s): task of delivering (1) or picking up (2) a bed, grouped task of delivering (3) or picking up (4) two beds and cleaning at the Thorax clinic (5). Although the transportation times for option 1 and 2 are probably almost the same, since transporting a clean bed one way and a dirty bed the other way will not differ much, they are still stored separately and used separately in the system depending upon the task sort. The same holds for task sort 3 and 4.

The duration of each finished task from or to the other zone than the bed cleaning department is stored and the average time of all these tasks finished in the history is taken as an estimation for the duration of a task from or to that zone. These estimated transportation times are stored in vectors

(v_s) for each sort of task (s) and each vector contains transportation times for each start or end zone other than the bed cleaning department (z).

Each new registered task has a planned arrival time (t_{plan}) and a start or end zone (z). The latest start time of a task (t_{lat}) is equal to the planned arrival time (t_{plan}) minus the estimated duration of a task to that zone ($v_s(z)$). The estimated actual arrival time (t_{act}) of a task is the earliest time a worker is finished plus the duration of the task, so basically it is the time the task will be finished if it is the next task to be assigned.

Whenever a worker receives a task on his mobile phone he first needs to go the start location if that is not his current location. The transportation times to get to the start location are fixed times based upon the number of floors he has to go up and down and the distance he has to cover. The assumption is made that a worker without a bed covers 70 meter in one minute which means 4.2 km in one hour. Depending on the number of floors up or down extra time is added for the taking the elevator and waiting for the elevator. The time added is 1 minute for one or two floors, 2 minutes for three to six floors and 3 minutes for more than 6 floors. These transportation times are stored in the start matrix ($startMatrix$), see Appendix C.

The process of a worker who finishes two tasks is shown in a flow chart in figure 5.

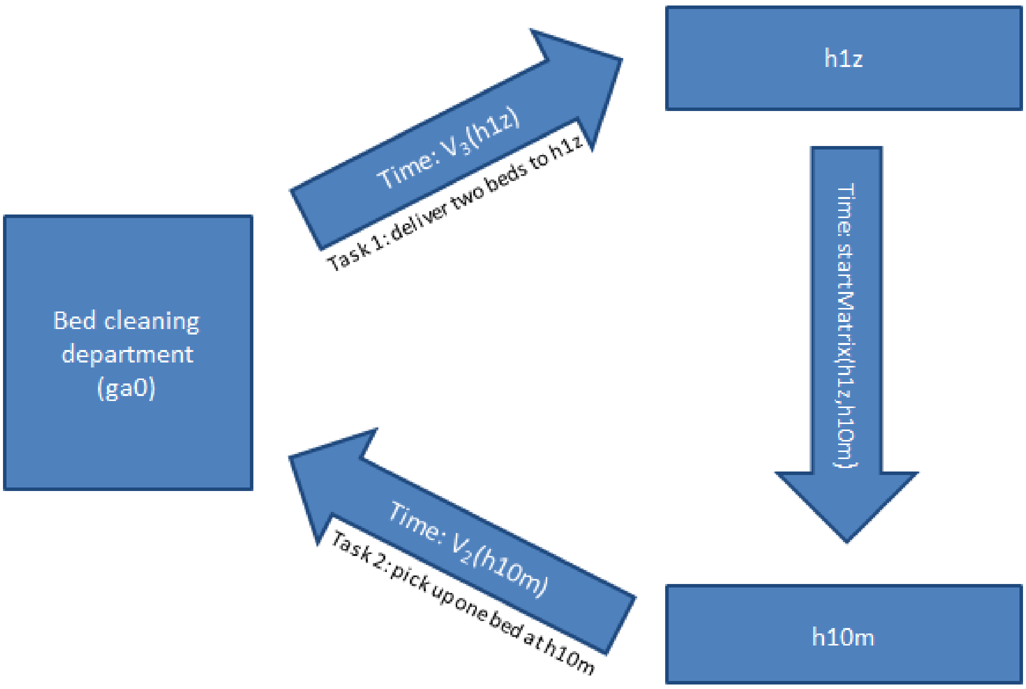

Figure 5: flow chart of the process of finishing two tasks

The worker starts at the beds zone and receives a task for delivering two clean beds, task sort 3, to a ward in zone h1z. The transportation time for this task is taken from the vector v_3. When the worker finished the task he uses a mobile phone to tell the system he is ready for a new task. The next task which is assigned to him is a task of picking up a bed at a ward in zone h10m and deliver this bed to the bed cleaning department. Therefore the worker first has to move to zone h10m without a bed and the transportation time needed for this is taken from the start matrix. When he gets there the dirty bed is picked up and delivered to the bed cleaning department, where the worker tells the system again he is ready for a new task. Notice that a new task could also be again picking up a task, which means he has to move to the start location of that task without a bed again.

When a worker has finished a task of delivering a bed to a zone, the system always checks if a bed from the same zone can be picked up, since the worker has to go back to the bed cleaning department in most cases anyway. Also when no more clean beds are left over at the bed cleaning department, tasks of picking up beds have to be done. If this happens in practice the workers often just pick up dirty beds from anywhere in the hospital without registering them, which makes it hard to take a good day with finished tasks as example day for the simulation. The same happens after the bed cleaning department is closed, because two workers will continue working after 16.30 and will only pick up dirty beds. This also means that there always will be clean beds ready to deliver in the morning. However, this picking up of beds which are not registered in the system makes the data less reliable.

Although we want to stay as close to the reality as possible we need to make some assumptions for this simulation which will be shown in the first section. In the second section the different definitions are shown and in the third section is explained how the simulation model is implemented. In the fourth section of this chapter the different options for the model will be explained and in the final section we explain the sensitivity analysis that will be done.

8.1 Assumptions

1. The circulation of beds is a push system, which means storage of the beds take place at the bed cleaning department instead of everywhere in the chain.

2. The bed cleaning department starts and ends the day with a fixed amount of clean beds b that are ready to be delivered.

3. Dirty beds are cleaned in exactly five minutes and this is done by cleaning machines, which means that the workers of the bed cleaning department do not have to spend any time on cleaning beds.

