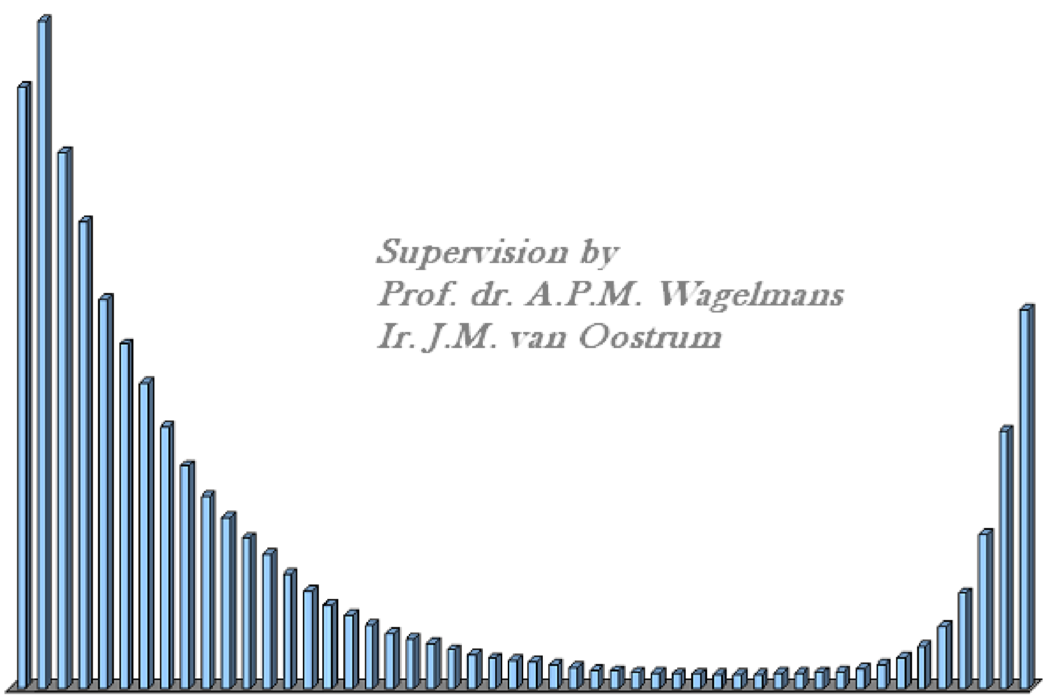
FEM 21031
Master's Thesis

Effects of the length of the planning horizon in a Master Surgical Schedule

Lanah Evers
264834



Supervision by
Prof. dr. A.P.M. Wagelmans
Ir. J.M. van Oostrum

Index

1. Introduction 4
2. Literature review 5
3. Problem description 8
4. Data 11
4.1 Patient data 11
4.2 Expected surgery times 11
4.3 MSS 11
4.4 Dummy capacity 12
5. Simulation 14
5.1 Simulating and scheduling patients 14
5.2 Input decisions 15
5.2.1 Arrival process 15
5.2.2 Lower bound on dummy space 16
5.2.3 Distinction between dummies 17
5.2.4 Shift 'late' A surgeries from dummy to A space 17
5.3 The basic scenario 19
5.4 Performance measures 19
5.4.1 Waiting time 19
5.4.2 Length of waiting list 20
5.4.3 Percentage A scheduled in A 20
5.4.4 Planned overtime 20
5.4.5 OR utilisation 21
5.4.6 Demand for hospital beds 21
5.5 Warm-up period 21
5.6 Run length and number of simulations 23
6. Results 25
6.1 Effects of the planning horizon 26
6.1.1 Waiting time 26
6.1.2 Length of waiting list 28
6.1.3 Percentage A scheduled in A 28
6.1.4 Planned overtime 29
6.1.5 OR utilisation 29
6.1.6 Demand for hospital beds 30
6.2 Effects of other input decisions 31
6.2.1 Arrival process 32
6.2.2 Lower bound on dummy space 32
6.2.3 Distinction between dummies 32
6.2.4 Shift 'late' A surgeries from dummy to A space 32
6.3 Time series of bed occupation levels 35
7. Discussion 38
8. Conclusion 39

A. References 40
B. Appendix 40
B.1 Gamma and Poisson distributions 40
B.2 Time series in relation to warm-up period 41
B.3 95% confidence intervals of the performance measures 42
B.4 Distribution of waiting time for all surgery types 43
B.5 Distribution of waiting time per surgery type 49
B.6 Distribution of bed occupation within cycle 50
B.7 QQ-plots: normality of demand levels for hospital beds 51
B.8 Time series identification of bed occupation levels 52

1. Introduction

The administrative department of a hospital has to face several rather complex tasks regarding the construction of time schedules, one of which is scheduling surgeries of a varying flow of patients. Currently hospitals construct operating room schedules over and over again, given the list of elective patients. Elective patients are patients that require surgical procedures which are not emergent and therefore can be scheduled some period ahead in time. The current way of scheduling is a time consuming and complex task. Furthermore, it results in unstable operating room schedules and in variability in occupation of hospital beds in wards and intensive care units.

The construction of a cyclical operating room schedule is one way to make improvements with regard to these issues. Such a cyclical operation room schedule is called a Master Surgical Schedule (MSS). The MSS recurrently assigns time slots for types of frequently performed elective surgical procedures to the available OR-days within the cycle. The term OR-day is used to denote the capacity of one Operating Room per day. The implementation of an MSS is beneficial to the hospital since it ensures most information about the future surgery schedule to be known on forehand.

Applying the MSS, a patient will be scheduled, if possible, within one of the available time slots for the appropriate surgery type within the planning horizon. A planning horizon contains a certain number of cycles of the MSS. In case all available time slots of a certain surgical procedure are occupied within the planning horizon, the patient can be scheduled within so-called dummy space of the MSS. Using a dummy time slot for such a patient means that the patient can still be scheduled within the planning horizon, as opposed to assigning the patient to a waiting list, in which case the patient will not be scheduled before the end of the planning horizon. Thus, the use of available dummy space for a patient results in a reduction of the waiting time for those patients. Extending the length of the planning horizon on the other hand, provides us with more available time slots which are explicitly reserved for the appropriate surgery type. Scheduling the patient in one of those slots provides us with a more stable schedule, but increases the patients waiting time.

Excessive lengths of waiting lists are a great issue in the Dutch healthcare system. On the one hand, short waiting lists seem to be desirable, while on the other hand longer waiting lists provide the schedule maker with a higher probability that all available time slots can be filled with their appropriate surgery type, which means providing a more stable schedule. Factors like stability in the schedule and in the amount of required hospital beds can help healthcare costs decrease significantly.

This research consists of the following parts:

• First of all, the influence of the length of the planning horizon on several factors of interest for both patients and hospital will be measured. These factors include measures like waiting time, stability of the schedule and demand for hospital beds.
• Secondly, the effects of other input factors in the planning process are investigated. By other effects we mean for example including the possibility of rescheduling.
• Also, we are interested in the relations in the demand pattern for hospital beds between consecutive days.

2. Literature Review

The interest in applications of operations research to healthcare scheduling problems has increased strongly over the last decade. Models suggested in literature to provide stable cyclical operating room schedules (MSS), are even more recent. They can be found in [1] to [5], although with slightly different definitions of the MSS. The definition used in this paper is the one formulated in [5] by Van Oostrum et al.: ‘An MSS specifies for each OR-day (i.e. operation room on a day) of the planning cycle a list of recurring surgical procedure types that must be performed.’

While several approaches to improve healthcare planning are proposed in literature, the actual implementation of such tools in practice however, is rather scarce. This is due to managerial implications such as substantial influence of for instance surgeons and anaesthesiologists. Van Oostrum et al. [6] argue that the MSS, contrary to other planning approaches, is very suitable to be actually implemented in hospital organisations.

From an operations research point of view an MSS in operating room scheduling can be more or less compared to an MPS in production planning. The MPS (Master Production Schedule) assigns production moments and production amounts to the available moments in time in order to meet time varying demand. Resulting from the MPS, an MRP (Material Requirements Planning) can be constructed. The MRP gives a detailed description about the need for components and raw materials necessary to produce the final products, present in the MPS. The similarities between the MPS to the MSS with respect to these features are the following:

- The MPS is used to decide when and how much to produce in order to meet time varying demand. The MSS is used to decide when surgeries are scheduled, based on a varying flow of patients. Thus, the ‘demand’ in the MSS is the list of surgeries that need to be performed.
- In the MPS production amounts are assigned to one of all possible time slots on one of all available machines. In the MSS surgeries are assigned to one of all possible time slots within one of all available operation rooms.
- The MRP is a time schedule regarding the material inputs, necessary for the MPS to be carried out. The ‘inputs’ necessary for the surgeries of the MSS to be performed, are the surgeons and assisting hospital personnel. Like the MRP is a result of the MPS, the hospital personnel schedule is a result of the MSS. Also, the demand level for hospital beds is a direct result of scheduling patients based on the MSS.

The main difference between an MPS and an MSS can be found looking at the way time varying demand is handled. The MPS is constructed based on the information about demand sizes of the coming period. Variation in future demands can result in variation in the production amounts and production moments prescribed by the MPS. Frequent changes in the MPS can induce major changes in the detailed MRP schedules, a phenomenon referred to as nervousness. Using the MPS, production amounts of several different future time periods can be produced all at once in advance. The products needed in order to meet future demand maintain in inventory, until they reach their due date. For operating room scheduling the situation is different. The MSS is a fixed time schedule that contains time slots for surgery types that may or not may be occupied by the appropriate surgery types because of the varying flow of patients. When no patients are present, no surgeries can be performed. In that case, time slots of the MSS will remain empty. In other words, no extra work can be performed in periods of low demand in order to compensate future periods of high demand when no patients are 'available'. Thus, an unstable MPS results in cost increase in the MRP process for machine scheduling, while an unstable output of the MSS results in instability of demand levels for hospital beds, operating room schedules and time schedules of hospital personnel. In order to reduce schedule instability in operating room scheduling, it might be helpful to look at some ways of improving schedule stability in production scheduling mentioned in literature.

In production scheduling, one way to resolve nervousness is freezing a part of the planning horizon. Once the production times and production amounts within the frozen horizon are decided upon, they will not be changed, even when information about future demands beyond the planning horizon becomes available. In this way, using a frozen interval gives the manufacturer some certainty about the production amounts necessary in the near future. In [7] Shirdharan et al. use simulation in order to investigate the effect of the freezing method, the effect of the length of the frozen interval and the length of the planning horizon. The effects are measured by means of production and inventory cost and deterioration in customer service. The authors show that freezing up to 50 percent of the planning horizon has marginal effect on the measures just mentioned. In a next article [8], the same authors use the same input decisions, but now in order to measure the effect on schedule stability. In operating room scheduling, freezing a part of the planning horizon can be applied to improve the knowledge of the hospital management about the operating room and personnel schedules in the coming period. On the other hand, the frozen interval results in an increase in the patients' average waiting time to be scheduled for surgery, since time slots of the frozen horizon are not available for scheduling new arrivals.

In [9] Campbell and Mabert investigated cyclical scheduling in order to improve schedule stability in the MPS. Cyclical scheduling indicates that the time between production periods for each item is constant, contrary to varying intervals between production moments as a result of cost minimization. In fact, this same feature is already incorporated in the MSS: for each surgery type, the possible surgery moments are fixed and are repeated every cycle. Campbell and Mabert investigated the additional cost of cyclical scheduling. They developed a mathematical programming model which showed that cyclical schedules result on average in only 4.4 percent higher cost than non-cyclical ones. This is a bit hard to relate to operating room scheduling since we are mainly interested in measures like patients' waiting time and schedule stability. We don't encounter a trade-off between set-up costs and inventory cost like in production scheduling. Applying a cyclical schedule in production planning can also result in a decrease in customer satisfaction since it increases the average time the customer has to wait before receiving the requested products. The same applies to the MSS: the cyclic nature of the MSS results in an increase in the patients' average waiting time, since surgeries cannot be scheduled simply at the first available OR-day.

The articles just mentioned about improving schedule stability considered a fixed dynamic future demand. Uncertainty about the demand levels was neglected. In reality however, demand forecasts of feature periods might include forecast errors. In [10] Neng-Pai and Krajewski investigated the effects of demand uncertainty on several factors, including the stability of the MPS. Uncertainty in future demand results in additional costs. When customer satisfaction plays an important role, a reasonable amount of safety stock is required in order to compensate demands which exceed the expected demand level. The use of safety stock in the MPS can be related to the use of dummies in the MSS. The amount of patients requiring a certain surgery type might exceed the expected amount of patients of that surgery type. Scheduling such a patient within dummy space can be compared to the use of safety stock in production planning. Without extending the patient's waiting time, the patient can still be scheduled for surgery, comparable to the customer that can still receive his products without experiencing additional waiting time. So demand uncertainty is already incorporated in the MSS by the use of dummy space. Neng-Pai and Krajewski showed that one way of improving schedule stability is increasing the amount of safety stock, which can be compared to adding dummy space in the MSS.

An additional way of increasing schedule stability is considering rescheduling. Different ways of rescheduling in production planning are mentioned in [11] to [13]. For the MPS rescheduling means adding or deleting orders. It involves minimizing costs resulting from schedule changes related to lower-level items. The higher the number of levels the MRP contains, the more complicated the rescheduling optimization will get. This is a problem we will not encounter when rescheduling in the MSS, since increasing stability in operating room schedules merely comes down to trying to fill all time slots with the appropriate surgery type. As the stability in the operating room schedule increases, the stability in the resulting personnel schedules and hospital bed occupation levels will increase as well.

An analytical approach of deriving performance measures of appointment driven systems, like hospitals and doctors offices, is described in [14] and [15]. Both articles use a vacation model in order to derive the performance measures of the system. In the first article two queuing systems are combined within one analytical model. First, the queuing system that starts when making the appointment up to the arrival at the service facility is considered. Second, the customer will arrive into a next queue at the service facility itself. Article [15] considers appointment driven systems without that second queue. Another difference between these two articles is the kind of arrival distribution that is used in the model.

3. Problem description

Let OR-day (*i,j*) denote operating room number *j* on the *i*th day within the cycle. We define the MSS as a cyclical operating room schedule which specifies for each OR-day (*i,j*) of the planning cycle a list of recurring surgical procedure types that must be performed. This schedule contains only types of surgery which occur on average at least once per cycle (category A surgeries). Slack is located within the MSS for the surgery types which occur on average less than once per cycle time (category B) and for emergency operations (category C). The slack for category B surgeries is represented in the MSS by so called dummies. Resulting from these definitions, the choice of the length of the cycle determines the number of category A surgeries and thus the amount of planned slack for category B surgeries.

Consider an MSS which contains surgery types of different specialties. Suppose all possible surgeries within each specialty are divided into a number of different surgery types. We will introduce the following notations, which will be used throughout this report.

- I — The set of surgery types
- S — The set of specialties
- a_{is} — Surgery type $i \in I$ of specialty $s \in S$, belonging to category A
- b_{is} — Surgery type $i \in I$ of specialty $s \in S$, belonging to category B
- λ_{is} — The expected arrival rate of surgery type i of specialty s
- n_{is} — The number of time slots of type a_{is} included in the MSS

Figure 1 graphically shows the process of using available time slots when scheduling patients for surgery by means of the MSS.

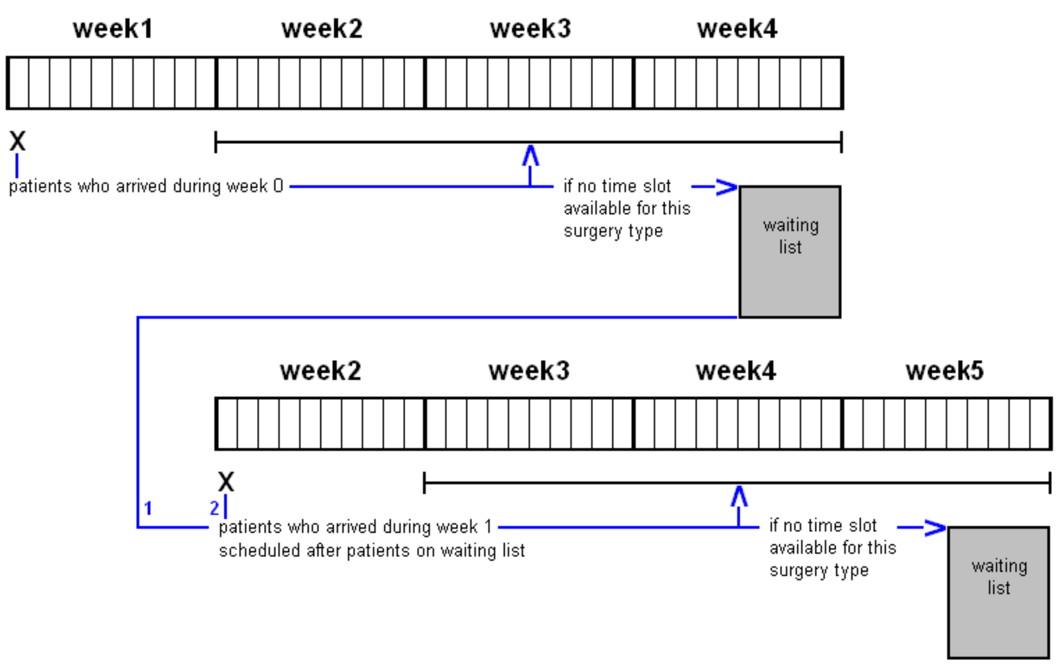


Figure 1: available OR days in the rolling horizon planning process, planning horizon 4 weeks

The cycle length of the MSS which will be considered in this research is one week. This means that after each week the schedule containing available time slots for category A, B and C surgeries repeats itself. The length of future period that is considered to schedule the current list of patients, is called the planning horizon. The planning horizon contains an integer number of cycles which is a variable factor in this research. Patients that don't fit within the planning horizon end up on the waiting list. As mentioned before, elective surgeries (category A and B) are scheduled some time ahead. In this research, the minimum period between the arrival of a patient and the planned surgery date is one week, which corresponds to one cycle length of the MSS. Thus, we will make use of a frozen horizon with the length of one week. All patients arriving during a certain week are accumulated. Next, they will all be scheduled, if possible, within the planning horizon, but without using the time slots of the upcoming week. Since patients can't be scheduled within the first cycle, the planning horizon should contain at least two cycles in order to be able to assign time slots of the MSS to these patients.

The length of the planning horizon in figure 1 is four weeks. The patients that have arrived during the previous week are accumulated. So at the start of week 1 we have a list of patients that have arrived during week 0. These patients can be scheduled starting from week 2, up to and until the end of week 4. Each patient of category A will be scheduled within the first available time slot of its own surgery type a_{is}. If the total of all n_{is} time slots for type a_{is} are occupied within this planning horizon, the patient will be scheduled in the first available dummy space. A patient of category B on the other hand, will be scheduled, if possible, in the dummy space mentioned right away. A distinction is made between dummies assigned to different specialties.

In case a patient cannot be scheduled within the available dummy space, again all dummies will be checked, but now by adding some planned overtime. This feature is included, because it reflects reality. Consider for example a surgery which has an expected duration of 80 minutes. If an OR-day has an available dummy space of only 70 minutes, in reality this OR-day is likely to be used for this surgery nevertheless. In doing so, one ends up with a planned overtime of ten minutes. We have decided upon a maximum use of planned overtime of 50% of the capacity that is assigned the category C surgeries within the OR-day.

If a patient cannot be scheduled within the planning horizon, the patient will be assigned to a waiting list. Next, when simulating the flow of patients at week 2, the end of the planning horizon will move to the end of week 5. Before trying to schedule the patients just simulated, we first check for newly available time slots to schedule the patients from the waiting list in. Continuing in this matter, at each following week the schedule is extended by one week again. This extending feature, while maintaining the surgeries scheduled so far, is called a rolling horizon. This process is illustrated in figure 2. The ranks 1 and 2 denote the order in which the patients are scheduled. Patients on the waiting list are scheduled before the new arrivals.

Figure 2 shows a decision chart, containing all scheduling decisions just mentioned. The list of simulated patients contains both category A and B patients. It contains no category C patients, since emergency patients cannot be scheduled on forehand. Starting from the top of the list, the patients are scheduled in descending order.

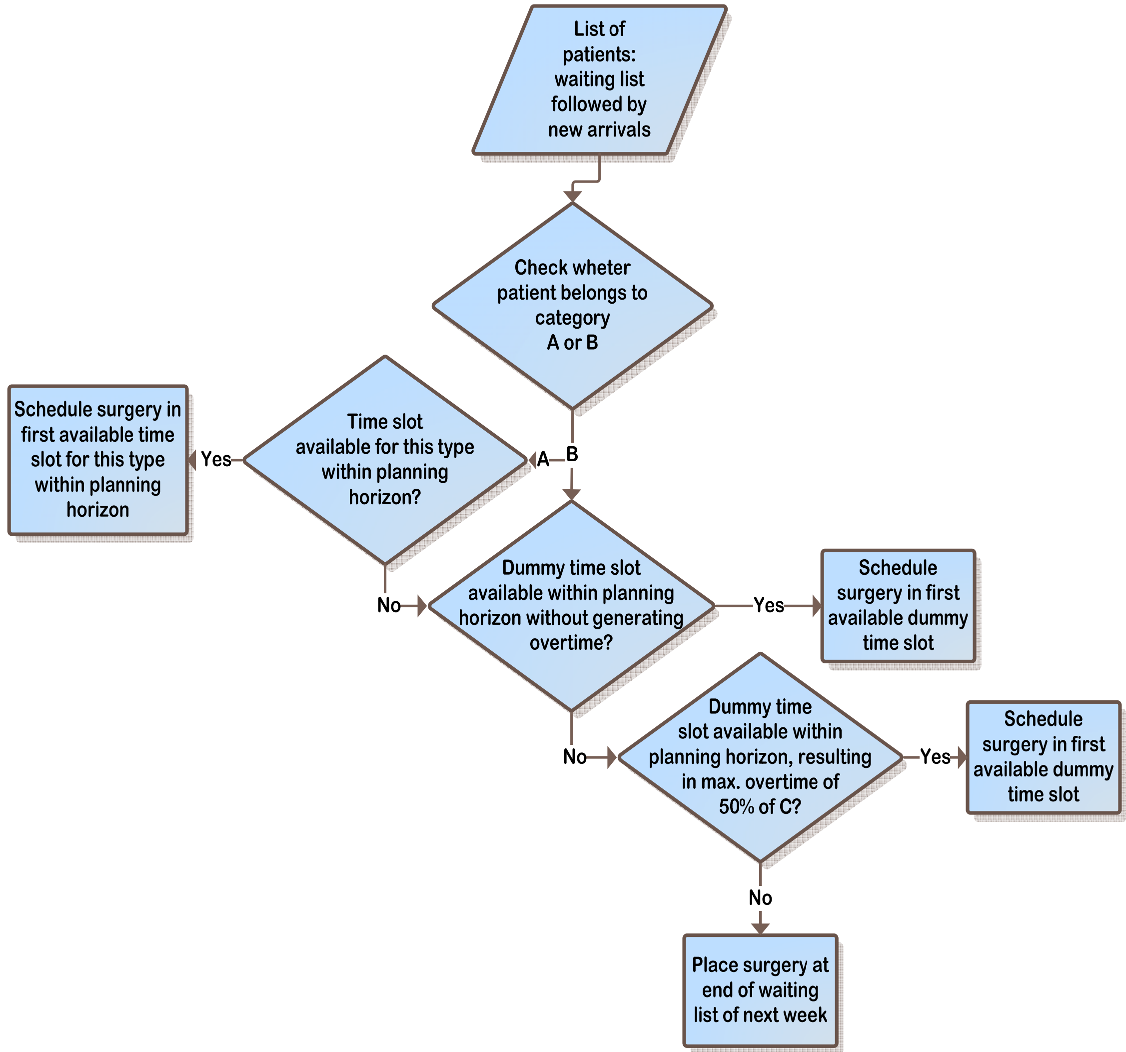


Figure 2: decision chart for scheduling patients

The main object of this research is investigating the effect of the length of the planning horizon on several performance measures. The following measures will be considered:

• The patient's waiting time (average value and distribution)
• The length of the waiting list
• The percentage of category A surgeries scheduled in A time slots
• Planned overtime (the amount of times it occurs and its average duration)
• OR utilisation, which indicates the percentage of used capacity
• Demand for hospital beds (average value and distribution)

The exact definitions of these performance measures will be discussed in section 5.4.

4. Data

The data that is used in this research originates from the urology and general surgery department of the Dutch Beatrix Hospital. First, the main features of the patient data are described in section 4.1. Next, a division of the surgery types of both departments is shown in section 4.2, each with their own expected surgery duration. Based on this specific data, an MSS has been designed which will be presented in section 4.3. Finally, the available dummy capacity resulting from the MSS and the capacity of both departments will be given in section 4.4.

4.1 Patient data

Since we consider data originating from two different departments, we have $s \in \{1,2\}$ in this research, were $s = 1$ denotes general surgery and $s = 2$ denotes urology. Both departments are categorized into thirteen different surgery types, $i \in \{1,2,\ldots,13\}$. Data about the surgeries that took place in both departments during one year is available. For each patient, information is given such as: specialty (urology or general surgery), surgery type, surgery duration and the time spent in the ward. The data file contains information about a total of 1862 patients.

4.2 Expected surgery times

The following table shows the expected surgery durations and yearly frequencies for each surgery type of both the urology and general surgery department. The values are based on the one-year-patient data.

	Duration (minutes)		Frequency (per year)	
Surgery type	**General surgery**	**Urology**	**General surgery**	**Urology**
1	196	54	59	137
2	41	108	418	3
3	217	37	10	93
4	67	65	3	94
5	223	186	10	3
6	60	121	463	6
7	285	298	8	6
8	146	124	58	2
9	85	53	156	60
10	65	169	7	14
11	70	421	184	2
12	83	333	50	10
13	423	131	2	4

Table 1: Expected surgery durations and yearly frequencies

4.3 MSS

An MSS specially designed for the urology and general surgery department of the Beatrix Hospital is used in this research. The MSS has been constructed by means of optimisation with regards to the OR utilisation and the stability of its resulting demand for hospital beds. The OR-days within the MSS contain time slots that will be used for category A patients and those that will be used as dummy space. All OR-days contain a certain amount of time that is supposed to remain unoccupied for the emergency surgeries. A distinction is made between the dummies assigned to the urology department and the dummies assigned to the general surgery department. No adjustments to this schedule will be made. Only the effects of the ways of applying this MSS will be investigated. The cycle length of this MSS is exactly one week, it contains 40 surgeries divided over a total of ten OR-days, see table 2.

	Day	1	1	2	3	3	4	4	5	5	5
	OR	2	3	3	2	4	2	4	2	3	4
specialism	type										
General surgery	1			1							
General surgery	2		2	1	5			1			
General surgery	6	6	1	2	1						
General surgery	8						1				
General surgery	9								2	1	
General surgery	11						3	1			
General surgery	12							1			
General surgery	dummy							1	1		
Urology	1										3
Urology	3					2					
Urology	4					1					1
Urology	9					1					
Urology	dummy										1

Table 2: The MSS

4.4 Dummy capacity

The MSS just given, contains dummies specifically assigned to certain OR-days. However, we checked all OR-days for possible overcapacity which could be added to the dummy space as well. The available dummy capacity per OR-day is determined by the difference between the total capacity and the planned capacity for type A and C surgeries. The planned capacity for type A surgeries follows directly from the expected durations of the surgeries within the MSS. The time assigned to type C surgeries, the emergency surgeries, results from the following restriction:

$$\textit{Capacity for } C = (1 - \textit{norm utilisation}) \cdot \textit{Total Capacity}$$

For the urology and general surgery department a norm utilisation of respectively 81% and 75% is used. OR-days (3,4) and (5,4) are used for urology procedures. The other eight OR-days are assigned to the general surgery department. Applying this information on the given MSS, combined with available data about the total capacity per OR-day, one gets the following result for the available time left for dummy surgeries, shown in table 3.

Day	1	1	2	3	3	4	4	5	5	5
OR	2	3	3	2	4	2	4	2	3	4
total capacity	480	210	480	480	270	480	480	480	210	480
type A capacity	357	141	356	264	192	357	195	171	85	225
type C capacity	120	53	120	120	51	120	120	120	53	91
time left for type B capacity	3	16	4	96	27	3	165	189	72	164

(light grey)	General surgery dummy space
(dark grey)	Urology dummy space

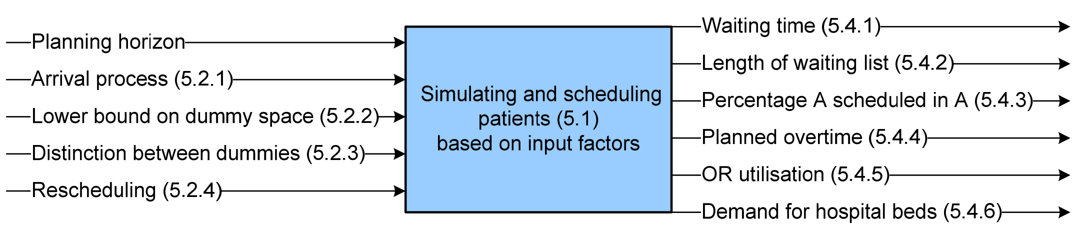
Table 3: division of scheduled time between surgery types per OR-day

The MSS prescribes that the dummy time that results in OR-days (4,4) and (5,2) will be used exclusively for dummies from the general surgery department, while OR-day (5,4) is meant to be used only for dummies from the urology department. How to handle the available capacity in the remaining OR-days is discussed in section 5.2.2.

5. Simulation

We have decided upon the use of simulation, because the problem is too complex to derive all desired performance measures analytically. Although the situation considered in this research is an appointment driven system which resembles the system mentioned in [15], the methods suggested in the article are not suitable for this research. First of all, the vacation models do not include the use of dummy space for appointments. Since the use of dummy space is a main feature of this problem, deriving analytical performance measures disregarding the use of dummies does not seem appropriate. Next, the vacation models include certain aspects that don't apply to this research, like the possibility of arrivals that receive immediate treatment if they arrive during the appropriate time slots (vacations) in case the queue is empty. This is certainly not the case for this problem, since we assume a frozen horizon of the first week in this research. Therefore, the solution methods of the vacation models are not appropriate to evaluate this case analytically.

The simulation process consists of the following steps. First of all, we will explain how a weekly flow of patients is simulated and scheduled within the planning horizon (section 5.1). Next, we will define input and output variables of the simulation process (section 5.2 and 5.3). The input factors are decisions and assumptions regarding the scheduling process. Different input factors result in different output factors. By varying between the possible input factors we will create different scenarios of which the performance will be measured by means of the output factors. A choice is made for all input factors in order to construct a basic scenario (section 5.4). Next, we will explain how much cycles are included within one simulation run. This is done by determining the warm-up period and the run length. The warm-up period contains at least the amount of cycles the system needs in order to arrive at a stable situation (section 5.5). After the warm-up period a certain amount of cycles and simulations is used to evaluate the effect of the inputs in the different scenarios (section 5.6). The figure below gives an overview of the inputs and outputs that will be discussed in the next sections.



5.1 Simulating and scheduling patients

First, assumptions with regard to the arrival process of the patients will need to be made. For our data, the distribution should have an expected value of 40 patients per week (rounded down to an integer number of patients). Then, for each cycle, a realisation out of the arrival distribution is drawn to represent the amount of patients that will be randomly selected out of the patient data. The choice of simulating only once per cycle time, as opposed to simulating a continuous flow of patients, is because of the use of a frozen first cycle in the planning process. All patients arriving during a certain week, are accumulated at the start of the next week. Thus, simulating a continuous flow of patients is not necessary.

This research covers only the scheduling of elective patients, which means that only the flow of category A and B patients will be simulated. In the MSS a certain amount of capacity will be left open explicitly for category C, the emergency patients. For each patient that is selected out of the data in the simulation, its surgery type is known. And for each type of surgery the expected surgery time is given, as in table 1. In this research the expected surgery time will not be simulated. Instead, the surgery is assumed to require exactly the amount of time expected. Since this MSS is already assumed to be optimal with respect to the expected OR utilisation, taking empirical values of the surgery durations in the simulation is not of interest. After we have simulated the patients that have arrived during one week, these patients need to be scheduled for surgery. This is done as described in chapter 3.

5.2 Input decisions

Several decisions and assumptions regarding the input of the simulation and scheduling process have to be made. By selecting one of the possible options for each of those input variables, we create a certain scenario. In the next subsection, one of all possible scenarios is chosen to be the so-called 'basic scenario'. After evaluation of the performance of this basic scenario, the next scenarios will be obtained by differing only one of the input factors, ceteris paribus. In this way the effect of the individual input parameters on the performance measures can be evaluated. Since the main object of this research is investigating the effect of the length of the planning horizon, we will first create scenarios by only changing the length of the planning horizon. Next, the effect of other input decisions, which will be mentioned in this subsection, will be measured likewise.

5.2.1 Arrival process

A common assumption in queuing theory is that the arrival process follows a Poisson distribution. Applying this assumption to this research, we set the expected value of the arrival distribution equal to the average number of patients arriving during one cycle. A drawback of the Poisson distribution is that its first and second moment are the same. That is, the variance equals the expected value. The data does not allow us to make any judgements about the variance of the arrival process, since it contains only surgery dates and no arrival dates. A way to implement a variance that differs from the expectation is by applying another commonly used arrival distribution. In the scenarios we will select between the use of either the Poisson or the Gamma arrival distribution. The Gamma distribution has two parameters which result in separate values for the first and second moment.

The expected value of the arrival distribution, obtained by the data of this research is rounded down to the integer number of 40 patients per week. The Beatrix Hospital performs elective surgeries only during a total of 46 weeks within one year. The MSS has been constructed, using this information. The total number of surgeries performed in one year, was divided by 46 in order to obtain the average number of 40 surgeries performed per week. Arrivals of elective patients can only take place during the 46 weeks in which the hospital is open. Furthermore, the length of the waiting list at the end of the year was about the same as at the beginning of the year. Thus, the average number of arrivals per week equals the average number of performed surgeries per week.

For a Poisson distributed variable X we know that $E(X) = Var(X) = \lambda$, which means that using a Poisson distribution with $\lambda = 40$ implies not only a mean of 40 patients per week, but also a variance of 40. In case X is Gamma(α, β) distributed it holds that $E(X) = \alpha\beta$ and $Var(X) = \alpha\beta^2$. Applying this to our data with $E(X) = 40$ we will define scenarios using the following three settings:

$$\left.\begin{array}{l}\alpha = 80 \\ \beta = 0.5\end{array}\right\} Var(X) = 20 \qquad \left.\begin{array}{l}\alpha = 20 \\ \beta = 2\end{array}\right\} Var(X) = 80 \qquad \left.\begin{array}{l}\alpha = 10 \\ \beta = 4\end{array}\right\} Var(X) = 160$$

We do not simulate a gamma distribution with a variance of 40 since this resembles the Poisson distribution with $\lambda = 40$ quite a lot. Furthermore, we decided to include a benchmark scenario at which the constant value of 40 patients will be simulated every cycle, which implies zero variance. The probability density functions of the above three Gamma distributions are shown in figure 11 in appendix B.1. Figure 12 in the same appendix shows the similarity between the probability density functions of a Poisson and a Gamma distribution, both with $E(X) = Var(X) = 40$.

5.2.2 Lower bound on dummy space

The MSS prescribes that the available dummy time that results in OR-days (4,4) and (5,2) will be used exclusively for dummies from the general surgery department, while OR-day (5,4) is meant to be used only for dummies from the urology department (see table 2, the MSS). Using only the available dummy capacity of these three OR-days, would be a waste of the available time in the other OR-days shown in table 3. OR-day (3,2) for example contains an unplanned capacity of 96 minutes which could well be used for scheduling dummy surgeries with a relatively short expected duration. On the other hand, table 3 also shows that the capacity of certain OR-days is almost fully assigned to type A and type C surgeries. OR day (1,2) for example, contains no more than three minutes of unplanned capacity. Because we would like to avoid planned overtime as much as possible, it seems reasonable to prescribe a lower bound on the available unplanned capacity before assigning it to the dummy space.

For our data, the shortest expected surgery duration is 37 minutes (urology, surgery type 3). In practice it does happen that hospitals allow some planned overtime when scheduling surgeries. If the surgery with the expected duration of 37 minutes would be scheduled in OR-day (3,4) one would allow 10 minutes of planned overtime. Since cases like these happen in practice, it seems reasonable to define a minimum number of minutes of free capacity that an OR-day should contain, in order for it to be used for dummy surgeries. Another way to define the lower bound could be to take a certain percentage of the shortest expected surgery duration. Including the 27 minutes of OR-day (3,4) in the dummy space and excluding the OR-days with less time available for dummy space could be a choice for this research case. Adding a maximum allowed overtime of 50 percent of the type C durations to the dummy space that remains after applying the lower bound, one obtains the available dummy capacities given in table 4. Six OR-days remain in which dummy scheduling is allowed. Three of which have a specific department to which they are assigned, while the other three can be used for surgeries out of both the urology and the general surgery department.