4. Workers accept tasks until 9.15, 12.00 and 15.15. Then they first finish their tasks and have a break for respectively 15, 30 and 15 minutes.

5. Two workers will be busy from 7.30 until 8.00 and from 14.00 until 14.30 with cleaning beds at the Thorax clinic.

6. Every task gets registered in the system two hours before the planned arrival time, which means a task is unknown until two hours before it has to be finished.

7. Workers will always be busy when there are unassigned tasks in the system, because every task is allowed to be completed before the planned arrival time.

8. There is only one type of bed, which is the most common used bed called PLANO, which means all tasks for other beds will be seen as a PLANO bed as well.

9. All addresses in the same zone have the same transportation times.

10. A worker is only 79.7% of his duty busy with completing tasks, this percentage comes from the historical data (see section 7.2).

11. A worker covers 70 meters in a minute in the hospital when he is moving empty handed.

8.2 Definitions

Input variables:

w: number of workers
b: number of beds available at the start of the day

Model variables:

z: start/end zone which is not the bed cleaning department
s: tasks sort (=1,2,3,4 or 5)
v_s: vector of transportation times for task sort s, where $v_s(z)$ is the transportation time for zone z
$startMatrix$: matrix of travel times without a bed from each zone A to another zone B
n_{tasks}: number of tasks on a day
t_{plan}: planned arrival time of task
t_{lat}: latest start time of a task = t_{plan} - $v_s(z)$
t_{act}: actual arrival time of a task
$late$: lateness of a task = $t_{act} - t_{plan}$

$tard$: tardiness of a task $= \max(late, 0)$

n_{late}: number of tasks finished late

n_{on}: number of tasks finished within one hour

n_{fin}: number of finished tasks $= n_{late} + n_{on}$

Output variables:

P_{fin} = percentage of finished tasks $= \frac{n_{fin}}{n_{tasks}} \cdot 100$

P_{late} = percentage of finished late tasks $= \frac{n_{late}}{n_{fin}} \cdot 100$

P_{on} = percentage of finished tasks within one hour $= \frac{n_{on1}}{n_{fin}} \cdot 100$

$T =$ average late time of a late task $= \frac{1}{n_{late}} \sum_{i=1}^{n_{fin}} tard_i$

M = maximum lateness of task $= \max(tard_i)$

8.3 Implementation

We start by taking an example day with tasks out of the historical data from I-transport. For each task the following information is known: planned arrival time (t_{plan}) and the task sort (s). The estimated duration time of a task ($v_s(z)$) can be taken from the transportation time vectors and depends on the start and end zone. The next section gives an answer to the following research question: How can we estimate the total time needed to transport a bed to (or from) a specific ward from (or to) the bed cleaning department?

8.3.1 Transportation times

In the simulation model the average of the historical transportation times are used which are shown in table 14. In reality these transportation times will fluctuate, but in the model they will be kept fixed. The reason for this is that the data behind these averages is unknown and therefore no distribution can be fit to this data and use this distribution to sample the variation. The impact of using a distribution to sample the variation in the transportation times will be discussed in chapter 11: Discussion.

The bed cleaning department is shown as zone **ga0**. Since a task always starts or ends at the bed cleaning department, there are no transportation times for zone **ga0**. However, more times are missing (*****) in table 14. For transportations to these zones no tasks of that type have been registered yet. For example, there are no tasks in the history of picking up beds at the zones **h1z** and **h2m.** Also there are no grouped tasks finished yet to or from the zone **h2n**. The reason could be that are not many tasks for **h2n** and therefore there will not be two tasks to/from this zone with similar planned arrival times, which means tasks to/from this zone cannot be grouped.

Table 14 also shows a difference in transportation times between picking up beds and delivering beds. Although the differences are only small for most zones, this does not hold for every zone. For example, for zone **bd4** the transportation time of picking up a single bed is twice as large as the time of delivering a bed. The reason for this big difference is unknown and this might be because only a few months of historical data is used for these transportation times. It could be that the pick-up time for zone **bd4** is based upon only one (or a few) task(s) and these task(s) happened to take longer than normally because the hallway was very busy on that day.

In the model we do not change any of these times that might be wrong, since the system that is now running uses these same times. In the future this could be further investigated and implemented in another way. For example, only take the average transportation time if there are at least 5 tasks and otherwise just take the transportation time form the other direction or from the start matrix. The reason that only few tasks are registered for some zones might be that new transportation times are not calculated automatically in the system. The system only calculates new values when someone explicitly tells the system do this and this happens only scarcely in reality.

Zones	V_1	V_2	V_3	V_4
ba0	8	6	*	*
bd1	10	10	*	*
bd2	10	10	*	*
bd3	16	12	*	*
bd4	11	22	*	*
ca2	12	9	*	*
d3	10	7	14	6
dp1	20	19	*	*
dp2	16	16	*	*
dp3	15	15	*	*
ga0	**	**	**	**
h1z	6	*	16	*
h2m	6	*	16	*
h2n	8	10	*	*
h3m	7	6	11	8
h3n	8	7	14	10
h3z	12	11	15	14
h4m	7	5	11	8
h4n	8	6	14	10
h4z	9	6	14	11
h5m	8	7	14	10
h5n	8	6	14	10
h5z	9	6	16	9
h6m	10	9	17	*
h6n	10	6	15	10
h7m	8	10	14	10
h7n	9	8	16	13
h7z	9	9	14	10
h8m	8	8	14	10
h8n	9	11	16	10
h9m	10	9	16	11
h9n	10	7	17	10
h9z	10	7	14	10
h10m	12	9	13	11
h10n	10	7	15	11
h10z	8	9	14	15
sv1	11	10	18	15
sv2	12	12	22	16

Table 14: average of historical transportation times for different task sorts (in minutes)

The transportation times of the start matrix are shown in tables 22 and 23 in Appendix C. These times are used for the movements without a bed from the end zone of a finished task to the start zone of a next task. Transposing this matrix will give us the exact same matrix, since these times are only based upon the number of floors and the distance that has to be covered through the hallway.

The assumption is made that workers are only 79.7% of their duty busy with tasks, assumption 10, and therefore the choice is made to add extra time to the transportation times and start matrix. By dividing the transportation times and start matrix by 0.797. The time a worker is not busy with tasks is equally divided over a day.