Day	1	1	2	3	3	4	4	5	5	5
OR	2	3	3	2	4	2	4	2	3	4
• dummy space with lower bound 20 minutes	0	0	0	96	27	0	165	189	72	164
• dummy space with lower bound & 50% of C for planned overtime	0	0	0	156	53	0	225	249	99	210

(light grey)	General surgery dummy space
(dark grey)	Urology dummy space

Table 4: available dummy space, without and with the possibility of planned overtime

In general, for deciding upon the available dummy space within the MSS, a lower bound needs to be defined on the remaining unassigned capacities in the OR-days. Within the scenarios we will vary this lower bound between two absolute levels: 20 minutes (which includes OR-day (3,4)) and 40 minutes (which excludes OR-day (3,4)).

Looking at the available dummy capacity in table 4 and the expected surgery durations of each surgery type, we notice five surgery types of which the expected duration exceeds the maximum available dummy space of 225 minutes. These five are all category B surgeries and they do not occur very often, as the yearly frequencies show. Still, a way to handle such surgeries has to be determined. We have decided upon scheduling these surgeries in the fist OR-day which has one of the three highest dummy capacities. This means using OR-day (4,4), (5,2) or (5,4) only if no dummy capacity of the OR-day has been already used for other surgeries. To ensure available capacity for these surgeries, we schedule these types before any other type within the planning process.

5.2.3 Distinction between dummies

In the MSS a distinction is made between the OR-days that are explicitly assigned to one specialty (see table 4). By allowing other specialties to use the restricted dummy space as well, one obtains more flexibility in the use of the MSS. It's questionable whether this relaxation has that much influence on the performance measures. This will be investigated by adding a scenario which disregards the distinction between urology and general surgery dummies.

5.2.4 Shift 'late' A surgeries from dummy to A space

Once surgeries have been scheduled for the coming period, it could be beneficial to consider rescheduling some of those surgeries. In particular, we would like to consider scenarios in which category A surgeries that were initially scheduled in dummy time slots are rescheduled, if possible, into newly available category A time slots. New time slots are the ones that are present in the cycle which is added to the future planning schedule, the moment the horizon 'rolls' forward. The implementation of this replanning feature is considered because it provides a more predictable (or stable) schedule. Waiting time on the other hand increases for the particular patients.

Surgery a_{is} was initially scheduled within dummy space only if all time slots for a_{is} within the planning horizon were already occupied. When more than one surgery a_{is} is initially scheduled within dummy space, selecting the one which was scheduled the furthest ahead in time gives the least increase in waiting time. For implementing this feature, we need to define a replanning horizon. This is the number of cycles which are considered to reschedule surgeries from, starting from the end of the planning horizon and counting backwards the number of cycles of the replanning horizon.

To illustrate this replanning feature, figure 3 shows rescheduling surgery type a_{42} for a planning horizon of four weeks and a replanning horizon of two weeks. The available time slots are located in the MSS on the fifth and tenth OR-day, (3,4) and (5,4).
The order of scheduling is indicated in figure 3 by ranks 1, 2 and 3. This means that rescheduling takes place in case one or more type a_{42} time slots remain empty after the regular new arrivals have been scheduled. In this example, only one type a_{42} patient arrived during week 1, which results in one OR-day left empty in week 5. The first type a_{42} surgery considered for rescheduling is the one originally scheduled in dummy space of week 4. Thus, this surgery is rescheduled to the available type A time slot of week 5. In case both OR-days of week 5 would have been empty, the surgery scheduled in dummy space of week 3 would have been rescheduled as well. It is possible that type A time slots of the last week remain empty, even after rescheduling. This happens when the number of available a_{is} time slots after scheduling the new arrivals exceeds the number of a_{is} surgeries present in the replanning horizon. We will construct scenarios by defining different replanning horizons containing an integer number of weeks. Note that the length of the replanning horizon cannot exceed the length of the planning horizon -1 because of the rolling horizon.

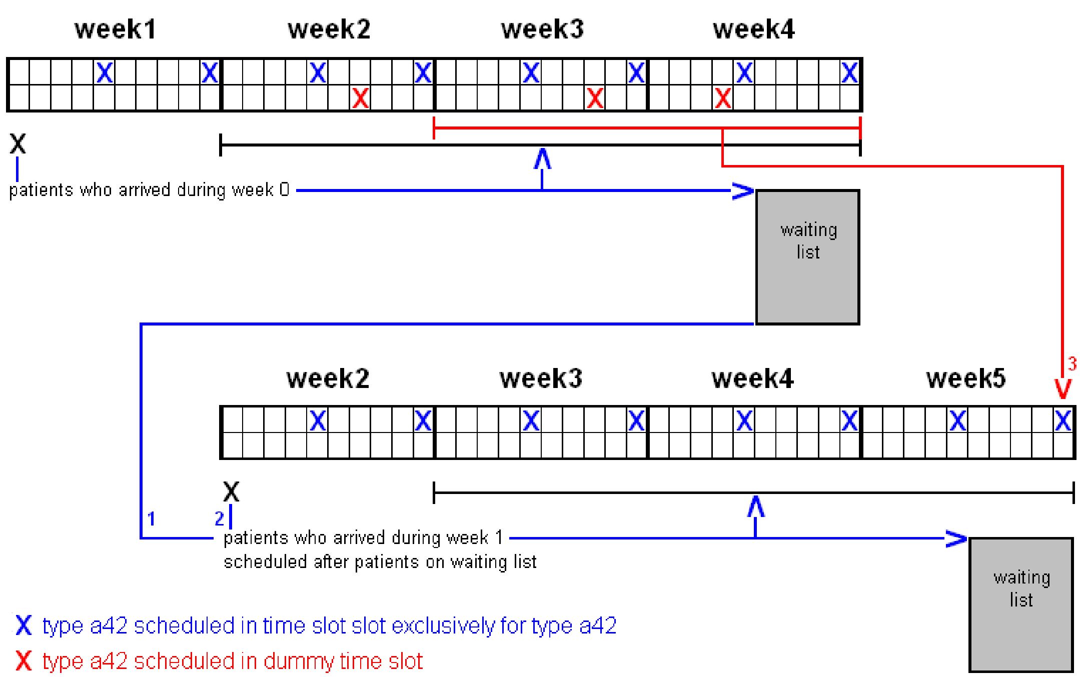


Figure 3: Rescheduling a_{42} from dummy to A slot. Planning horizon: 4. Replanning horizon: 2.

5.3 The basic scenario

For all of the previously mentioned input factors in the simulation and planning process, one of the possible options is selected in order to obtain the basic scenario. The input parameters we have selected to define the basic scenario are the following:

- The arrival process is assumed to be Poisson distributed, with the parameter for expected value and variance of 40 patients per week.
- The lower bound on available unscheduled capacity for it to be used as dummy space is set at 20 minutes. Applying this to our data, OR-day (3,4) is included in the available dummy space.
- The distinction between urology and general surgery dummies is implemented as prescribed by the MSS.
- No shifting of category A surgeries from dummy space to newly available category A time slots will take place.
- The planning horizon chosen for the basic scenario is five weeks.

5.4 Performance measures

Since we are interested in the effects of the length of the planning horizon and the other input variables just mentioned, we need to define the measures by means of which these effects will be evaluated.

With the use of a relatively long planning horizon two issues occur. On the one hand, we are likely to obtain a stable schedule, since category A patients are provided with more available time slots of their own type and thus they are less likely to be scheduled in a dummy OR-day. On the other hand, the waiting time for those patients is likely to be larger than it would have been with the use of a shorter horizon. Since short waiting times and stability of the schedule are the two conflicting issues here, we are interested in measurements of the patients waiting times, the amount of patients that end up on a waiting list and the number of category A patients do and do not end up being scheduled within a time slot of their own surgery type. Also, we will measure the amount of planned overtime, the OR utilisation and the demand pattern for hospital beds resulting from the scheduling process. These performance measures are described in the next subsections.

5.4.1 Waiting time

The waiting time of the patients will be expressed by the number of weeks it takes between the arrival of a patient and the scheduled date of surgery. Consider a patient arriving at week 1. The patients that have arrived during week 1 are accumulated at the beginning of week 2 and are scheduled in available time slots starting from week 3. The reason for this is the prescribed minimum period of one week between the date of arrival and the surgery date. On average though, if a patient arriving at week 1 is scheduled for surgery at week 3, the patient has been waiting for two weeks. During the simulation we will keep track of the waiting times in integer number of weeks, which implies a minimum simulated waiting time of two weeks for any patient. We will not only consider the average waiting time per patient, but also the distribution of the waiting time. Finally, we will also consider the proportion of patients that have to wait eight weeks or less. We include this last measure because hospitals often agree with health insurance companies upon an upper bound of this proportion.

5.4.2 Length of waiting list

The length of the waiting list is another measure to make judgments about a scenario. Like mentioned earlier, long waiting list don't seem desirable from the patients' point of view. For the schedule maker however, the more patients on the waiting list, the less likely it gets that time slots will not be filled. In other words, long waiting lists are more likely to provide a stable schedule.

Because of the problem description we only consider patients to be on the waiting list when they don't fit within the current planning horizon. The consequence of this definition is that a short waiting list or no waiting list at all, does not necessarily imply that the patients don't have to wait relatively long before their surgery, especially when a long planning horizon is used. Thus, this performance measure is one that is only of interest for the schedule maker.

When the capacity of all OR-days exceeds the expected OR utilisation, which is the case for our data, we expect the waiting list to decrease when we increase the planning horizon. A patient ends up on the waiting list when its surgery type does not fit within the planning horizon. For longer planning horizons, weekly arrivals which exceed the number of available time slots within one cycle, can be compensated by lower arrival amounts of other weeks, all within the same planning horizon. Statistically, adding different realisations out of the same arrival distribution results in a decrease in the standard deviation of the average of the realisations. Thus, the expected amount of patients that end up on the waiting list decreases, especially when the total available capacity exceeds the expected need for OR capacity.

5.4.3 Percentage of A scheduled in A

One way to express the stability of the schedule is by keeping track of the percentage of type A surgeries that will be scheduled in the type A time slots of the MSS. Type A patients that don't get to be scheduled within category A time slots are the ones that were placed in dummy space because the A time slots were already fully occupied within the horizon. As we just explained, the expected amount of patients ending up on the waiting list decreases as the planning horizon increases. The same holds for the proportion of type A patients ending up being scheduled in a dummy time slot. Thus, we expect increasing rates of type A patients that *do* fit within type A time slots, as we increase the planning horizon.

5.4.4 Planned overtime

Since we use the expected surgery durations in the simulation, the amount of overtime resulting from category A surgeries within the appropriate time slots prescribed by the MSS, is left out of consideration. The expected overtime in the OR-days containing dummies on the other hand, is a feature that defers per simulation, depending on the surgeries that are scheduled in the OR-day. Like mentioned before, each OR-day contains a certain amount of time that is assigned to the possible arrival of emergent patients. A dummy OR-day contains planned overtime when some of the time for category C patients is used for scheduling the elective patients. Because of this possibility, we will keep track of the total amount of this kind of expected overtime and the number of times it occurs.

5.4.5 OR utilisation

We define OR utilisation as the percentage of OR capacity that is used to schedule elective surgeries in. Since surgery durations are not taken stochastically in the simulation, we measure the expected surgery time of the scheduled surgeries relative to the available OR capacity. We will distinguish three types of OR utilisation:

1. The proportion of the total available minutes of OR capacity of all OR-days (the first row of table 3) that is used for scheduling all types of surgery.
2. The utilisation of the time capacity of the type A slots in the MSS (the second row of table 3). These slots can only be occupied by type A surgeries.
3. The utilisation of the dummy time slots, including the extra capacity of 50 percent of the capacity for type C surgeries (the second row of table 4). The dummy slots are used to schedule both type A and type B surgeries in.

We expect no change in the overall OR utilisation when changing the planning horizon since the overall expected surgery time only depends on the expected amount of arriving patients. For the OR utilisation of type A and B slots on the other hand, we do expect a changing pattern when changing the planning horizon. All elective surgeries are scheduled at least one week ahead. It is possible that the upcoming week contains certain type A time slots that remain empty. When this happens the hospital experiences lower utilisation of the type A time slots. The longer the planning horizon, the less likely this is to happen. For a planning horizon of two weeks, the only arrivals available to occupy the time slots of the upcoming week are last week's arrivals. For longer planning horizons however, the arrivals of previous weeks also could have possibly helped avoiding empty type A slots of the upcoming week.

5.4.6 Demand for hospital beds

In the simulation process, each cycle a number of patients is randomly drawn from the given patient data. For each of these patients the number of days the patient spent in the hospital to recover from surgery is given. These individual periods are used in the simulation process by means of which the daily number of required hospital beds can be reported in the end. The levels of required hospital beds were taken into account when the MSS was designed. The goal is to keep the demand for hospital beds within a certain range. Whether or not this range will also result from applying the MSS to the simulated patient data, will be shown by the performance measures which give the mean and standard deviation of the demand for hospital beds. For the demand levels we make a distinction between weekends and the days from Monday to Friday. The average levels are likely to be different since no patients are scheduled for surgery during weekends.

5.5 Warm-up period

Plotting the weekly results of a performance measure, one can visually determine the time period it takes before the performance remains in a certain stationary situation. Stationarity means that the unconditional mean, unconditional variance and autocorrelations of the observations are constant over time (p.130 [16]). Applying this to all performance measures and taking the maximum of the required time periods for the process to reach stationarity, gives us the warm-up period.

Of course, some performance measures will be more suitable than others when it comes to identifying the warm-up period. The performance measure ‘length of the waiting list’ for example doesn’t seem suitable because the waiting list might stay empty forever, especially when using a long planning horizon. A more suitable measure would be to look at the number of patients that are scheduled at least two weeks ahead. Consider for example a planning horizon of four weeks. Patients simulated at the start of week i are scheduled between the start of week $i+1$ and the end of $i+3$, if possible. At the beginning of the simulation process, say at week 1, it is not likely that patients simulated at the very first cycle end up being scheduled somewhere in week 4 already. But as time ‘rolls on’ the use of time slots beyond the upcoming week gets more likely. Thus, the number of patients scheduled at least two weeks ahead, which means in this case between week $i+2$ end $i+4$, seems a good measure for determination of the warm up period. Note that this measure doesn’t exist when a planning horizon of two weeks is used.

The longer the planning horizon is chosen, the more time the system is expected to take to become stationary. Therefore, we have measured the weekly number of scheduled patients beyond the upcoming week for a planning horizon of 50 weeks. This is the highest length of the planning horizon that we will investigate. We’ve plotted four of these time series to determine the warm-up period.

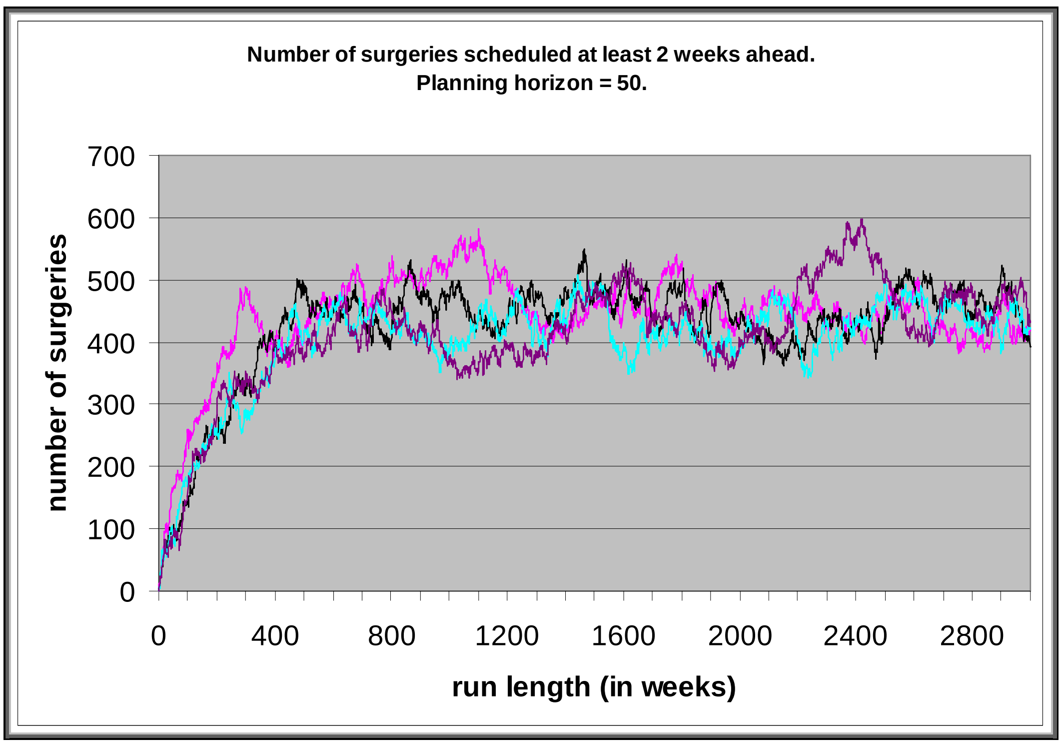


Figure 4: times series of the number of surgeries scheduled at least two weeks ahead of four independent simulations with a planning horizon of 50 weeks. Warm-up estimated at 800 weeks.

Based on figure 4 we have determined a warm-up period of 800 weeks for a planning horizon of 50 weeks. Appendix B.2 shows that the plotted time series for the planning horizons of 5 and 15 have lower required warm-up periods, as expected. Since we are looking for a warm-up period which can be used for all scenarios, we define a warm-up period of 800 weeks.

5.6 Run length and number of simulations

One simulation consists of a warm-up period followed by a certain run length. Of course, the longer the run length is chosen, the more accurate the estimates of the performance measures will become. Still, it is not sufficient in this case to perform only one simulation per scenario. The issue here is the dependence between the performance measures of successive cycles. The demand for hospital beds for example, depends for a great part on the number of hospital beds occupied at the previous day. We have chosen a run length of 2000 cycles. There is no specific reason for this amount other than the goal of increasing accuracy, still obtainable within reasonable amount of runtime of the simulation program.

To make justified statistical judgments about the performance the following holds: After simulation *i* one can determine the sample mean X_i for a certain performance measure. When *n* separate independent simulations have been performed, a $(1-\alpha)\%$ confidence interval can be constructed by:

$$\left[\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right] \quad (1)$$

Where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$

and $z_{\alpha/2}$ the z-statistic of the standard normal distribution. That is, in case $n \geq 30$, otherwise a t-statistic is required.

Note that X_i does not denote the value of a performance measure at *cycle i*, but the sample *mean* of the performance measure over all cycles within simulation *i*, in our case a simulation with warm-up period of 800 and run length of 2000. Doing so, $(1-\alpha)\%$ of all intervals constructed in this manner will contain the actual mean.

Equation (1) shows that increasing the number of independent simulations *n* decreases the size of the confidence interval around the mean of the performance measure by a factor $\sqrt{n}$. Determining the size of *n* depends on one's desire for the absolute size of the intervals. When a test size of n_1 does not give the desired interval yet, one can increase n_1 to n_2 such that $z_{\alpha/2}\, s_1 \big/ \sqrt{n_2}$ has the desired size, where s_1 is the variance of the n_1 sample means [17]. For the construction of the confidence intervals, we have set $n = 100$, which can be adjusted if requested.

For the performance measure which gives the standard deviation of the demand pattern for hospital beds we need to define X_i as the variance of the demand pattern of simulation *i*. Doing so, we first calculate the average variance $\overline{Var}$ using the *n* simulations by:

$$\overline{Var} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ , with } X_i \text{ the variance of simulation } i.$$

In the end, we can report the result by transforming $\overline{Var}$ to a standard deviation $\overline{sd}$ again. Taking the average of the standard deviations of the *n* demand patterns won't be correct. To see this, take for example *n* = 2 and A and B two demand patterns of separate simulations. The following applies for the calculation of the average standard deviation:

$$\overline{Var} = \overline{Var(A) + Var(B)} = \frac{1}{2}\left(Var(A) + Var(B)\right)$$

$$\overline{sd} = \sqrt{\frac{1}{2}\left(Var(A) + Var(B)\right)} \neq \frac{1}{2}\left(\sqrt{Var(A)} + \sqrt{Var(B)}\right) = \frac{1}{2}\left(sd(A) + sd(B)\right).$$

6. Results

Table 5 summarizes the mean values of the performance measures for all scenarios constructed, based on 100 simulations with warm-up 800 and run length 2000.

input variables					performance measures												
AP	LB	DD	RH	PH	WT	8W	WL	AA	NPO	DPO	OR	ORA	ORB	DB	(sd)	DBwe	(sd)
P(40)	20	Y	0	2	2.24	99.86	9.44	83.95	4.45	48.14	71.12	86.33	86.73	17.13	3.57	14.35	3.37
P(40)	20	Y	0	3	2.58	100.00	1.18	90.42	2.46	58.76	71.13	92.70	71.64	17.16	3.22	14.30	3.17
P(40)	20	Y	0	4	3.04	100.00	0.34	93.01	1.62	79.71	71.13	95.62	64.72	17.18	3.05	14.24	3.03
P(40)	20	Y	0	5	3.52	100.00	0.13	94.23	1.40	88.00	71.12	97.05	61.34	17.19	2.97	14.21	2.96
P(40)	20	Y	0	6	3.99	100.00	0.03	94.91	1.34	90.82	71.12	97.85	59.44	17.20	2.94	14.20	2.92
P(40)	20	Y	0	7	4.47	100.00	0.01	95.33	1.32	92.27	71.12	98.33	58.30	17.20	2.91	14.19	2.90
P(40)	20	Y	0	8	4.94	100.00	0.00	95.61	1.31	92.90	71.12	98.65	57.55	17.20	2.90	14.19	2.89
P(40)	20	Y	0	9	5.41	87.71	0.00	95.80	1.30	93.19	71.12	98.87	57.03	17.20	2.89	14.18	2.87
P(40)	20	Y	0	10	5.87	77.26	0.00	95.95	1.30	93.32	71.13	99.03	56.66	17.20	2.88	14.18	2.87
P(40)	20	Y	0	11	6.33	68.96	0.00	96.06	1.30	93.40	71.13	99.15	56.38	17.20	2.88	14.18	2.86
P(40)	20	Y	0	12	6.79	63.71	0.00	96.15	1.30	93.44	71.13	99.24	56.16	17.20	2.87	14.18	2.86
P(40)	20	Y	0	13	7.24	60.00	0.00	96.23	1.30	93.47	71.12	99.31	55.99	17.20	2.87	14.18	2.85
P(40)	20	Y	0	14	7.67	56.80	0.00	96.28	1.30	93.49	71.12	99.37	55.85	17.21	2.86	14.18	2.85
P(40)	20	Y	0	15	8.10	55.03	0.00	96.33	1.30	93.50	71.12	99.41	55.74	17.21	2.86	14.18	2.85
P(40)	20	Y	0	25	11.59	45.89	0.00	96.58	1.29	93.52	71.12	99.61	55.19	17.20	2.85	14.17	2.83
P(40)	20	Y	0	50	18.97	44.07	0.00	96.70	1.29	93.92	71.12	99.72	54.96	17.19	2.84	14.16	2.82
C(40)	20	Y	0	5	3.50	100.00	0.06	94.50	1.37	89.50	71.16	97.37	60.73	17.22	2.90	14.21	2.90
G(80,0.5)	20	Y	0	5	3.51	100.00	0.09	94.33	1.38	88.84	71.18	97.21	61.17	17.21	2.95	14.22	2.93
G(20,2)	20	Y	0	5	3.54	100.00	0.25	93.92	1.48	85.15	71.15	96.79	62.06	17.19	3.06	14.24	3.02
G(10,4)	20	Y	0	5	3.58	100.00	0.67	93.37	1.67	79.64	71.18	96.29	63.36	17.19	3.22	14.26	3.14
P(40)	40	Y	0	5	3.52	100.00	0.13	94.23	1.40	87.99	71.12	97.05	61.34	17.19	2.97	14.21	2.96
P(40)	20	N	0	5	3.49	100.00	0.07	94.23	1.38	88.44	71.12	97.05	61.34	17.19	2.99	14.21	2.97
P(40)	20	Y	1	5	3.57	100.00	0.12	94.42	1.31	88.60	71.12	97.35	60.62	17.18	3.01	14.25	3.00
P(40)	20	Y	2	5	3.58	100.00	0.13	94.41	1.35	87.81	71.12	97.37	60.59	17.18	3.00	14.23	2.99
P(40)	20	Y	3	5	3.64	100.00	0.19	94.49	1.39	87.22	71.12	97.46	60.36	17.18	2.98	14.23	2.96
P(40)	20	Y	4	5	3.74	100.00	0.24	94.66	1.35	87.11	71.12	97.64	59.93	17.18	2.95	14.23	2.93
P(40)	20	Y	1	2	2.26	99.86	9.44	86.84	3.94	50.05	71.12	89.35	79.59	17.15	3.43	14.30	3.28

AP	Arrival Process: P for Poisson (parameter1), C for Constant (parameter1), G for Gamma (parameter1, parameter2)
LB	Lower Bound (in minutes) on available unplanned capacity in order for it to be assigned to dummy space
DD	Distinction made between urology and general surgery Dummy space: Y(es) or N(o)
RH	Replanning Horizon (0 if no rescheduling takes place)
PH	Planning Horizon
WT	Average Waiting time (in weeks)
8W	Percentage of patients that experience a waiting time of Eight Weeks or less
WL	Average length of the waiting list (number of patients)
AA	Percentage of category A surgeries scheduled in type A time slots
NPO	Average Number of times Planned Overtime is scheduled
DPO	Average Duration of Planned Overtime
OR	OR utilisation
ORA	OR utilisation of the time scheduled for category A surgeries
ORB	OR utilisation of the time scheduled for category B surgeries, including 50% of time for C
DB	Average (and standard deviation) of daily Demand for hospital Beds during week days
DBwe	Average (and standard deviation) of daily Demand for hospital Beds during weekends

Table 5: Mean values over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks.

The values of the performance measures given in table 5 are the averages of the total of n = 100 simulations, each performed containing a warm-up length of 800 cycles and a run length of 2000 cycles. Table 6 in the appendix contains the 95% confidence intervals around these averages of the performance measures, constructed as described in section 5.6.

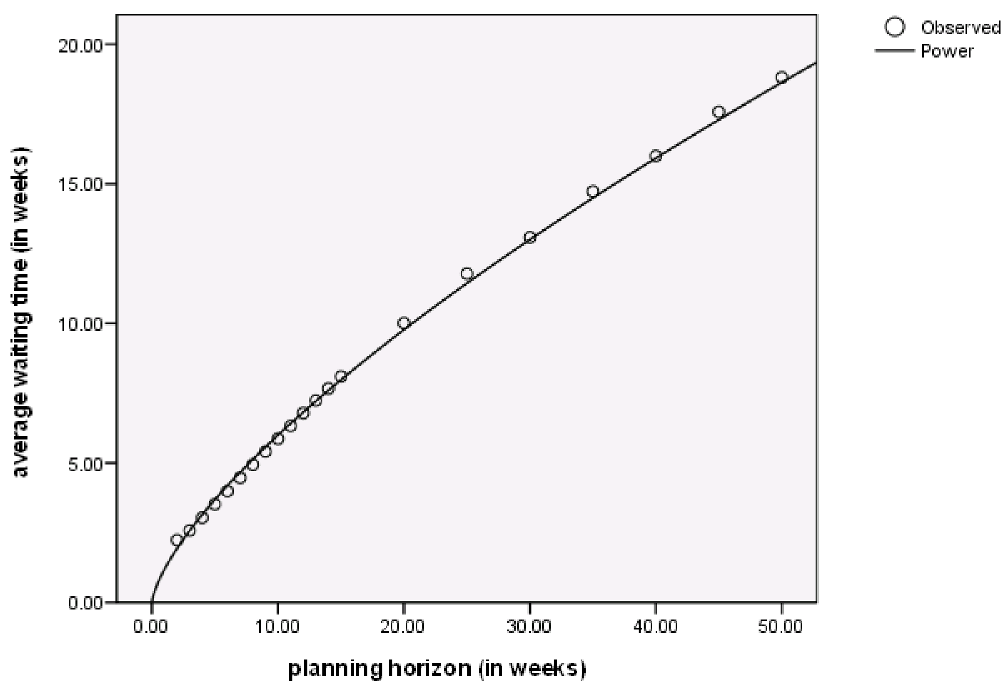
The interpretation of the results in table 5 will be discussed in this chapter. The scenarios in the upper part of table 5 are constructed by only adjusting the length of the planning horizon from the basic scenario. Using the results of these scenarios section 6.1 discusses the influence of the planning horizon. Next, new scenarios were constructed by changing one of the other input variables from the basic scenario, ceteris paribus. The influence of these input variables is discussed in section 6.2.

6.1 Effects of the planning horizon

The influence of the planning horizon is discussed in this section, using the results in the upper part of table 5 and by looking at some effects in more detail using additional descriptive statistics.

6.1.1 Waiting time

Clearly, as shown in table 5, the longer the planning horizon, the longer a patient has to wait on average before surgery. The next figure shows the dependence of these two variables graphically. Using the statistical software SPSS we have estimated the relation between planning horizon (P) and the average waiting time (W) by means of the function $W = 1.187P^{0.704}$.



Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	B1
Power	.996	4412.315	1	19	.000	1.187	.704

Figure 5: relation between planning horizon and average waiting time

To illustrate the distribution of the waiting time figure 6 shows the proportion of patients that experienced a waiting time of size $i \in \{2,3,\ldots 50\}$ weeks, when a planning horizon of 50 weeks is used. The probabilities are estimated using 100 simulations, each consisting of a warm-up period of 800 followed by a run length of 2000. Appendix B.4 contains the waiting time distributions for smaller planning horizons.

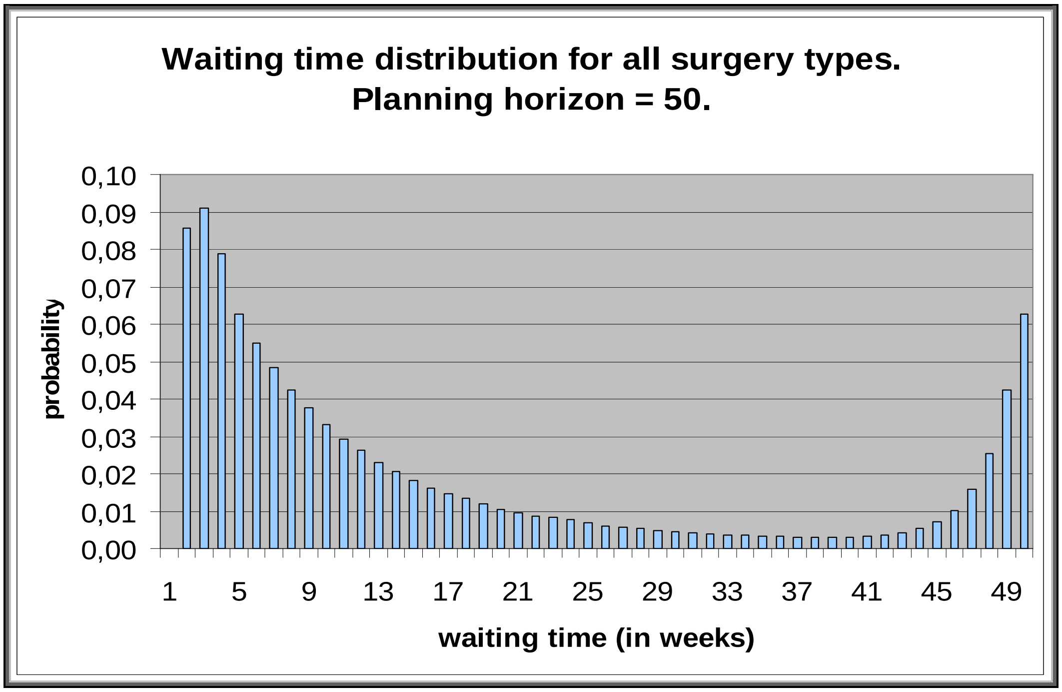


Figure 6: Waiting time distribution for all types of surgery. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. Planning horizon: 50 weeks.

The bar chart indicates that most patients either have to wait relatively long or relatively short. To investigate this issue further we look at the distribution of individual surgery types separately. The two charts in appendix B.5 give the distribution of the waiting time for surgery types a_{91} and a_{12} respectively. Clearly, there is a difference between both distributions. The chart of surgery type a_{91} shows increasing amounts of long waiting time near the end of the planning horizon, while the next chart shows that increasing waiting times for type a_{12} surgeries have decreasing probabilities. The explanation of this phenomenon can be found looking at the construction of the MSS. If we calculate the weekly frequencies λ_{is} (using the yearly frequencies from table 1, divided by 46 weeks), we find that the frequency λ_{91} was rounded down to obtain the number of surgeries in the MSS (n_{91}), while the weekly frequency λ_{12} was rounded up. On average surgery type a_{91} occurs 3.39 times per cycle and surgery type a_{12} 2.98 times. Rounding these frequencies results in $n_{is} = 3$ surgeries included in the MSS for both type a_{91} and type a_{12} surgeries (see table 2).

In general, if λ_{is} exceeds n_{is}, all time slots specifically meant for type a_{is} within the planning horizon will gradually become occupied as time rolls on. In that case, the only time slots available specifically for type a_{is} are the new ones of the MSS that are added at the end of the planning horizon as time rolls on. In case that number of available time slots is not sufficient for the arrivals within one cycle, some patients of type a_{is} will be scheduled within the fist available dummy time slot. These dummy slots are situated near the beginning of the planning horizon. This process is reflected by the distribution of the waiting time for surgery type a_{91}. When λ_{is} lies below n_{is}, the probability of a certain waiting time seems to decline exponentially towards zero. Because the amount of available dummy capacity exceeds the expected required dummy capacity, the same applies for waiting times for surgeries scheduled in dummy time slots. Combining the two possibilities just mentioned, results in the distribution pattern of figure 6. These findings show that, judging from the patients' point of view, using a long planning horizon is not beneficial. The possibility that a type a_{is} patient with $\lambda_{is} > n_{is}$ has to wait (almost) as long as the length of the planning horizon is rather large.

Finally, the proportion of patients that experience waiting times of eight weeks or less, is directly obtained by cumulating the probabilities of waiting times from two to eight weeks in the waiting time distributions. As the average waiting time increases when the planning horizon increases, this proportion shows a decreasing pattern.

6.1.2 Length of waiting list

The length of the waiting list declines to zero as we increase the planning horizon. Because the available OR capacity exceeds the expected demand for OR capacity, the probability of shortage of capacity declines as the planning horizon increases, which explains the pattern just mentioned.

The length of the waiting list can be related to the distributions of waiting time in appendix B4. The figures show that for small planning horizons some patients have waiting times which exceed the length of the planning horizon. This can only be the case for patients that did not fit within the planning horizon they faced right after their arrival. If that happens, these patients are placed on the waiting list. Thus, we have a positive average length of the waiting list, if and only if the waiting time distribution shows positive probability of waiting time longer than the planning horizon.

6.1.3 Percentage A scheduled in A

Table 5 clearly shows an increase in percentage of type A surgeries scheduled in A slots as we increase the planning horizon. The shorter the planning horizon, the higher the expected probability gets that the total amount of arrivals of type A patients within the planning horizon exceeds the available type A time slots within the planning horizon, resulting in dummy use for type A patients. Thus, using higher planning horizons decreases this 'risk' and increases the percentage of type A surgeries scheduled in type A time slots.

Of course, the percentage of A surgeries scheduled in A slots will never exceed 100 percent. Suppose the expected arrival rates λ_{is} for all surgery types $i \in I$ and $s \in S$ lie below n_{is}. In that case, the higher the planning horizon is chosen, the lower the probability becomes that type A patients won't fit within their time slots of the MSS within the planning horizon. Thus, the more likely it gets that all type A patients are scheduled within type A slots. Thus, when all type A surgeries have expected arrival rates $\lambda_{is} < n_{is}$, the percentage of A scheduled in A has an asymptotic upper bound of 100 percent. Now, suppose $\lambda_{is} > n_{is}$. The percentage of type a_{is} scheduled in a_{is} time slots will asymptotically reach un upper bound upper of n_{is} / λ_{is}. If the MSS contains amounts of n_{is} of which some were obtained by rounding up the value of λ_{is} and others by rounding down the value of λ_{is}, combining both types of upper bounds results in an overall asymptotic upper bound of U, where

$$U = \frac{\sum_{i\in I}\sum_{s\in S}\min(\lambda_{is}, n_{is})}{\sum_{i\in I}\sum_{s\in S}\lambda_{is}}. \qquad (2)$$

Looking at the results for AA in table 5 we recognize the above mentioned asymptotic behaviour. Applying formula (2) for U to our data, results in an asymptotic upper bound of 97.15 percent[1] of A surgeries scheduled in A slots as we increase the planning horizon. This value is the result of the fact that rounding the expected arrival rates of our data resulted in values of n_{is} almost all below the expected values λ_{is}.