8.3.2 Event based simulation

The I-transport system supports a tool which can run real-time simulations for assigning the tasks to workers. The disadvantage of this is that only one simulation can be done in a day and therefore analyzing different options would be very time consuming. The simulation model built for this thesis can be identified as an event based simulation, which means the simulation time jumps forward to the time a next event happens. This makes the simulation fast and therefore a lot of options can be analyzed for different simulation days. Therefore this simulation is more useful compared to a real time simulation.

There are three different possible events to which the jump in time can be made:

1. A worker finished a task
2. Start/finish of a break (morning, lunch or afternoon)
3. Thorax cleaning

Whenever a worker finished a task, he is ready for a new task and therefore a new task can be assigned to him. The way of assigning these tasks depends upon the chosen way of assigning out of the options explained in the next section. As long as there are free workers and unassigned tasks the systems keeps assigning tasks.

Another event is a break for the workers and these are at 9.30, 12.15 and 15.30. This means that a jump in time to these break times is always made and whenever these times are reached all workers get a break. This is implemented in such a way that a worker first finishes his current tasks and then takes a break for 15 minutes for a coffee break in the morning and afternoon and 30 minutes for a lunch break. When the worker finishes his break, he is immediately ready to start completing tasks again.

The third and last event is the cleaning of Thorax. Two workers have to clean the beds at the Thorax location once in the morning and once in the afternoon. This cleaning is done around 7.30 and 14.00 and it takes two workers half an hour to clean the beds. For example, this means that when the simulation time reaches 14.00 the first two workers that have finished their task will go to Thorax to clean beds. These tasks will take exactly half an hour and when these workers are finished they will start doing other tasks again.

8.4 Options

In this section the different options of the model are explained. We start with a basic model in which grouping of tasks is not possible, which means that a task is always linked to only one bed. This basis model will be used to test different ways of assigning tasks. The best way of assigning the tasks is chosen and the results will be compared with adding the possibility of grouping tasks. The last and final option that is tested is what the influence will be of giving deliver tasks a higher priority to pick up tasks. In testing the ways of assigning and the different options for the system, it is assumed that ten clean beds are ready in the morning to be delivered (b=10) and 7 workers are working on a day (w=7). The effect of different values for these input variables will be analyzed later.

8.4.1 Ways of assigning tasks

There are a lot of different ways to assign the tasks, however the effect of the different ways to assign these tasks is limited. The reason for this is that tasks are only known two hours before their planned arrival time. This means that only few tasks are known at a certain moment in time and tasks cannot be planned in such an efficient way as in the case in which all tasks are known for the whole day.

The first possibility, option 1, is to assign tasks in such a way that it is as efficiently as possible for the workers. This means that when a worker is finished with a task, he gets the best task for him, i.e. the task of which the start location is the closest to his current location. The main disadvantage of this way of assigning is that some tasks might be finished much too late, because workers never get close to the start location of that particular task.

That is the reason for creating option 2 in which priority is added. This means that tasks are still assigned as efficiently as possible for the workers, but late tasks have a priority. The priority starts at zero and when a task is late, the priority will be increased. For each 5 minutes the task is more too late, the priority is increased by 1. Then first the tasks with the highest priority are the only possible tasks to be assigned to workers. After all tasks with the highest priority are assigned, then tasks with less priority will be available to be assigned and so on, until all workers are busy again.

Another way to assign tasks, option 3, is to assign the task with the earliest latest start time to a worker. If more workers are available the worker which is the nearest to the start location of the task will get the task. The focus of this way of assigning tasks is on making as many tasks as possible finish on time, but workers might have to cover more distance and finish fewer tasks in a day.

I-transport is created a long time ago and therefore the makers are not sure in what way tasks are assigned in the system, but most likely a way is used which is similar to option 3.

8.4.2 Grouping tasks

The second option is the possibility of grouping tasks: combining two tasks which have a start location in the same zone and an end location in the same zone. Grouping is only possible for tasks which are known in the system and therefore only tasks which have a planned arrival time which differ less than two hours can be grouped.

The bed cleaning department had the possibility for grouping tasks implemented for some months and these months yielded the data shown in table 14. Currently the system does not allow grouping anymore and therefore we want to analysis what the benefits would be for implementing it again. The data shows that the average factor for which the transportation of a single task is multiplied to get the grouped task is 1.63 and 1.42 respectively for deliver and pick up tasks. Since these factors are smaller than two, a better performance is expected when allowing grouped tasks.

8.4.3 Priority to deliveries

The last option that is tested with the simulation is to give delivering of clean beds a higher priority than picking up dirty beds. The motivation behind this option is that an admission needs a bed on time to be sure that when the patient enters the hospital the bed is ready. For dirty beds however this is less relevant, since the beds can be placed on the hallway to be picked up. This option implies that tasks are less efficiently assigned, since this is an extra restriction on assigning the tasks. With this analysis we want to find out how badly the results are influenced by this option. With this option the number of clean beds will often reach zero at the bed cleaning department, which means that dirty beds have to be picked up.

8.5 Sensitivity analysis

Until now the number for workers (w) and the number of beds available in the morning (b) have been kept constant at 7 and 10 respectively. To see how the results respond to changes in these input variables a sensitivity analysis will be used in which the number of beds and the number of workers will vary. The number of workers will vary from 3 to 10, since less or more workers will not give us any interesting information. Less than 3 workers will never happen at the bed cleaning department and more than 10 workers will most likely not influence the results anymore, because all tasks will probably be on time with such an overcapacity.

For this sensitivity analysis the simulation model will be used where grouping is allowed and deliver task have no priority compared to pick up tasks. Grouping is allowed, because a better performance is expected when grouping is allowed. Deliver tasks have no priority, because this makes it hard to evaluate the results. With this sensitivity analysis we want to find out the positive impact of having more beds ready in the morning. However, when using priority for deliveries, less bed could be delivered on time when having more beds ready in the morning. More clean beds ready in the morning means more deliver tasks will be done before pick up tasks, which could be less efficient.