6.1.4 Planned overtime

Since we included the possibility of scheduling surgeries using at most 50% of the capacity for type C surgeries, we are interested in the effect of the planning horizon on the average duration of overtime (DPO) and the amount of times scheduling, using planned overtime occurs (NPO). The results show a decrease in number of times planned overtime is scheduled and an increase in the average duration of planned overtime as we increase the planning horizon.

The explanation for this pattern is related to the percentage of A surgeries scheduled in A slots. As we just discussed, increasing the planning horizon, increases the percentage of A surgeries scheduled in A slots. Or, increasing the planning horizon results in a decrease of dummy use for type A. Type B surgeries are always scheduled in dummy space, so the amount of overtime resulting from type B surgeries is not influenced by the planning horizon. Increasing the planning horizon, the number of times planned overtime is scheduled (NPO) decreases as a result of a decrease of dummy use for type A surgeries.

On average the duration of type A surgeries is shorter than the duration of type B surgeries, especially since some type B surgeries have expected durations exceeding the maximum available dummy space per OR-day. Because of a decrease in dummy use of type A surgeries when increasing the planning horizon, the greater the influence of relatively large overtime caused by type B surgeries becomes. This explains the increase of the average duration of overtime (DPO), shown in table 5.

6.1.5 OR utilisation

As we increase the planning horizon the average total OR utilisation remains equal. This statement is justified looking at the 95 percent confidence intervals, which all show an overlap. This is not surprising, because the expected total required surgery time of all arrivals is independent of the planning horizon. All surgeries, no matter what the waiting time might be, will eventually be scheduled.

---

[1] The weekly frequencies were obtained by first dividing the yearly frequencies from table 1 by 46 weeks. Next, these values were multiplied by (40/40.48) in order to obtain the values λ_{is} because in the simulation we use the rounded overall arrival rate of 40 patients per week, instead of the arrival rate of 40.48 patients per week resulting from the data.

The utilisation of the time slots for type A on the other hand, increases as the planning horizon increases. The reason for this was mentioned before. The longer the planning horizon, the less likely it gets that type A slots of future cycles remain empty. Since for longer planning horizons more type A surgeries are scheduled within A time slots, the use of dummy space for A surgeries occurs less often. Therefore, the utilisation of type B time slots decreases as the planning horizon increases.

Note that there is a difference between the asymptotic upper bound for the percentage of type A surgeries scheduled in type A slots and the expected utilisation of OR capacity for type A slots. The more the arrival rate for surgery type a_{is} exceeds the number of available time slots in the MSS for this surgery type n_{is}, the lower the asymptotic upper bound for AA gets. The OR utilisation of type A time slots on the other hand, is more likely to reach towards 100 percent when we face more expected arrivals than the number of available time slots for type A surgeries. When considering arrival rates λ_{is} below n_{is}, we get the opposite effect of an upper bound of AA of 100 percent and an expected OR utility below 100. We get:

- $AA \xrightarrow[ph \to \infty]{} \frac{n_{is}}{\lambda_{is}}$ and $ORA \xrightarrow[ph \to \infty]{} 100\%$ when $\lambda_{is} > n_{is}$
- $AA \xrightarrow[ph \to \infty]{} 100\%$ and $ORA \xrightarrow[ph \to \infty]{} \frac{\lambda_{is}}{n_{is}}$ when $\lambda_{is} < n_{is}$
- $AA \xrightarrow[ph \to \infty]{} 100\%$ and $ORA \xrightarrow[ph \to \infty]{} 100\%$ when $\lambda_{is} = n_{is}$

Where $\xrightarrow[ph \to \infty]{}$ denotes the asymptotic behaviour of the performance measures 'percentage of type A surgeries scheduled in type A time slots' and the 'OR utilisation of the type A time slots' respectively, as we increase the planning horizon.

6.1.6 Demand for hospital beds

In table 5 a distinction is made between the average daily demand for hospital beds from Monday through Friday and the average daily demand during weekends. Taking the weighted average results in 16.34 beds occupied on average for all days of the week, for all planning horizons. When we calculate the intervals of the weighted values for all planning horizons in the same manner, all intervals coincide. This indicates no influence of the planning horizon on the average daily demand for hospital beds. Like the overall OR utilisation, the average daily demand for hospital beds also only depends on the arrival rate of patients.

Although the average demand level over *all days of the week* doesn't depend on the planning horizon, the distribution of demand *between week days and weekend days* does seem to be influenced by the planning horizon. Comparing the use of planning horizons 2 and 3 to higher planning horizons, the demand for hospital beds is significantly lower from Monday through Friday and significantly higher during weekends. The reason for this is the following. When a short planning horizon is used, relatively a lot of surgeries are scheduled within dummy space, which we already concluded when evaluating the levels of OR utilisation of dummy space (ORB). Since most of the dummy space is located at the end of the week (see table 4), the demand of hospital beds is high near the end of each week, relative to the use of longer planning horizons.

Furthermore, looking at the standard deviations, the fluctuation in demand for hospital beds seems to decrease as we increase the planning horizon. This intuitively makes sense since we achieve higher schedule stability when increasing the planning horizon. Appendix B.6 shows the average demand for hospital beds from Monday to Sunday for planning horizons 2 and 50. Planning horizon 50 shows less deviation between the days of the week than planning horizon 2.

When measuring the relative frequencies of the demand levels we get the following distribution, obtained over a total of 100 simulations, each with once again a warm-up period of 800 and a run length of 2000. The frequencies show a normal distribution, for which the first and second moment can be estimated using the average and standard deviation for the appropriate planning horizon, obtained from the simulation. To verify the normality of these distributions appendix B.7 shows QQ-plots of the demand levels for hospital beds for planning horizon 2. The distributions for higher planning horizons have a comparable shape, but have expected values end variances which correspond to the values reported in table 5.

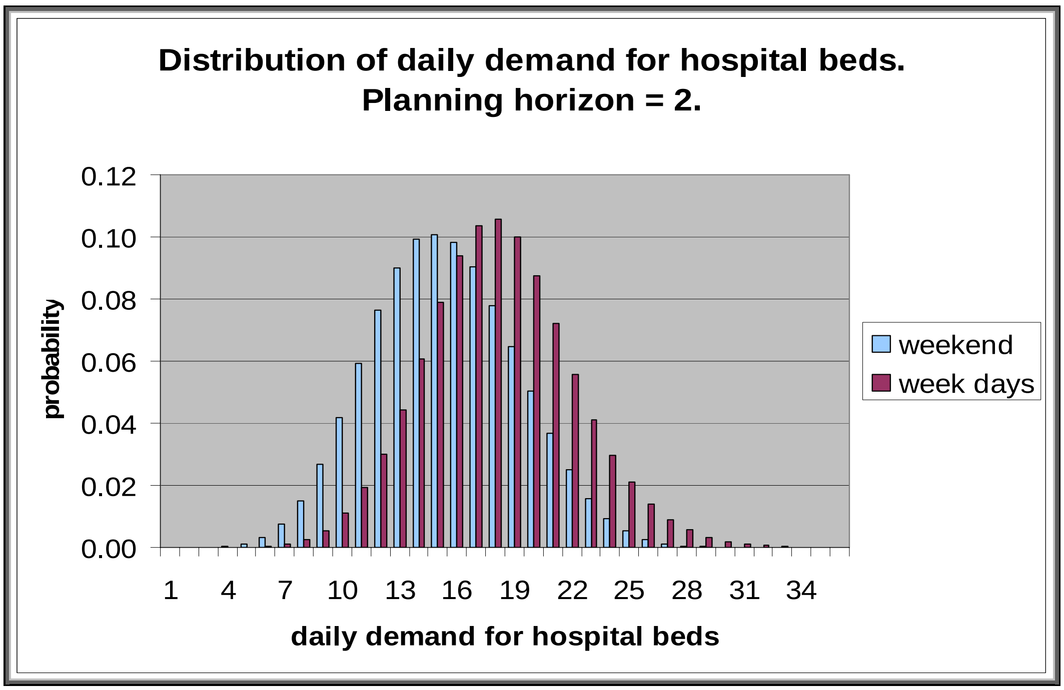


Figure 7: distribution of daily demand for hospital beds. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. Planning horizon: 2 weeks.

6.2 Effect of other input decisions

Now that we've evaluated the effect of the planning horizon on the regular scheduling process, we will look at the effects of adjustments to this scheduling process. That is, we will look at the effect of the input parameters mentioned in section 5.4. Because no information is given about the distribution of the arrivals, we will first look at the influence of different arrival distributions. The next three input parameters involve decisions that have to be made regarding the way of scheduling. We will look at the decision when to include capacity of an OR-day to the available dummy space, the effect of relaxing the prescriptions regarding the exclusive dummy space and the effect of rescheduling A surgeries from dummy to A time slots.

6.2.1 Arrival process

A clear pattern in the effect of the variance of the arrival distribution is shown by the fist four rows of the second part of table 5. The constant arrival function (with zero variance) is followed by the gamma distributions with variances of 20, 80 and 160 respectively. The basic scenario has variance 40. Higher variances have negative effects on all performance measures. By negative effect we mean a less desirable effect for both patients and hospital. The average waiting time as well as the average length of the waiting list increase, which is unbeneficial for patients. A higher variance also results in less use of type A time slots, more cases of planned overtime and a higher variance in the demand level for hospital beds, all unbeneficial effects for the hospital. Note that the variance of the arrival function is not a parameter that can be adjusted in the scheduling process, like the other input parameters. The reason for evaluating the effect of different variances is merely to investigate the effect of an unknown parameter about reality.

6.2.2 Lower bound on dummy space

The dummy capacity was defined as the time left after taking the capacity for type A and C surgeries into account, but by restricting this capacity by means of a lower bound. After applying the lower bound, 50 percent of the capacity for type C surgeries was added to the dummy capacity, allowing the possibility of scheduling with overtime. Considering the dummy capacity, the scenario which prescribes a lower bound of 40 minutes excludes only one OR-day more then the basic scenario, which prescribes a lower bound of 20 minutes. The result show no significant difference between both scenarios, which indicates that hardly any surgeries were scheduled in OR-day (3,2) in the basic scenario to begin with.

6.2.3 Distinction between dummies

Disregarding the distinction between urology and general surgery dummies results in slightly improved results for the average waiting time and waiting list. Also, the amount of times planned overtime occurs, decreases slightly. When the distinction between the two types of dummy capacity is a highly valued factor of schedule stability for the hospital, the positive effects of disregarding this distinction are probably not beneficial enough to be seriously considered.

6.2.4 Shift 'late' A surgeries from dummy to A space

As for the possibility of rescheduling type A surgeries from dummy to A space in the MSS, we have considered all possible replanning horizons in case a planning horizon of five weeks in used. That is, we have considered the type A surgeries scheduled within dummy slots of the last one to four weeks of the planning horizon for rescheduling. Of course, the longer the replanning horizon is chosen, the higher the number of surgeries considered for rescheduling. This results in an increase in both the percentage of A surgeries scheduled in A slots and the utilisation of the A capacity. As utilisation of A slots increases, utilisation of B slots decreases because of the shifts made from B to A space in the MSS. This effect is beneficial to the hospital since it increases schedule stability. For the patients on the other hand, the results show a slight increase in the average waiting time as the replanning horizon goes up. Looking at the process of rescheduling we know that the maximum increase in individual waiting time is at most the size of the replanning horizon, since surgeries before the replanning horizon will not be rescheduled.

Comparing the length of the waiting list when rescheduling to the length of the waiting list of the basic scenario, the waiting list is shorter on average when a replanning horizon of 1 is used and longer on average when a higher replanning horizon is used. This is due to two different aspects of rescheduling. On the one hand, rescheduling a surgery from a dummy to a type A slot provides an newly available dummy. This dummy can be occupied by several different surgery types, while the type A slot to which the surgery was moved was exclusively available for one specific surgery type. A surgery that could have ended up on the waiting list, can now possibly be scheduled within the newly available dummy. This decreases the average length of the waiting list. On the other hand, the possibility of the dummy remaining unoccupied exists, because the horizon eventually rolls on, not containing the dummy location anymore. The closer the newly available dummy was situated near the beginning of the planning horizon, the more likely this gets. Consider a type A surgery of the same type as the one that was rescheduled earlier. When such a surgery arrives, it encounters one less available time slot of its own type. If the dummy location, from which the other surgery was rescheduled, is not available within the planning horizon anymore, the possibility of this surgery ending up on the waiting list exists. When this happens, it increases the average length of the waiting list.

Besides rescheduling at a planning horizon of five weeks, we have added one extra scenario of the use of a planning horizon of two weeks for which one week is the only possible replanning horizon. This results in a relatively high gain in schedule stability. The process of rescheduling is depictured below for surgery type a_{42}, which has its time slots in the MSS on OR-days 5 and 10.

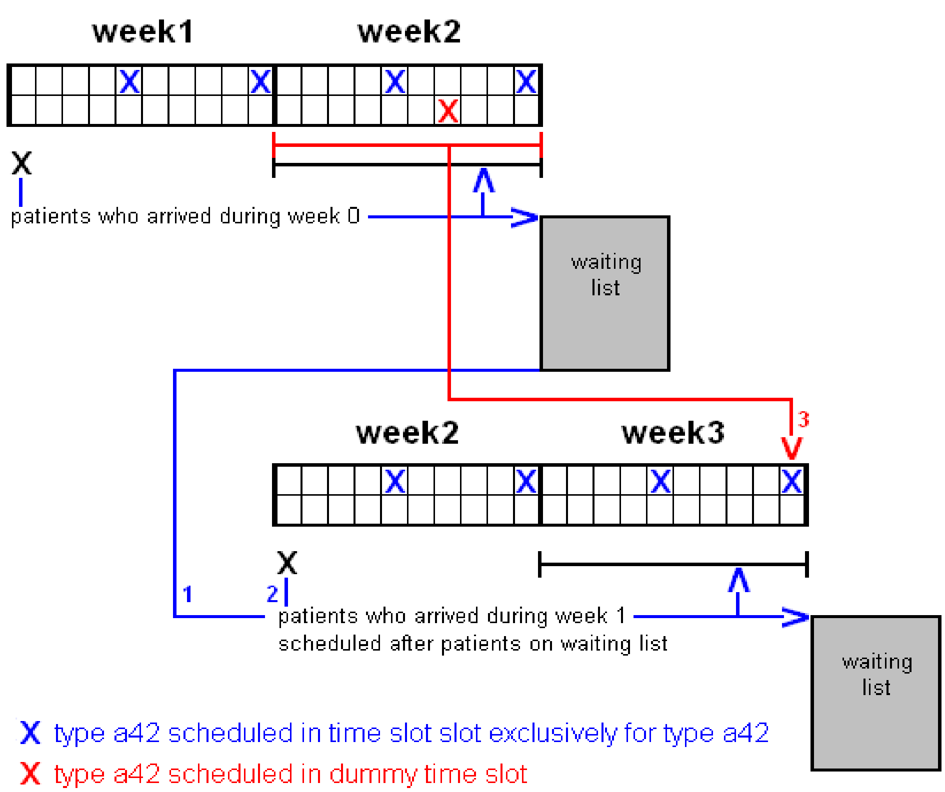


Figure 8: Rescheduling a_{42} from dummy to A slot. Planning horizon: 2. Replanning horizon: 1.

In a scenario without rescheduling, one or two time slots for type a_{42} in week 3 will remain empty in case the number of patients of type a_{42} arriving at week 1 (and the ones on the waiting list) is less than two. This gives loss of utilisation of type A time slots, since patients arriving at week 2 will be scheduled starting from week 4 and the slots of week 3 would remain empty.

Now, if rescheduling of type A surgeries from dummy to A slots is allowed, the type A time slots which would otherwise have remained empty, will be occupied if week 2 contains type a_{42} surgeries initially scheduled in dummy space. The results show quite an improvement in the percentage of type A surgeries scheduled in type A slots, which raises from 83.95 percent to 86.84 percent. The OR utilisation of type A time slots also raises about three percent, only with an average increase in waiting time of 0.028 week (rounded to three decimal accuracy). This increase of 0.028 indicates that on average 2.8 patients out of 100 patients are rescheduled from dummy to A space. This reasoning only applies when a planning horizon of two weeks is used, because in that case the increase in waiting time always is one week per patient when rescheduling. The data shows that on average about 95 percent of all patients are of type A (see table 1). So, rescheduling at a planning horizon of two weeks results in an increase in the percentage of A patients scheduled in A slots ($2.8 / 95 \cdot 100$) about 2.9 percent. This explains the relation between the increase in average waiting time and the increase in percentage of type A surgeries scheduled in type A time slots, for planning horizon 2 and replanning horizon 1. The same kind of relation can be found between waiting time and the OR utilisation of type A. Since the MSS contains 37 type A time slots out of a total of 40 time slots (which is 92.5 percent), we have an increase of ($2.8 / 92.5 \cdot 100$) about 3.0 percent. These calculations are meant to give an impression about the relationship between these three performance measures for this specific case of rescheduling.

The gain in schedule stability is relatively high when rescheduling is implemented for a planning horizon of two weeks and a replanning horizon of one week. There are two reasons for this relatively high gain:

1. If rescheduling takes place for a planning horizon of two weeks, the surgery is always shifted to the upcoming week. Using a longer planning horizon, this is not the case (see for example figure 3 in section 5.2). Rescheduling to the upcoming week will occupy type A slots that would otherwise have remained empty.
2. Secondly, without rescheduling, the only patients available for occupying type A slots of week $i+2$ are the ones arriving at week i and possibly a few patients from the waiting list (see figure 8). This means that loss of utilisation of A slots is more likely to occur for a planning horizon of two weeks than for a higher planning horizon. Thus, relatively there is more opportunity for gain of utilisation of A slots when rescheduling at a planning horizon of two weeks, than for example when rescheduling at five weeks. The use of a planning horizon of five week already has a high utilisation of type A slots to begin with. The results show a utilisation of type A slots for planning horizon five without rescheduling of already 94.23 percent, as opposed to 'only' 83.95 percent without rescheduling for planning horizon two.