9. Simulation: results

In this chapter the results of the different ways of assigning, of the different options for the simulation and of the sensitivity analysis will be shown. In the first section we will explain the used data for the simulation and the following sections will show all results.

9.1 Data

For the simulation six different days are chosen from the historical data of I-transport. The reason for taking these specific days is there are a lot of tasks on these days, which means also some tasks were not finished on time or not finished at all. Therefore these days are interesting to test the different options of the model on. The chosen days will not contain any transportations to zones for which we do not have a transportation time yet, since these days contain the tasks used to create the vectors of transportation times shown in table 14. Information over these simulation days is shown in the following table.

Date	Day	# tasks (n_{tasks})	# deliver tasks	# pick-up tasks	# different zones
17-06-2013	Monday	259	130	129	28
01-07-2013	Monday	284	129	155	30
18-07-2013	Thursday	231	119	112	29
19-07-2013	Friday	237	113	124	27
16-08-2013	Friday	256	124	132	32
28-08-2013	Wednesday	272	141	131	26

Table 15: details information over the chosen simulation days from the historical data

For evaluating the different options the output variables of the model will be used. The first variable is the percentage of tasks that are finished: P_{fin}. This is also the first variable which is looked at when comparing the different options, since as many tasks as possible should be finished before comparing the percentage of late tasks: P_{late}. For these percentages only the number of finished tasks is taken as denominator, because tasks which are not finished at all are left out. The bed cleaning department aims at finishing a task within one hour and therefore also the percentages of tasks that are finished within one hour are shown: P_{on}.

The first three output variables are related to a number of tasks that are finished or finished within a certain timeframe. However, next to these percentages we also want to know the actual time tasks are late. Therefore the last two output variables are related to lateness of tasks. The first variable, T, is the sum of the lateness of all tasks together divided by all tasks that are late. This shows the average lateness of a late task. The lateness of tasks cannot be too large, since admissions really need a bed. Therefore we also want to know the maximum time a task is late: M.

9.2 Ways of assigning tasks

The next table shows the results of option 1: assigning task as efficiently as possible for the workers.

Date	P_{fin}	P_{late}	P_{on}	*T (min)*	*M (min)*
17-06-2013	100%	44.79%	21.24%	70.6	442.7
01-07-2013	100%	24.30%	17.96%	155.5	470.4
18-07-2013	100%	3.90%	72.29%	32.7	78.1
19-07-2013	97.05%	8.02%	47.25%	47.9	273.7
16-08-2013	100%	8.20%	40.23%	78.4	316.7
28-08-2013	100%	19.49%	33.09%	83.2	354.3

Table 16: results of simulation for option 1 of assigning tasks

For all dates expect 18-07-2013 all tasks are finished. For 18-07-2013 the percentage of finished tasks, P_{fin}, is still close to a 100%, which means only a few tasks are not finished for this date. Therefore the focus in this evaluation can be on the number of tasks that are late and the lateness of these tasks. To evaluate the percentages of late tasks, P_{late}, they should be compared with the other models, but the 44.79% for 17-06-2013 seems to be high. Also the lateness of the tasks and especially the maximum lateness of a task, M, is high. For the first data a task is more than 7 hours late, which means a task at the beginning of the day is constantly skipped when assigning tasks and is finished at the end of the day when all other tasks are finished. Tasks with such lateness are unacceptable for the hospital and therefore we take a look at the next model in which priority is added to make sure late tasks are assigned sooner. The results of option 2 as way of assigning tasks are shown in table 17.

Date	P_{fin}	P_{late}	P_{on}	*T (min)*	*M (min)*
17-06-2013	100%	74.13%	12.36%	38.5	182.6
01-07-2013	98.59%	73.93%	1.79%	35.6	146.3
18-07-2013	100%	4.33%	71.86%	9.1	20.0
19-07-2013	97.05%	13.50%	44.30%	6.7	20.4
16-08-2013	100%	7.03%	30.47%	6.0	20.6
28-08-2013	100%	38.60%	21.69%	24.7	74.3

Table 17: results of simulation for option 2 of assigning tasks

The percentages of late tasks have increased for almost all days. The largest increase is even from 24.30% to 73.93%, which is a big difference. Although more tasks are late now, there are no longer tasks for which the lateness is more than 4 hours. However, for 17-06-2013 the maximum lateness is still a little bit more than 3 hours, which is still too high and unacceptable. Therefore in the third and last option of assigning the tasks the focus is only on finishing tasks on time. The results are shown in the following table:

Date	P_{fin}	P_{late}	P_{on}	*T (min)*	*M (min)*
17-06-2013	100%	65.25%	3.09%	44.9	95.9
01-07-2013	100%	76.41%	1.76%	36.7	86.9
18-07-2013	100%	0.43%	74.89%	5.6	5.6
19-07-2013	97.05%	3.38%	41.35%	3.4	11.6
16-08-2013	100%	1.56%	18.75%	7.2	9.0
28-08-2013	100%	30.15%	20.59%	21.8	59.4

Table 18: results of simulation for option 3 of assigning tasks

Although the percentage of late tasks for 17-06-2013 and 01-07-2013 are still quite high, the maximum time a task is late has significantly decreased. Day 17-06-2013 still contains a task which is 95.9 minutes late. However, this is something which can be expected, because the busiest days are chosen for the simulation and only 7 workers are scheduled to work in the simulation. The other output variables are quite similar to the previous option, which shows the biggest gain with option 3 is reducing the maximum time a task is late. This way of assigning is therefore chosen as the best option and for analyzing the other options of the simulation this way will be used.

This result shows that in the current implementation of the system the way of assigning tasks should not be changed, because I-transport now has a similar way of assigning tasks as option 3. Although better ways of assigning tasks might be possible, these ways are not found in the tested options.

9.3 Grouping tasks

The results of the simulation when grouping is possible are shown in table 19.