6.3 Time series of bed occupation levels

We would like to investigate the demand pattern for hospital beds in more detail. Section 6.1.6 already showed that increasing the planning horizon, decreases the standard deviation of the demand levels. Appendix B.6 shows that the demand levels of the days within the cycle start to deviate less from one another as the planning horizon is chosen higher.

When we look at the outcomes of the simulation for the demand levels for hospital beds of consecutive days, we are in fact looking at a time series. We will investigate the properties of two of such time series: one resulting from a planning horizon of two weeks and the other resulting from a planning horizon of 50 weeks. Figure 9 shows two of such time series, containing the first 200 values after the warm-up period.

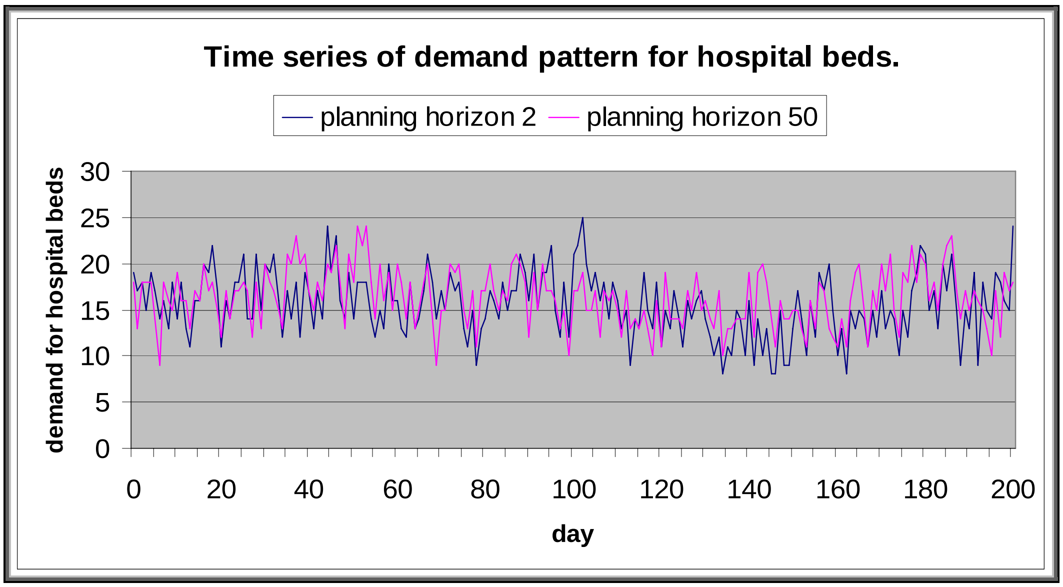


Figure 9: time series of the first 200 demand levels after a warm-up period of 800 weeks

A first step in identifying a time series model is investigating the (partial) autocorrelation function. Appendix B.8 shows these functions by means of a correlogram of 14000 demand levels resulting from simulation runs of both planning horizons 2 and 50. The empirical autocorrelation functions of both time series show a typical pattern. We find large positive values at lags 7, 14, 21 etc. This shows that bed occupation levels of the same day within different cycles are closely related, which is an indication of a seasonal pattern within the time series. This is not surprising since the occupation of hospital beds of a certain day within the cycle depends for a great part on the surgeries that are scheduled for that day by means of the MSS. The different values between the average daily levels within the cycle in appendix B.6 support this idea of seasonality. If the occupation of hospital beds would not depend on the day within the cycle, the daily levels in appendix B.6 would be closer to one another. Those average daily levels can be defined as μ_s, where $s \in \{1,2,\ldots,7\}$ denotes the day of the week, starting at $s=1$ denoting Monday. Subtracting the appropriate μ_s from the series Y_t we obtain a new deseasonalised time series $Y_t^* = Y_t - \mu_s$, when t corresponds to day s. Appendix B.8 also shows the correlograms of the deseasonalised time series of the demand levels for planning horizons 2 and 50.

The basic class of so-called autoregressive (AR) models takes the value of the time series at time t to be a linear function of the values at times *t-1* to *t-p*. For an AR(p) model the autocorrelation function has values which decline exponentially towards zero, while the partial autocorrelation function has values of zero for k-th order partial autocorrelations with $k > p$ [16]. Appling this to the correlograms of both deseasonalised time series, an AR(1) model seems to be appropriate. For our time series this results in the following seasonal AR(1) model, containing seven 'seasons'.

$$Y_t - \sum_{s=1}^{7} \mu_s D_{s,t} = \phi\left(Y_{t-1} - \sum_{s=1}^{7} \mu_s D_{s,t-1} \right) + \varepsilon_t \text{ , where} \tag{3}$$

- Y_t The demand level for hospital beds at day t
- μ_s The average value of the demand level at day $s \in \{1,2,\ldots,7\}$ corresponding to the values given in Appendix B.6
- $D_{s,t}$ Seasonal dummy variable: 1 if t corresponds to day s, 0 otherwise
- ε_t The error term at day t

The model represented by (3) has the unconditional mean demand level of $E(Y_t) = \mu_s$ when t corresponds to day s, which implies $E(Y_t^*) = 0$ for the deseasonalised time series. Regressing the deseasonalised series Y_t^* on Y_{t-1}^* gives us the estimate of ϕ, which represents the dependence between two consecutive days when disregarding the seasonal pattern. The results of the least squares regressions are included in appendix B.8. Rounding the estimate to two decimal accuracy gives us:

$$\hat{\phi} \approx 0.71$$

for both planning horizons. The similarity between this estimated parameter of the models of both planning horizons indicates that the speed at which a demand level returns to the average daily value is independent of the planning horizon. This can be related to the patients who are present in the wards and intensive care units at a certain point in time. Since the expected time a patient spends in the hospital for recovery depends only on its surgery type and not on the planning horizon that was used in the scheduling process, the rate at which patients leave the hospital is not influenced by the planning horizon. This means that when the hospital encounters days with excessively high demand for hospital beds, the rate at which the demand level will decrease towards the daily average is roughly the same for all planning horizons.

Furthermore, the decrease in the standard deviation of the demand pattern for hospital beds when increasing the planning horizon, is not only due to the deviation between the daily averages within one week. Judging from the estimation results of the AR(1) models in Appendix B.8, the variation between the daily levels of different weeks is larger for smaller planning horizons. This can be concluded from the standard error (S.E.) of the regression, which is larger for the AR(1) model of planning horizon 2 than for the AR(1) model of planning horizon 50.

Using the model (3) figure 10 shows the fitted deseasonalised time series along with the actual deseasonalised series for the first 500 days after a warm-up period of 800 weeks. A similar figure is shown in Appendix B.8 for planning horizon 50.

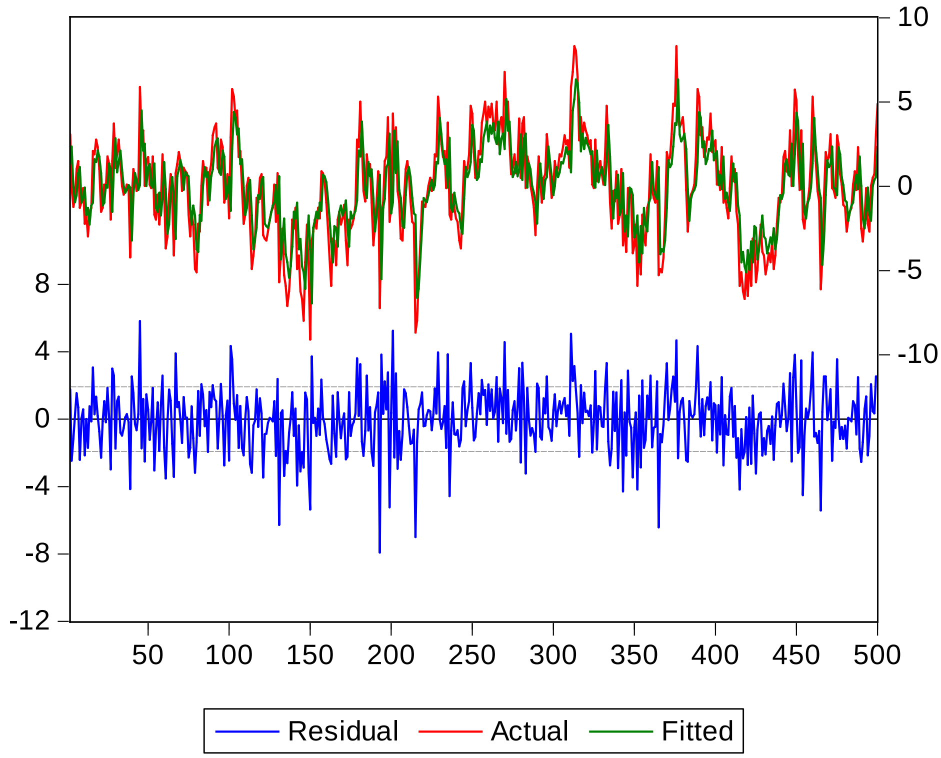


Figure 10: Actual, fitted and residual time series of the deseasonalised time series Y_t* for planning horizon 2, shown for the first 500 observations, after a warm-up period of 800 weeks.

Suppose the hospital departments only have 20 beds available. Demands exceeding this level result in shortages which could for example be handled by temporarily using beds of other departments. This is of course not a desirable situation. A useful application of the time series we just modelled, could be to estimate the time it will take before the demand decreases below a certain desirable level. Let's consider the following example of a incidentally high demand of 30 hospital beds at $t = T$, corresponding to a Monday. Suppose the hospital uses a planning horizon of two weeks. Then, the following expected demand pattern shows that the departments encounter a shortage of beds during an estimated period of five consecutive days.

- $Y_T = 30$, where T corresponds to $s = 1$ (representing Monday)
- $E(Y_{T+1}) = \mu_2 + \phi(Y_T - \mu_1) = 13.98 + 0.71(30 - 17.34) = 22.97$
- $E(Y_{T+2}) = \mu_3 + \phi(E(Y_{T+1}) - \mu_2) = 18.15 + 0.71(22.97 - 13.98) = 24.53$
- $E(Y_{T+3}) = \mu_4 + \phi(E(Y_{T+2}) - \mu_3) = 16.34 + 0.71(24.53 - 18.15) = 20.87$
- $E(Y_{T+4}) = \mu_5 + \phi(E(Y_{T+3}) - \mu_4) = 19.82 + 0.71(20.87 - 16.34) = 23.03$
- $E(Y_{T+5}) = \mu_6 + \phi(E(Y_{T+4}) - \mu_5) = 16.19 + 0.71(23.03 - 19.82) = 18.74$

7. Discussion

We have evaluated the relation between input and output factors of the scheduling process in a Master Surgical Schedule. The main feature we encountered looking at the results, was a trade-off between the patients' waiting time and schedule stability. We did not make any judgements about the quality of a scenario. The evaluation of the performance of a scenario is a subjective issue. From a patient's point of view, short planning horizons are preferred, since these result in short waiting times. The hospital management on the other hand, benefits from a stable (hence predictable) schedule. When the management knows what kind of surgeries to expect when and where, they can use this information in order to reduce costs. For example, stable operating room schedules will lead to stable time schedules of hospital personnel. This implies less overtime within the working hours of personnel, which reduces costs for the hospital. This also brings us to a beneficial effect of schedule stability for hospital personnel: they will experience less stress because of unstable working hours. In the end, the hospital management should select a scheduling procedure which provides the best balance between their view of the importance of the patients' waiting time versus schedule stability.

Some aspects about the data of this research forced us to make assumptions. These assumptions are part of the inputs of the scheduling process and might have had their influence on the results in some way. First of all, we have assumed an expected arrival rate of 40 patients per week. Since the hospital schedules elective surgeries only during a total of 46 weeks per year, one could argue that the expected number of arrivals per week will defer during the year depending on time, because of a temporary increased amount of patients who might arrive after certain holidays. This time conditional hetroskedastisity could be modelled by using a non-homogeneous Poisson distribution. This aspect was left out of consideration and due to further research.

Also, we made a decision how to handle type B surgeries with high expected surgery times. Those surgeries have an expected duration which exceeds the maximum available dummy space in one OR-day. We decided upon scheduling these surgeries within one of the three largest dummy spaces of the cycle, if empty, still causing quite a large amount of planned overtime. The effect of this decision is clearly visible in the results. The longer the planning horizon is chosen, the less the amount of type A surgeries gets which are scheduled in dummy space. Thus, for long planning horizons, the total planned overtime is mainly the result of the type B surgeries just mentioned. Since these surgeries have high planned overtime, the average amount of planned overtime increases towards the excess duration of these type B surgeries. It might be worthwhile considering other ways of handling these surgeries in further research. For example, once in a while an extra OR-day could be assigned to such surgeries, and by doing so, avoiding planned overtime for these surgeries.

The last assumption we would like to mention here, is the use of capacity for type C surgeries to extend the dummy space with. We decided upon a maximum use of 50 percent of the type C surgeries, allowing planned overtime when scheduling surgeries in dummy space. Combined with the measure just mentioned above, the effect of not allowing any planned overtime at all could also be a subject of further research.

8. Conclusion

In this research we have investigated the effect of inputs in a Master Surgical Schedule by means of simulation. The main feature of investigation was the effect of the planning horizon. As expected, increasing the planning horizon increases schedule stability but also increases the patients' average waiting time. The schedule stability was captured by the percentage of type A surgeries scheduled in time A slots, the amount and duration of planned overtime, the OR utilisation of both type A and B slots and the stability in the demand pattern for hospital beds. A remarkable pattern was shown in the distribution of the patients waiting time. With high probabilities patients seem to experience either relatively short or relatively long waiting times. The explanation for this phenomenon can be found by comparing the expected arrival rates to the number of available time slots within the MSS per surgery type. When the expected arrival rate exceeds the number of available time slots, all time slots within the planning horizon will gradually get occupied, resulting in surgeries being scheduled either in type A slots near the end of the planning horizon or in dummies near the beginning of the planning horizon. The difference between the expected arrival rate and the number of available time slots also effects the upper bounds of the percentage of type A surgeries scheduled in type A time slots and the OR utilisation of type A time slots.

Besides the effect of the planning horizon we also investigated the effect of the unknown variance of the arrival distribution and decisions regarding the scheduling process. The higher the variance of the arrival distribution, the less desirable effects are found for both patients and hospital: waiting times increase and schedule stability decreases. When defining the dummy capacity of the MSS, raising the upper bound of free capacity before assigning it to the dummy capacity might result in higher average waiting times. This was not the case for our data, since the dummy space with low time capacity seemed to be hardly ever used to begin with. Disregarding the distinction between dummy capacities of different specialties on the other hand did have slight effect on the average waiting time and length of the waiting list, which both decreased slightly. Rescheduling type A surgeries from dummy time slots to newly available type A time slots also effects the performance of the scheduling process. Slightly increased waiting times go along with an increase in schedule stability. This effect is most profitable when a planning horizon of two weeks is used with a replanning horizon of one week. With those settings, type A slots which would have remained empty without rescheduling, are then possibly used for shifting type A patients to. This only increases the waiting time of such patients by one week.

We modelled the time pattern of the demand level for hospital beds by means of a seasonal AR(1) model. The model captures the relation between demand levels of consecutive days, taking into account the expected level of the appropriate day within the cycle.

A. References

[1] J.T. Blake, J. Donald (2002). 'Mount Sinai Hospital Uses Integer Programming to Allocate Operating Room Time'. I*nterfaces*, 32 (2):63-73

[2] J.M.H. Vissers, I.J.B.F. Adan, J.A. Bekkers (2005). Patient mix optimization in cardiothoracic surgery planning: a case study'. *IMA Journal of Management* Mathematics, 16 (3):281-304

[3] M. van Houdenveen, J.M. van Oostrum, G. Wullink, E.W. Hans, J.L Hurink, J. Bakker, G. Kazemier (2007). 'Fewer intensive care units refusals and higher capacity utilisation by using a cyclical surgical case schedule'. *Journal of Critical Care*, 23 (2): 222-226

[4] J. Belien, E. Demeulemeester (2007). 'Building cyclic master surgical schedules with levelled resulting bed occupancy'. *European Journal of Operations Research*, 176 (2):1185-1204

[5] J.M. van Oostrum, M van Houdenhove, J.L. Hurink, E.W. Hans, G. Wullink, G. Kazemier (2008). 'A master surgical scheduling approach for cyclic scheduling in operation room departments'. *OR spectrum*, 30 (2):355-374

[6] J.M. van Oostrum, E.Bredenhoff, E.W. Hans (2007). 'Managerial implications and suitability of a Master Surgical Scheduling approach'. Submitted to *AOR-ORAHS'2007 special volume*

[7] V. Shidharan, W.L. Berry, V. Udayabhanu (1987). "Freezing the master production schedule under rolling planning horizons'. *Management Science*, 33(9): 1137-1149

[8] V. Shidharan, W.L. Berry, V. Udayabhanu (1988). 'Measuring master production schedule stability under rolling planning horizons'. *Decision Sciences*, 19(1):147-166

[9] G.M. Campbell, V.A. Mabert (1991). 'Cyclical schedules for capacitated lot sizing with dynamic demands'. *Management Science*, 37(4): 409-427

[10] N.P. Lin, L. Krajewski (1992). 'A model for Master Production Scheduling in Uncertain Environments'. *Decision Sciences,* 23(4): 839-861

[11] K. K. Yang, F. Robert Jacobs (1999). 'Replanning the master production schedule for a capacity-constrained Job Shop'. *Decision Sciences*, 30(3): 719-748

[12] O. Tang, R.W. Grubbstrom (2002). 'Planning and replanning the master production schedule under demand uncertainty'. *Elsevier*, 78(3): 323-334

[13] J.A. Hill, W.L. Berry, D.A. Schilling (2003). 'Revising the master production schedule in sequence dependent processes'. *International Journal of Production Research*, 41(9): 2021-2035

[14] S. Creemers, M. Lambrecht (2007). 'Queuing models for appointment-driven systems'. Submitted to *AOR-ORAHS'2007 special volume*

[15] S. Creemers, M. Lambrecht (2007). 'A queuing model to analyse appointment-driven service systems'.