Date	P_{fin}	P_{late}	P_{on}	*T (min)*	*M (min)*
17-06-2013	100%	0.77%	79.92%	3.4	3.4
01-07-2013	100%	0.70%	55.99%	3.6	3.6
18-07-2013	100%	0%	90.45%	0	0
19-07-2013	100%	2.53%	84.8%	4.7	11.6
16-08-2013	100%	1.17%	82.81%	1.1	1.3
28-08-2013	100%	0.37%	64.70%	0.9	0.9

Table 19: results of simulation with grouping tasks allowed

To get an idea of the impact of making grouping possible, the results are compared with the results without grouping, see table 18. The first difference is that in these results all tasks are finished for each day even for 19-07-2013. Next we notice that all percentages of late tasks are significantly lower. For 01-07-2013 it is a decrease of more than 75%. Also the percentages of tasks that are finished within one hour increased a lot and the same holds for the average lateness of a late task (T) and the maximum lateness (M). The differences are huge and therefore grouping should be an option that needs to be implemented in the system again. Grouping of tasks improves the results a lot and therefore grouping is allowed in producing the results of the following sections.

9.4 Priority to delivery

Table 20 shows the results of giving priority to delivery tasks instead of pick up tasks.

Date	P_{fin}	P_{late}	P_{on}	*T (min)*	*M (min)*
17-06-2013	100%	1.54%	80.31%	7	10.6
01-07-2013	100%	11.27%	69.72%	18.3	44.9
18-07-2013	100%	0%	89.61%	0	0
19-07-2013	100%	2.53%	84.81%	11.1	17.0
16-08-2013	100%	1.17%	82.42%	1.1	1.3
28-08-2013	100%	0.37%	68.01%	15	15

Table 20: results of simulation with priority for delivery

For most days only the variables T and M differ a lot with the results without priority to deliveries, see table 19, except for 01-07-2013. For 01-07-2013 the percentages of tasks that is late has increased with $11.27\% - 0.70\% = 9.43\%$ and also the variables T and M increased a lot. However, even for 01-07-2013 the results are not that bad, since more tasks are finished within an hour. For all other days the differences are small and therefore giving priority to deliveries does not have a really big influence on the results. These results show that giving priority to deliveries is a good idea and the negative impact is acceptable.

9.5 Sensitivity analysis

In this section the influence of the input variables are tested. The variables are the number of workers (w) and the number of clean beds at the bed cleaning department at the beginning of the day (b). To make the results for different values of the input variables comparable we make use of the simulation model with grouping of tasks, but without priority for deliveries. If this priority option would be taken, more beds ready in the morning could have a negative impact on the results, since more beds means more deliveries and this could be less efficient. We want to find out the positive effect of having more beds ready at the bed cleaning department and therefore the priority for deliveries is left out.

For the sensitivity analysis the data of 28-08-2013 is used. The number of beds will vary from zero to twelve and the number of workers will vary from 1 to 10. As result the percentage of beds on time is taken. This variable is the most relevant output variable, assuming all tasks are finished, and the one which can be easily compared. The results are shown in table 21.

$b\backslash w$	1	2	3	4	5	6	7	8	9	10
0	2.41%	6.62%	16.75%	21.04%	32.55%	87.02%	94.09%	95.09%	96.18%	96. 18%
1	3.18%	7.69%	17.52%	22.15%	34.12%	89.35%	95.82%	96.20%	96.58%	96.58%
2	7.69%	14.81%	17.90%	22.50%	34.88%	90.91%	96.59%	97.16%	97.73%	97.73%
3	7.69%	14.81%	18.35%	22.83%	35.19%	91.81%	96.98%	98.22%	98.87%	98.87%
4	7.69%	14.81%	18.35%	23.04%	35.19%	92.36%	97.74%	98.42%	98.87%	98.87%
5	7.69%	14.81%	18.35%	23.04%	35.19%	92.88%	98.13%	99.25%	99.25%	99.25%
6	7.69%	14.81%	18.35%	23.04%	35.19%	93.66%	99.25%	99.88%	99.88%	99.88%
7	7.69%	14.81%	18.35%	23.04%	35.19%	93.68%	99.25%	99.88%	99.88%	99.88%
8	7.69%	14.81%	18.35%	23.04%	35.82%	94.81%	99.63%	100%	100%	100%
9	7.69%	14.81%	18.35%	23.04%	35.82%	95.19%	99.63%	100%	100%	100%
10	7.69%	14.81%	18.35%	23.04%	35.82%	96.32%	99.63%	100%	100%	100%
11	7.69%	14.81%	18.35%	23.04%	35.82%	96.32%	99.63%	100%	100%	100%
12	7.69%	14.81%	18.35%	23.04%	35.82%	96.32%	99.63%	100%	100%	100%

Table 21: percentages of on time tasks for different values for the input variables

The first thing that can be noticed is that for more than nine beds, the results do not change anymore. The reason for this is that with more than ten beds the bed cleaning department never runs out of clean beds. The results also do not change any more for more than eight workers. Table 21 also shows this implies that for some values of b there will always be late tasks no matter how many workers are working. This is because at a certain moment in the simulation no more clean beds are available at the bed cleaning department and therefore some delivery tasks will always be late. These late tasks could be for example at the beginning of the day which means no matter how many workers are present the dirty beds cannot be picked up and cleaned on time to make the deliveries finish on time.

A big gap is present in the results between five and six workers scheduled on this day. The reason for this gap is that with less than six workers, finishing tasks will always lag behind. At the beginning of the day two workers are needed at the Thorax for cleaning beds and this means that the day is started with two workers less than scheduled to do all tasks. This implies that during the cleaning of beds at the Thorax location the bed cleaning department will have undercapacity and therefore tasks in the queue will pile up. During the whole working day this undercapacity at the beginning of the day is never overcome. With at least six workers the busy part of the beginning of the day can be handled and this explains the big difference in scheduling five or six workers. If the assumption is made that most days will have a similar pattern as the example day 28-08-2013, the bed cleaning department should make sure that they have at least six workers working on every day. If it is beneficial to schedule more than six workers depends upon the costs compared with the improvement in performance. This also depends upon the number of clean beds available in the morning, because with only one bed ready to be delivered ($b = 1$) the benefit of more than 6 workers ($w \geq 6$), is larger compared to having ten clean beds ready in the morning ($b = 10$).