[16] P.H. Franses, D. van Dijk (2008), Time Series Models for Business and Economic Forecasting, Second Edition. Prepared for *Cambridge University Press*

[17] D.D. Wackerly, W. Mendenhall, R.L. Scheaffer (1996), Mathematical statistics with applications, Fifth Edition. ISBN 0-534-20916-5

B. Appendix

B.1 Gamma and Poisson distributions

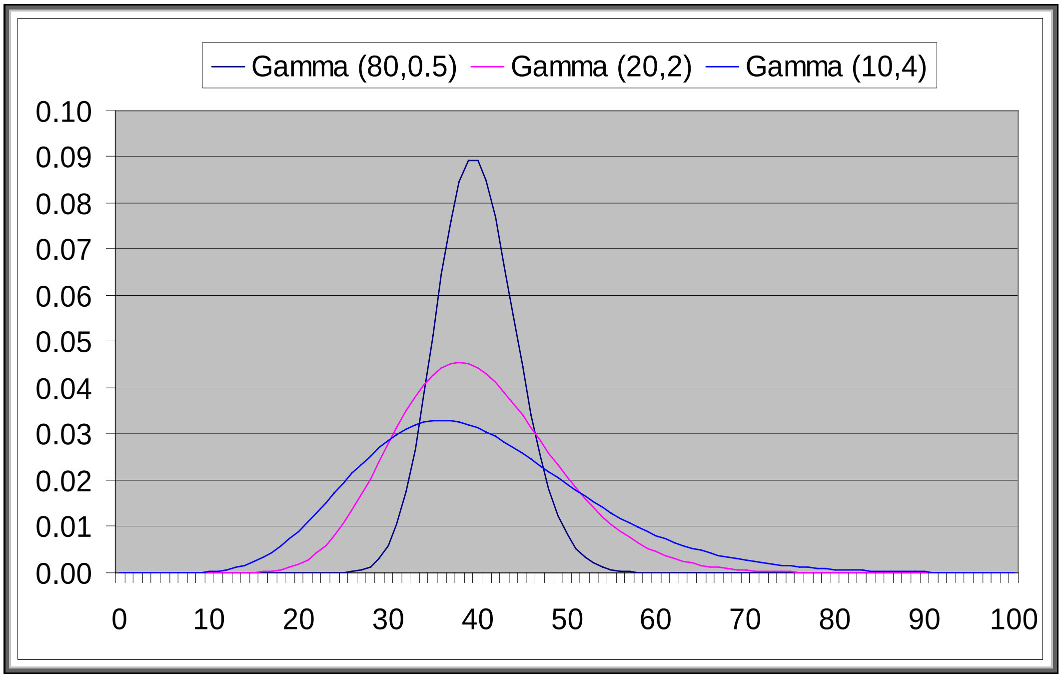


Figure 11: probability density functions of Gamma distribution with E(X) = 40 for all three and Var(X) = 20, Var(X) = 80 and Var(X) = 160 respectively.

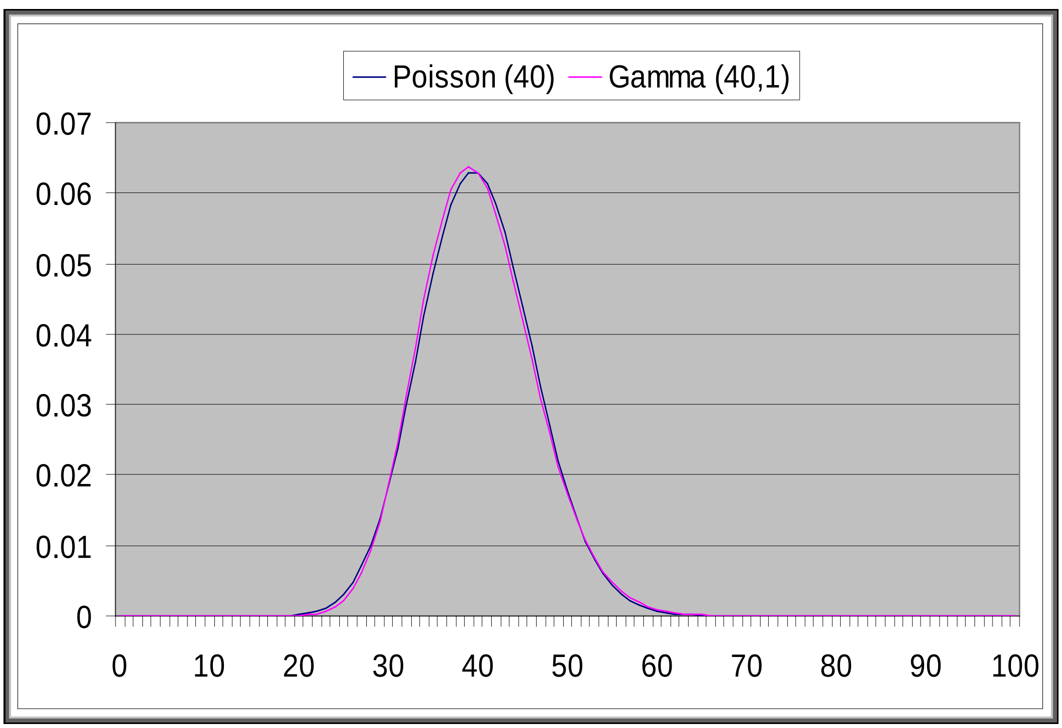


Figure 12: probability density functions of Gamma distribution and a Poisson distribution, both with E(X) = Var(X) = 40.

B.2 Time series in relation to warm-up period

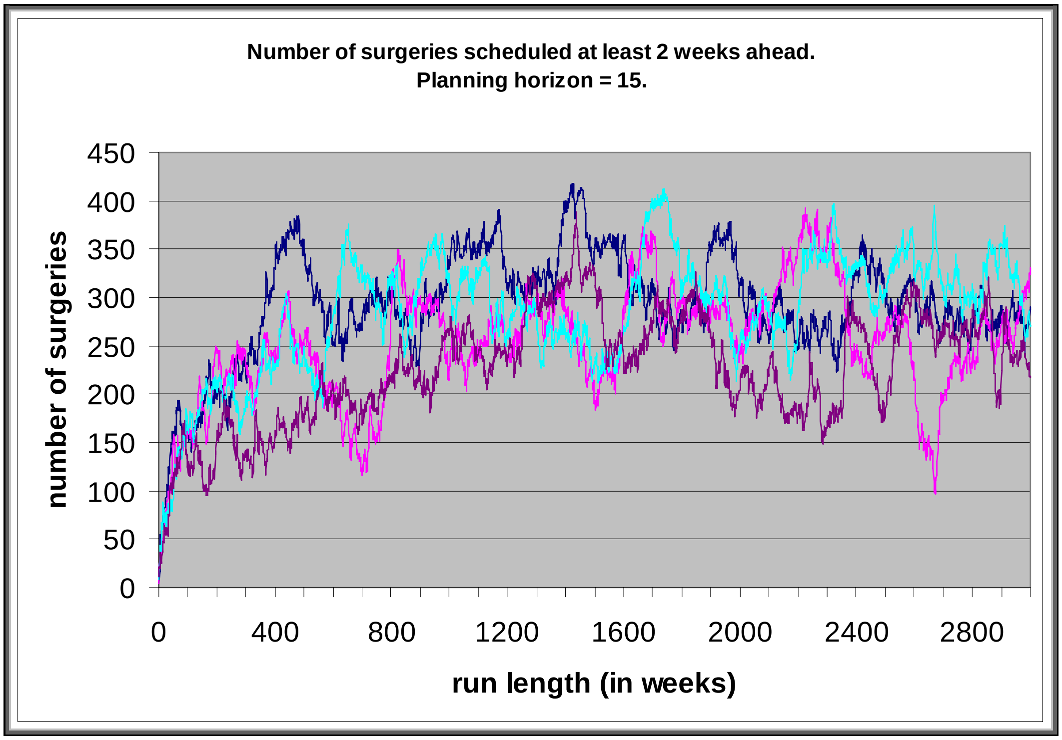


Figure 13: times series of the number of surgeries scheduled at least two weeks ahead of four independent simulations with a planning horizon of 15 weeks.

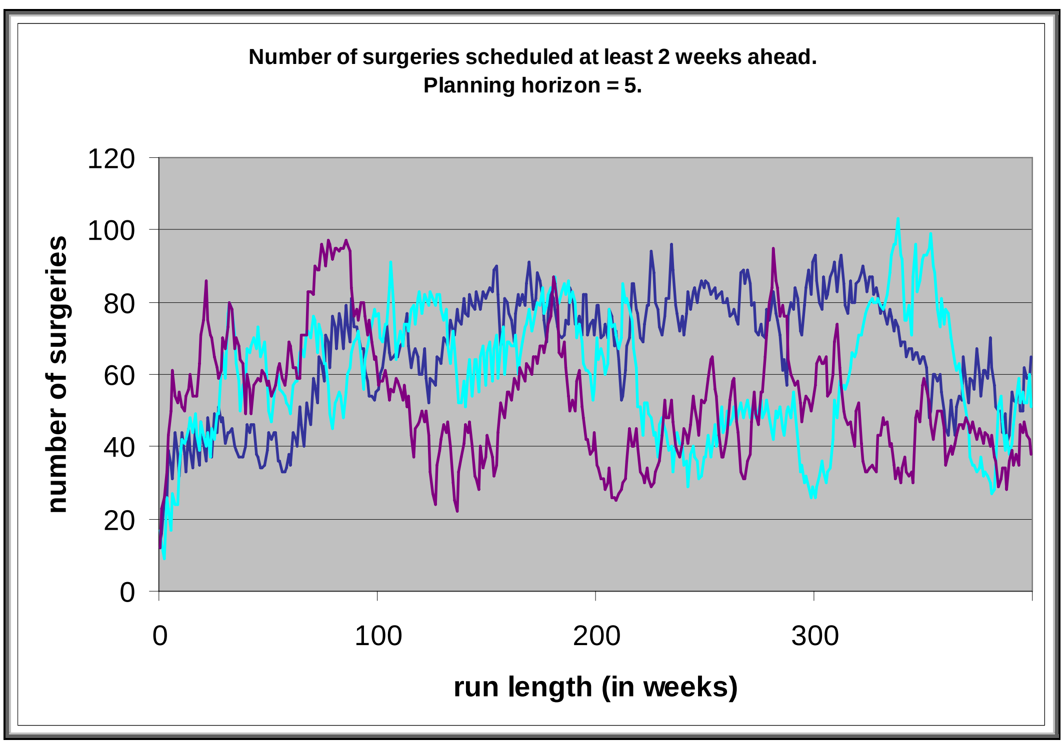


Figure 14: times series of the number of surgeries scheduled at least two weeks ahead of three independent simulations with a planning horizon of 5 weeks.

B.3 95% confidence intervals of the performance measures

input variables					performance measures																							
AP	LB	DD	RH	PH	WT		WL		AA		NPO		DPO		OR		ORA		ORB		DB		(sd)		DBwe		(sd)	
P(40)	20	Y	0	2	2.23	2.24	9.33	9.55	83.92	83.98	4.43	4.46	48.04	48.24	71.07	71.18	86.27	86.39	86.62	86.85	17.11	17.16	3.57	3.58	14.33	14.38	3.36	3.38
P(40)	20	Y	0	3	2.57	2.58	1.15	1.21	90.39	90.46	2.44	2.47	58.54	58.97	71.07	71.18	92.66	92.75	71.46	71.82	17.13	17.18	3.22	3.23	14.27	14.32	3.16	3.18
P(40)	20	Y	0	4	3.04	3.05	0.33	0.35	92.97	93.05	1.61	1.63	79.39	80.03	71.07	71.18	95.59	95.66	64.54	64.91	17.16	17.20	3.05	3.06	14.21	14.26	3.02	3.04
P(40)	20	Y	0	5	3.51	3.53	0.12	0.13	94.19	94.27	1.40	1.41	87.66	88.33	71.07	71.18	97.01	97.09	61.14	61.53	17.17	17.21	2.97	2.98	14.19	14.24	2.95	2.97
P(40)	20	Y	0	6	3.98	4.01	0.03	0.04	94.87	94.95	1.33	1.35	90.48	91.17	71.07	71.18	97.81	97.89	59.25	59.64	17.17	17.22	2.93	2.94	14.17	14.22	2.91	2.93
P(40)	20	Y	0	7	4.44	4.49	0.01	0.01	95.29	95.37	1.31	1.32	91.93	92.61	71.07	71.18	98.30	98.37	58.10	58.49	17.18	17.22	2.91	2.92	14.17	14.21	2.89	2.90
P(40)	20	Y	0	8	4.91	4.98	0.00	0.00	95.57	95.65	1.30	1.31	92.56	93.24	71.07	71.18	98.62	98.68	57.35	57.75	17.18	17.22	2.89	2.91	14.16	14.21	2.88	2.89
P(40)	20	Y	0	9	5.37	5.46	0.00	0.00	95.76	95.84	1.29	1.31	92.84	93.53	71.07	71.18	98.84	98.90	56.83	57.23	17.18	17.22	2.88	2.90	14.16	14.21	2.87	2.88
P(40)	20	Y	0	10	5.82	5.93	0.00	0.00	95.91	95.99	1.29	1.31	92.97	93.66	71.07	71.18	98.99	99.06	56.46	56.86	17.18	17.22	2.87	2.89	14.16	14.20	2.86	2.88
P(40)	20	Y	0	11	6.26	6.41	0.00	0.00	96.02	96.10	1.29	1.30	93.05	93.75	71.07	71.18	99.11	99.18	56.17	56.58	17.18	17.23	2.87	2.88	14.15	14.20	2.86	2.87
P(40)	20	Y	0	12	6.70	6.87	0.00	0.00	96.11	96.19	1.29	1.30	93.09	93.79	71.07	71.18	99.20	99.27	55.96	56.36	17.18	17.23	2.87	2.88	14.15	14.20	2.85	2.87
P(40)	20	Y	0	13	7.13	7.34	0.00	0.00	96.19	96.26	1.29	1.30	93.12	93.82	71.07	71.18	99.28	99.34	55.78	56.19	17.18	17.23	2.86	2.87	14.15	14.20	2.85	2.86
P(40)	20	Y	0	14	7.56	7.79	0.00	0.00	96.25	96.32	1.29	1.30	93.14	93.84	71.07	71.18	99.33	99.40	55.65	56.06	17.18	17.23	2.86	2.87	14.15	14.20	2.84	2.86
P(40)	20	Y	0	15	7.97	8.24	0.00	0.00	96.30	96.37	1.29	1.30	93.15	93.85	71.07	71.18	99.38	99.45	55.53	55.94	17.18	17.23	2.85	2.87	14.15	14.20	2.84	2.85
P(40)	20	Y	0	25	11.32	11.86	0.00	0.00	96.55	96.62	1.29	1.30	93.17	93.87	71.05	71.18	99.57	99.64	54.99	55.40	17.18	17.22	2.84	2.85	14.15	14.20	2.83	2.84
P(40)	20	Y	0	50	18.40	19.53	0.00	0.00	96.66	96.74	1.28	1.30	93.54	94.30	71.05	71.18	99.67	99.77	54.73	55.20	17.16	17.22	2.83	2.85	14.13	14.19	2.82	2.83
C(40)	20	Y	0	5	3.50	3.51	0.05	0.06	94.47	94.52	1.36	1.37	89.15	89.86	71.13	71.19	97.35	97.40	60.60	60.86	17.20	17.24	2.90	2.91	14.19	14.23	2.89	2.90
G(80,0.5)	20	Y	0	5	3.50	3.52	0.09	0.10	94.30	94.36	1.38	1.39	88.50	89.18	71.13	71.22	97.18	97.24	61.01	61.34	17.19	17.23	2.94	2.95	14.20	14.24	2.93	2.94
G(20,2)	20	Y	0	5	3.53	3.55	0.23	0.27	93.87	93.97	1.47	1.49	84.75	85.54	71.07	71.23	96.74	96.84	61.83	62.29	17.17	17.22	3.05	3.07	14.22	14.26	3.01	3.03
G(10,4)	20	Y	0	5	3.57	3.59	0.61	0.72	93.30	93.44	1.65	1.68	79.17	80.11	71.09	71.27	96.24	96.35	63.08	63.64	17.16	17.22	3.21	3.23	14.23	14.29	3.13	3.15
P(40)	40	Y	0	5	3.51	3.53	0.12	0.13	94.19	94.27	1.40	1.41	87.65	88.32	71.07	71.18	97.01	97.09	61.14	61.53	17.17	17.21	2.97	2.98	14.19	14.24	2.95	2.97
P(40)	20	N	0	5	3.48	3.50	0.06	0.07	94.19	94.27	1.37	1.39	88.11	88.76	71.07	71.18	97.01	97.09	61.14	61.53	17.17	17.21	2.98	2.99	14.19	14.24	2.96	2.97
P(40)	20	Y	1	5	3.56	3.58	0.12	0.13	94.38	94.47	1.30	1.32	88.22	88.98	71.07	71.18	97.32	97.39	60.42	60.82	17.15	17.20	3.00	3.02	14.23	14.27	2.99	3.01
P(40)	20	Y	2	5	3.57	3.59	0.12	0.14	94.37	94.45	1.35	1.36	87.46	88.16	71.07	71.18	97.33	97.40	60.39	60.78	17.16	17.21	3.00	3.01	14.21	14.25	2.98	3.00
P(40)	20	Y	3	5	3.63	3.65	0.18	0.19	94.45	94.53	1.39	1.40	86.85	87.58	71.07	71.18	97.43	97.50	60.16	60.55	17.16	17.20	2.97	2.98	14.21	14.26	2.95	2.97
P(40)	20	Y	4	5	3.72	3.75	0.23	0.25	94.62	94.70	1.34	1.36	86.70	87.51	71.07	71.18	97.61	97.68	59.74	60.13	17.16	17.20	2.94	2.96	14.21	14.26	2.93	2.94
P(40)	20	Y	1	2	2.26	2.27	9.33	9.55	86.81	86.87	3.93	3.96	49.93	50.17	71.07	71.18	89.30	89.39	79.44	79.75	17.13	17.18	3.42	3.43	14.28	14.33	3.27	3.29

Table 6: 95% intervals. 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks.