10. Conclusion

This paper started with making forecasts for admissions and discharges of patients. The best model we found fitted the historical data very well and can be used to make forecasts for the upcoming years. These forecasts can be useful in making decisions for the hospital on a strategic, tactical and operational level. On an operational level this information can be used for the logistical process of transporting clean and dirty beds. For this process the actual forecasts are not relevant, but only the part of the forecasted admissions and discharges which need a bed. However, the data about how many admissions need a bed is missing and therefore different scenarios are created. As soon as this data comes available the right scenario can be chosen and the crew schedule for the bed cleaning department can be based upon the results for that particular scenario. Without this missing data a good link between the forecasts and the number of tasks they generate for the bed cleaning department cannot be made and therefore only the historical data is used in analyzing the system of I-transport.

In analyzing this system different ways of assigning tasks are tested and different options for the system are tested. To do these analyzes we had to build a simulation which simulates days as close as possible to the way I-transport is used. With the results of the simulation show that in the way of assigning tasks the focus should be on making tasks finish on time instead of making it as efficiently as possible for the workers. This is also how the current implementation of I-transport is, so this should not be changed. Grouping tasks was allowed for some months in the implementation of I-transport, but this option is currently set off again. The results show that this option should definitely be set on again, because a lot more tasks can be finished on time when workers are allowed to handle two beds in one task. For every admission of a patient a new bed should be ready, while if really necessary a dirty bed can be put on the hallway until it is picked up. Therefore also the effect of giving priority to deliver tasks is analyzed and the results show that the negative impact on results is only small. This could be a reason to give all deliver tasks a higher priority. The performed sensitivity analysis shows that at least six workers should be scheduled on a day, because less than six workers will lower the performance drastically.

11. Discussion

Better and more data is needed to improve this analysis. Although a lot of data is available at the hospital, the main missing part for this analysis is data about the number of admissions that need a clean bed. The easiest way to collect this data is by letting every specialism count the number of admissions that need a bed. This data can then be used to split the forecasts of admissions over the different specialisms, because the percentages of admissions that need a bed will most likely vary over the different specialisms. Then the translation from forecasts to tasks can be made more precisely and the results will be more reliable.

Also the data of transportation times should be analyzed and corrected. As explained in section 8.3.1 there are transportation times which most likely are incorrect and influenced by outliers. The most obvious reason for this problem is that only a few observations are in the dataset for these zones. Therefore this dataset of transportation times should be extracted from the system and thoroughly analyzed. After doing this further research of data, some transportation times could be corrected or left out because they are outliers.

In the simulation model the average transportation times are used, because the data from which these averages are calculated is not available. Therefore the simulation model cannot be made stochastic. To make it a stochastic model, distributions need to be fit to this missing data. The impact on the results when using stochastic instead of deterministic transportation times will not be too large, because the variation in a transportation time of one zone to another zone will probably not be too high. For example, if on average the transportation time from zone A to zone B is 10 minutes, the minimum and maximum time will most likely not be smaller than 7 and 13 minutes respectively. However, to make sure this impact is not too large, this missing data should be obtained and analyzed.

Another part of future research could be to analyze the effect of combining the crews of patient and bed transport. These two crews are currently separated, but both work with the same system: I-transport. By combining the teams the two crews can help each other at the peak moments. This implies that both crews need workers which can transport beds and patients which is currently not the case. Most workers of the bed cleaning department are not skilled to transport patients and the other way around. If the analysis shows that combining these teams could mean a big cost reduction, this could be a reason to train both crews to make them more all-round workers.

12. References

Adan, I.J.B.F. & Vissers, J.M.H. (2002). *Patient mix optimisation in hospital admission planning: a case study*. International Journal of Operations & Production Management Vol. **22**(4): 445- 461

Benjamini, Y. (1988). *Opening the Box of a Boxplot*. The American Statistician Vol. **42**(4): 257-262

Brailsford, S.C., Harper, P.R., Patel, B. & Pitt, M. (2009). *An analysis of the academic literature on simulation and modelling in health care*. Journal of Simulation **3**: 130-140

Carter, M. (2002). *Diagnosis: Mismanagement of resources*. OR/MS Today **29**(2): 26-32

Centraal Plan Bureau (CPB). (2013). *Toekomst voor de zorg*. CPB Boek

Dawson, R. (2011). *How Significant Is a Boxplot Outlier?*. Journal of Statistics Education Vol. **19**(2): 19

Dumas, B. (1985). *Hospital Bed Utilization: An Implemented Simulation Approach to Adjusting and Maintaining Appropriate Levels.* Health Services Research Vol. **20**(1): 43-61

Garcia, K. A. (2009). *Using a Randomized Regression Approach to Estimate Hospital Admissions to Reduce Emergency Department Holding*. Master's thesis Computer Science, Florida Institute of Technology

Goldman, J., Knappenberger, H.A. & Eller, J.C. (1968). *Evaluating Bed Allocation Policy with Computer Simulation*. Health Services Research Vol. **3**(2): 119-129

Green, L. V. (2002). *How many hospital beds?*. Inquiry - Excellus Health Plan Vol. **39**(4): 400-412

Heij, C., de Boer, P., Franses, P.H., Kloek, T. & van Dijk, H.K. (2004). *Econometric Methods with Applications in Business and Economics*. Oxford University Press.

Kuhn, H.W. (2012) *A tale of three eras: The discovery and rediscovery of the Hungarian Method*. European Journal of Operational Research **219**: 641-651

Kil, S. (2011). *Bed Occupancy Optimization: Combining Wards to Gain Performance*. Master's Thesis Operations Research and Quantitative Logistics, Erasmus School of Economics

Machielsen, P.M. (2011). *De keten-DBC's gediagnosticeerd. Een explorerend onderzoek naar de invoering van keten-DBC's binnen algemene ziekenhuizen*. Afstudeerscriptie Zorgmanagement, Erasmus University Rotterdam

Malhotra, S & Zodpey S. (2010). *Operations research in public health*. Public Health Research Methods **54**: 145-150

Oostrum, J. van. (2009). *Applying mathematical models to surgical patient planning*. Erasmus Research Institute of Management

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley series in behavioral science : quantitative methods

Wilson, C.R.(1997). *Distortion effects and extreme observations in empirical research; An analysis of the incremental information content of cash flows*. Accounting and Finance **37**: 163-180

Appendix

A. DOT system

Knowing how the Dutch Healthcare system works, will help to understand the context of this research and therefore DBC/DOT system is briefly explained here. Until 2004 was every action on a patient in a hospital invoiced separately and directly after the action was taken. However, this was decided to be inefficient and therefore the Dutch government has decided in 2005 to make the medical sector more like a competitive market. Therefore they introduced a new financial system which consists out of DBC’s (diagnosebehandelingcombinaties). A DBC describes the diagnosis that has been made and the treatment that is needed. A DBC consists out of all actions and treatments that a patient has in the hospital for a certain period. Hospitals can only state a price for the patient for the DBC’s and not for any separate actions. The DBC’s consists out of two segments: segment A in which the price of a DBC is fixed for all hospitals and this price is calculated by an external organization. And segment B where the price of the DBC’s can be negotiated between the health insurance company and the hospitals. However, in 2012 a new even more transparent system was developed which is called DOT (diagnosebehandelingcombinatie onderweg naar transparantie) by decreasing the number of DBC combinations from 30,000 to 4,400 healthcare products. By making it more transparent the hospitals are more comparable and therefore they will compete more with each other.

B. Extra figures of data analysis

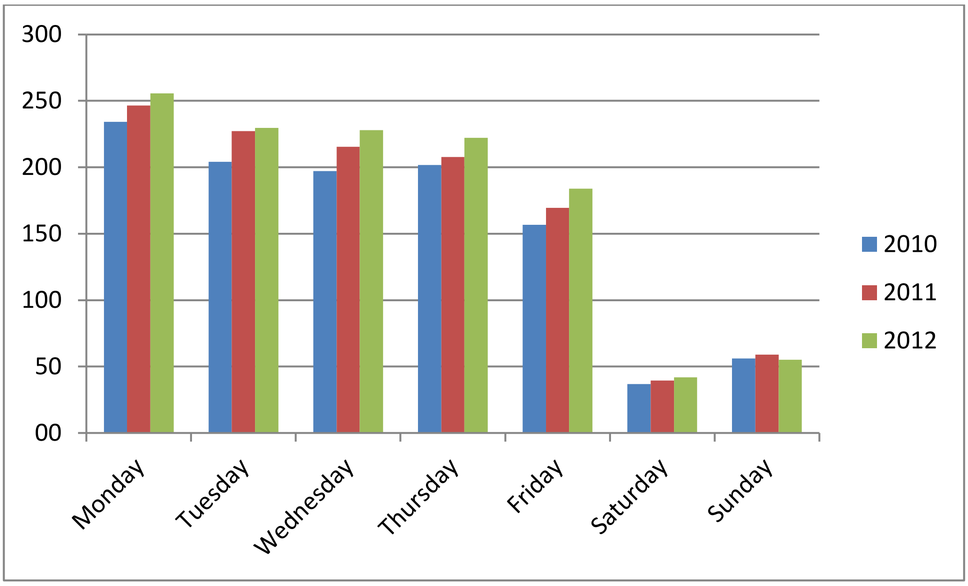

Figure 6: average number of admissions for each day

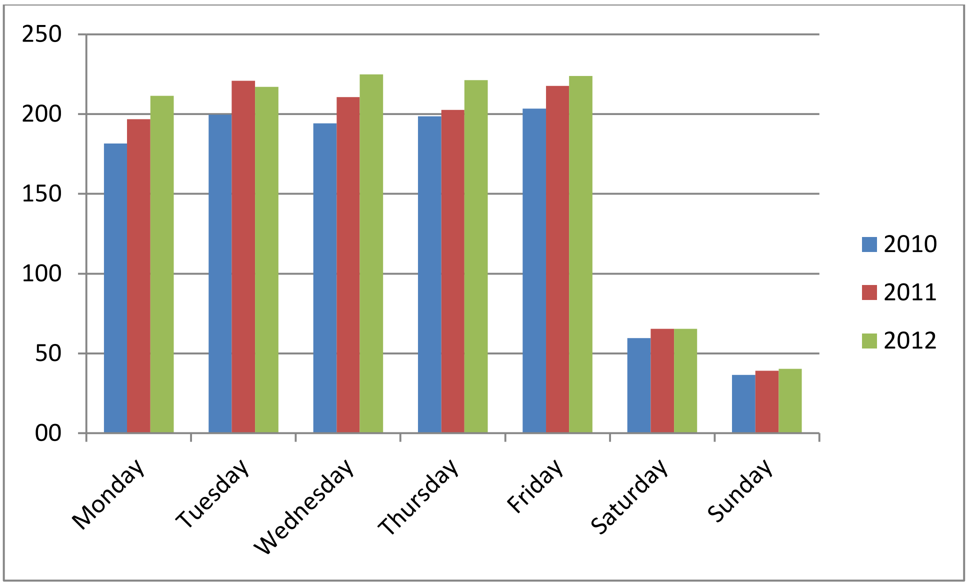

Figure 7: average number of discharges for each day

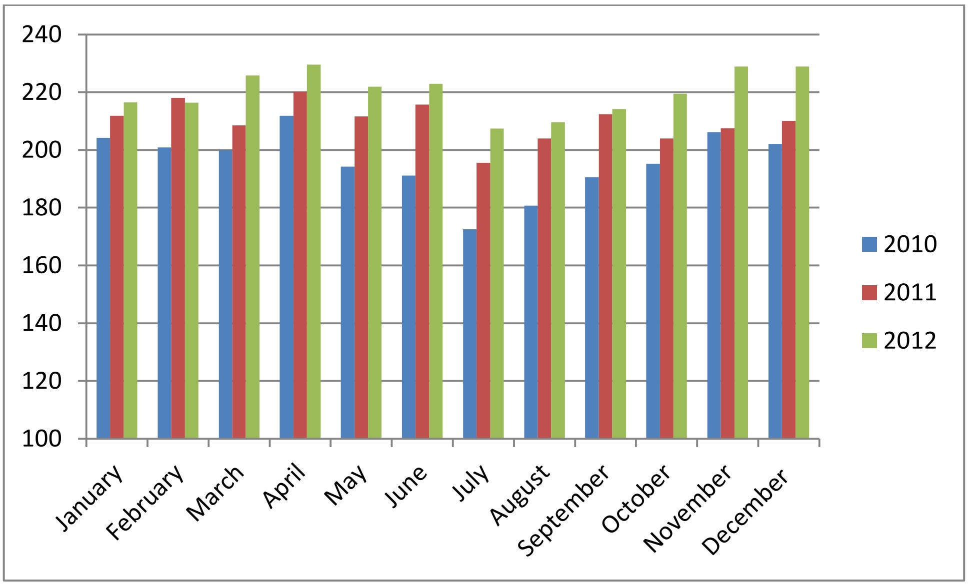

Figure 8: average number of discharges on a weekday for each month

C. Start matrix

Zones	ba0	bd1	bd2	bd3	bd4	ca2	d3	dp1	dp2	dp3	ga0	h1z	h2m	h2n	h3m	h3n	h3z	h4m	h4n
ba0	0	1	2	3	3	2	2	3	3	3	8	2	2	2	3	3	3	3	3
bd1	1	0	1	2	2	4	2	3	3	3	8	2	3	3	3	3	3	3	3
bd2	2	1	0	1	1	4	2	3	3	3	8	3	3	3	3	3	3	3	3
bd3	3	2	1	0	1	4	2	3	3	3	8	3	3	3	3	3	3	3	3
bd4	3	2	1	1	0	4	2	3	3	3	8	3	4	3	3	3	3	3	3
ca2	2	4	4	4	4	0	2	3	3	3	2	3	3	3	3	3	3	3	3
d3	2	2	2	2	2	2	0	3	3	3	2	2	2	2	2	2	2	2	2
dp1	3	3	3	3	3	3	3	0	1	1	3	3	3	3	3	3	3	3	3
dp2	3	3	3	3	3	3	3	1	0	1	3	3	3	3	3	3	3	3	3
dp3	3	3	3	3	3	3	3	1	1	0	4	3	3	3	3	3	3	3	3
ga0	8	8	8	8	8	2	2	3	3	4	0	2	2	2	2	2	2	2	2
h1z	2	2	3	3	3	3	2	3	3	3	2	0	1	1	1	1	1	2	2
h2m	2	3	3	3	4	3	2	3	3	3	2	1	0	1	1	1	1	2	2
h2n	2	3	3	3	3	3	2	3	3	3	2	1	1	0	1	1	1	2	2
h3m	3	3	3	3	3	3	2	3	3	3	2	1	1	1	0	1	1	1	1
h3n	3	3	3	3	3	3	2	3	3	3	2	1	1	1	1	0	1	1	1
h3z	3	3	3	3	3	3	2	3	3	3	2	1	1	1	1	1	0	1	1
h4m	3	3	3	3	3	3	2	3	3	3	2	2	2	2	1	1	1	0	1
h4n	3	3	3	3	3	3	2	3	3	3	2	2	2	2	1	1	1	1	0
h4z	3	3	3	3	3	3	2	3	3	3	2	2	2	2	1	1	1	1	1
h5m	3	3	3	3	3	3	2	3	3	3	2	2	2	2	1	1	1	1	1
h5n	3	3	3	3	3	3	2	3	3	3	2	2	2	2	1	1	1	1	1
h5z	3	3	3	3	3	3	2	3	2	2	2	2	2	2	1	1	1	1	1
h6m	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	1	1
h6n	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	1	1
h7m	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h7n	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h7z	3	3	3	3	3	3	2	5	2	2	2	2	2	2	2	2	2	2	2
h8m	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h8n	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h9m	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h9n	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h9z	3	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h10m	2	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h10n	2	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
h10z	2	3	3	3	3	3	2	3	2	2	2	2	2	2	2	2	2	2	2
sv1	3	5	5	5	5	3	3	5	2	2	3	2	2	2	2	2	2	3	3
sv2	3	5	5	5	5	3	3	3	2	2	3	2	2	2	2	2	2	2	2

Table 22: start matrix which contains transportation times for movements without a bed (part I)

Zones	h4z	h5m	h5n	h5z	h6m	h6n	h7m	h7n	h7z	h8m	h8n	h9m	h9n	h9z	h10m	h10n	h10z	sv1	sv2
ba0	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	3
bd1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5	5
bd2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5	5
bd3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5	5
bd4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5	5
ca2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
d3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3
dp1	3	3	3	3	3	3	3	3	5	3	3	3	3	3	3	3	3	5	3
dp2	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
dp3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ga0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3
h1z	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h2m	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h2n	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h3m	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h3n	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h3z	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
h4m	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h4n	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h4z	0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h5m	1	0	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h5n	1	1	0	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h5z	1	1	1	0	1	1	2	2	2	2	2	2	2	2	2	2	2	3	2
h6m	1	1	1	1	0	1	1	1	1	2	2	2	2	2	2	2	2	3	2
h6n	1	1	1	1	1	0	1	1	1	2	2	2	2	2	2	2	2	3	2
h7m	2	2	2	2	1	1	0	1	1	1	1	2	2	2	2	2	2	3	2
h7n	2	2	2	2	1	1	1	0	1	1	1	2	2	2	2	2	2	3	2
h7z	2	2	2	2	1	1	1	1	0	1	1	2	2	2	2	2	2	3	2
h8m	2	2	2	2	2	2	1	1	1	0	1	1	1	1	2	2	2	3	2
h8n	2	2	2	2	2	2	1	1	1	1	0	1	1	1	2	2	2	3	2
h9m	2	2	2	2	2	2	2	2	2	1	1	0	1	1	2	2	2	3	2
h9n	2	2	2	2	2	2	2	2	2	1	1	1	0	1	2	2	2	3	2
h9z	2	2	2	2	2	2	2	2	2	1	1	1	1	0	1	1	1	3	2
h10m	2	2	2	2	2	2	2	2	2	2	2	2	2	1	0	1	1	3	2
h10n	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	0	1	3	2
h10z	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	0	1	2
sv1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	0	1
sv2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	0

Table 23: start matrix which contains transportation times for movements without a bed (part II)