B.4 Distribution of waiting time for all surgery types

(for planning horizons 2 to 10 and 15)

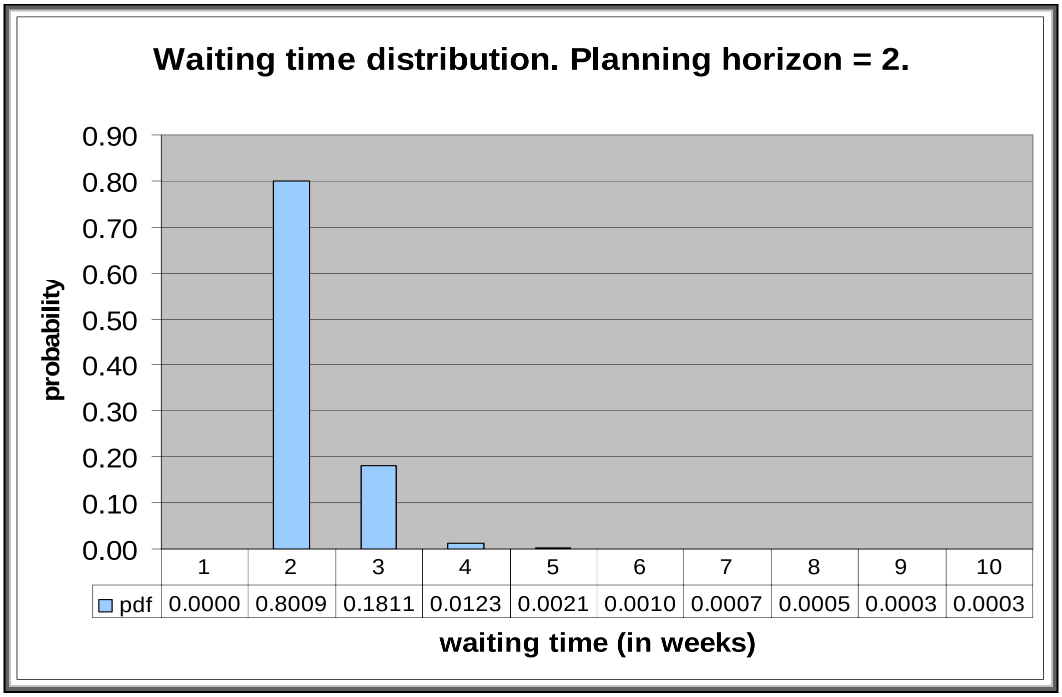


Figure 15: Waiting time distribution for all types of surgery. Planning horizon: 2 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. 0 < P(waiting time = t, 10 < t ≤ 40) < 0.0003. P(waiting time = t, t >40) = 0.

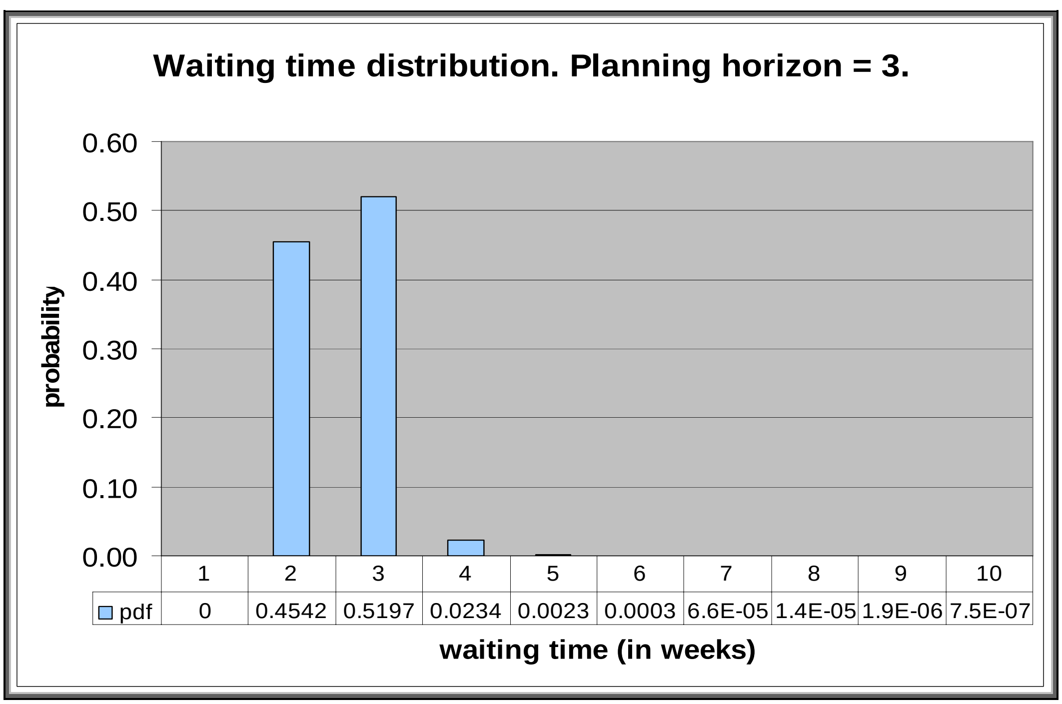


Figure 16: Waiting time distribution for all types of surgery. Planning horizon: 3 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. 0 < P(waiting time = t, 10 < t ≤ 12) < $1.4 \cdot 10^{-7}$. P(waiting time = t, t >12) = 0.

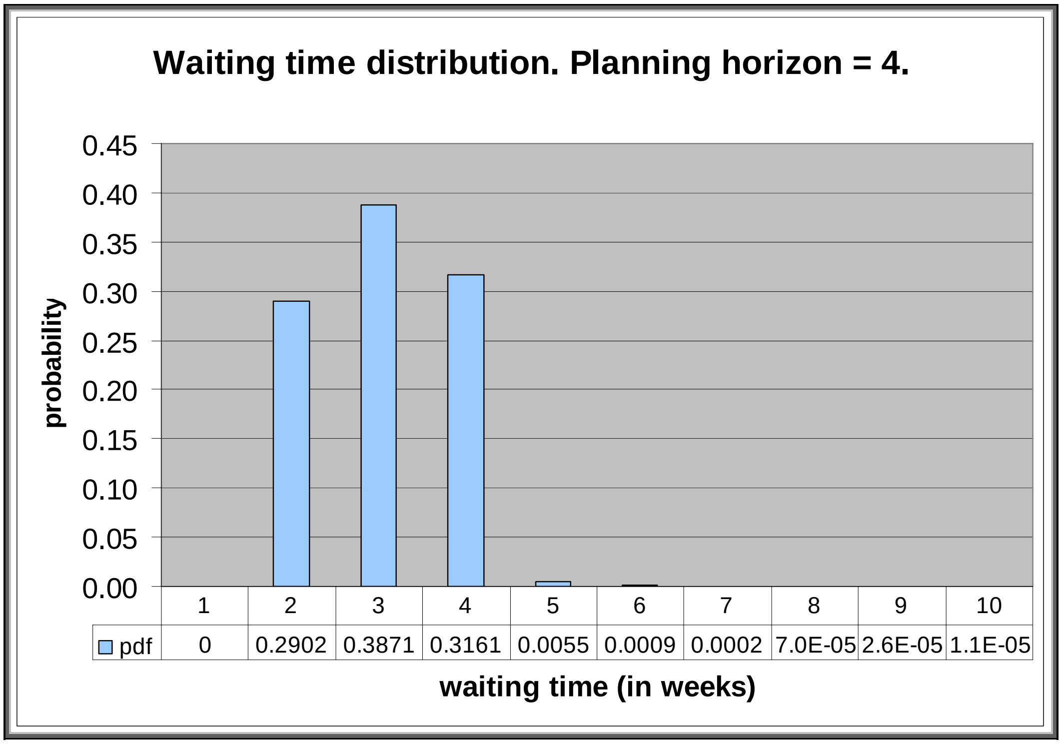


Figure 17: Waiting time distribution for all types of surgery. Planning horizon: 4 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 13) < 4.9 \cdot 10^{-6}$. $P(\text{waiting time} = t, t > 13) = 0$.

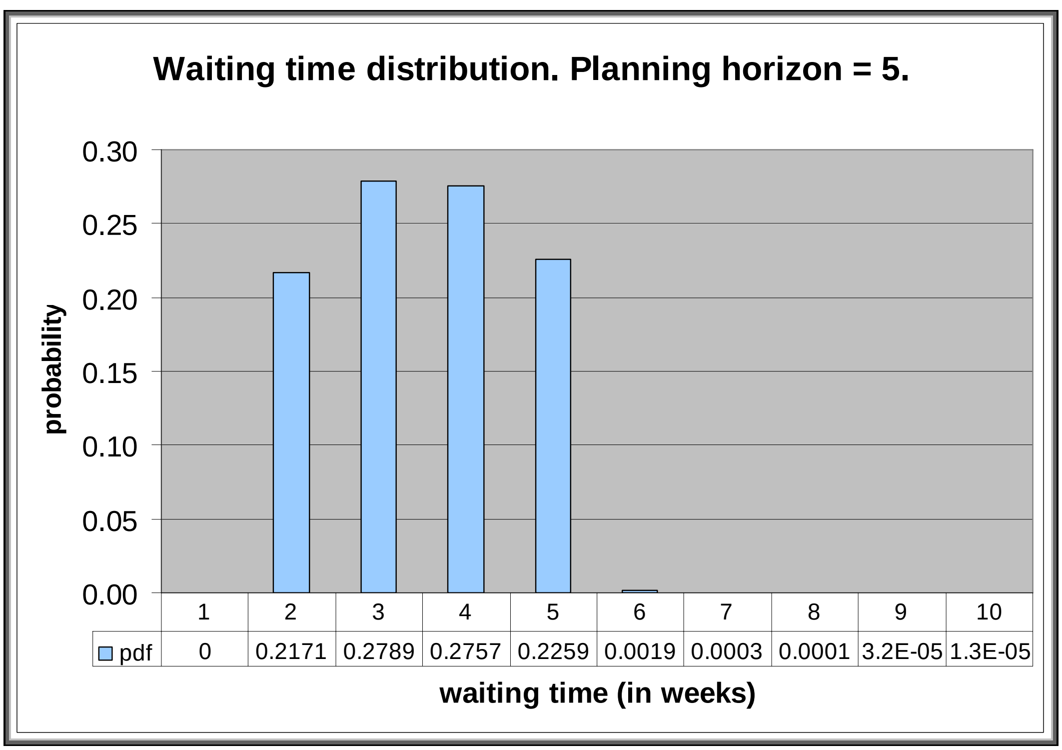


Figure 18: Waiting time distribution for all types of surgery. Planning horizon: 5 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 12) < 6.4 \cdot 10^{-6}$. $P(\text{waiting time} = t, t > 12) = 0$.

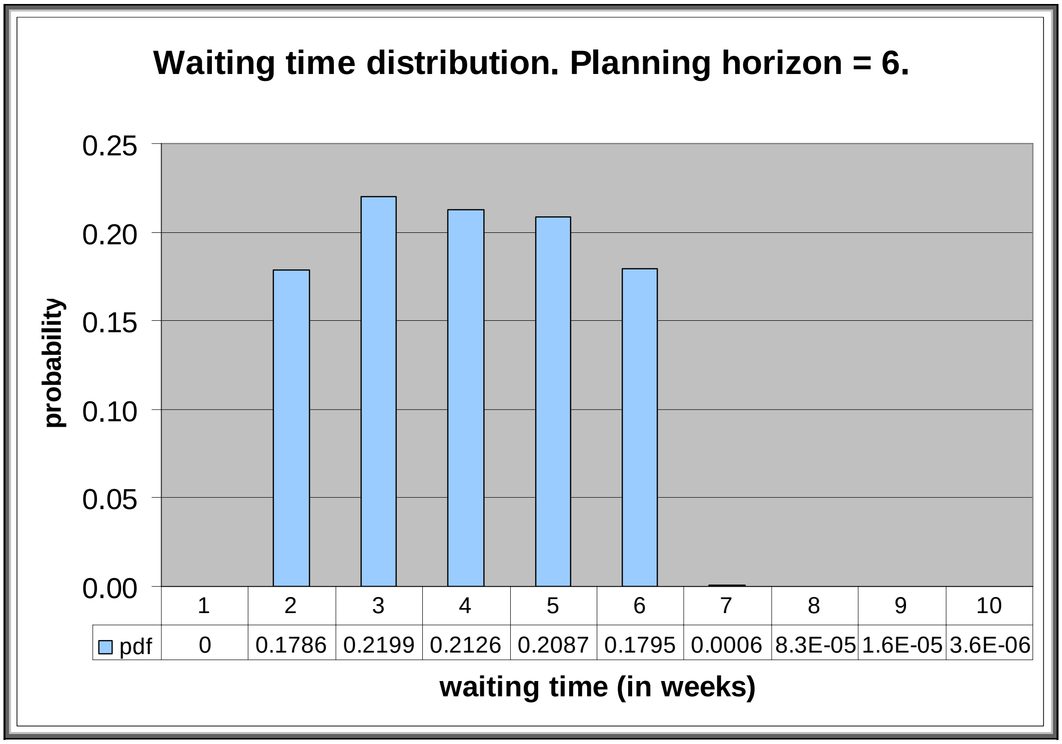


Figure 19: Waiting time distribution for all types of surgery. Planning horizon: 6 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 12) < 6.4 \cdot 10^{-7}$. $P(\text{waiting time} = t, t > 1) = 0$.

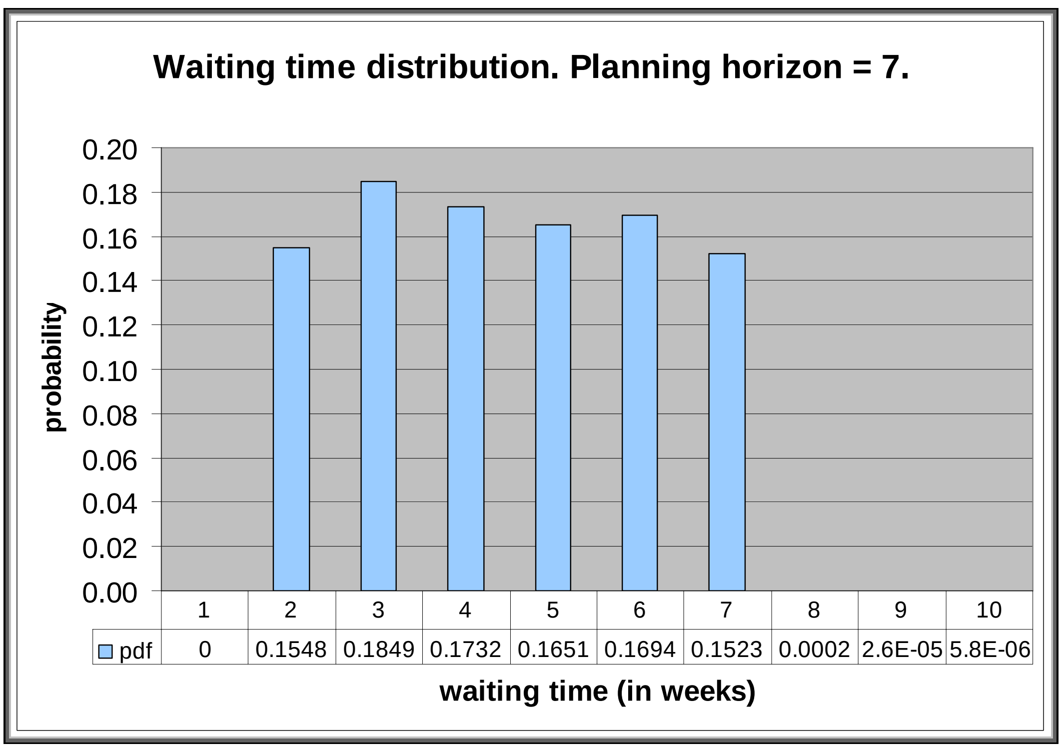


Figure 20: Waiting time distribution for all types of surgery. Planning horizon: 7 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 13) < 1.2 \cdot 10^{-6}$. $P(\text{waiting time} = t, t > 13) = 0$.

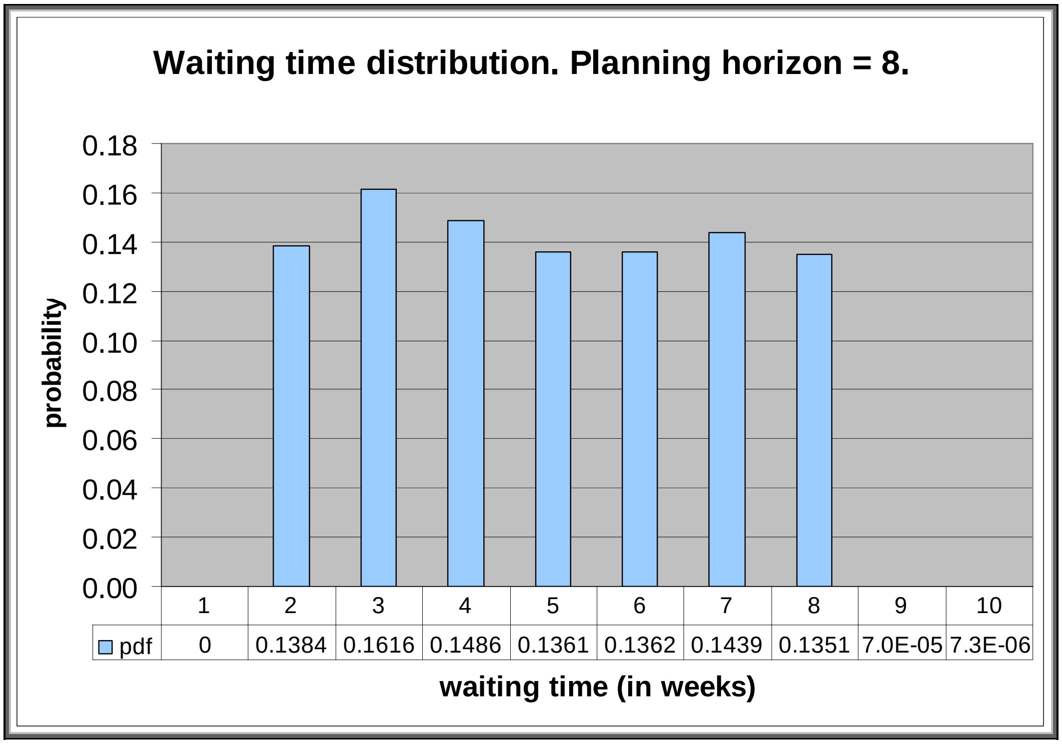


Figure 21: Waiting time distribution for all types of surgery. Planning horizon: 8 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 13) < 1.2 \cdot 10^{-6}$. $P(\text{waiting time} = t, t > 13) = 0$.

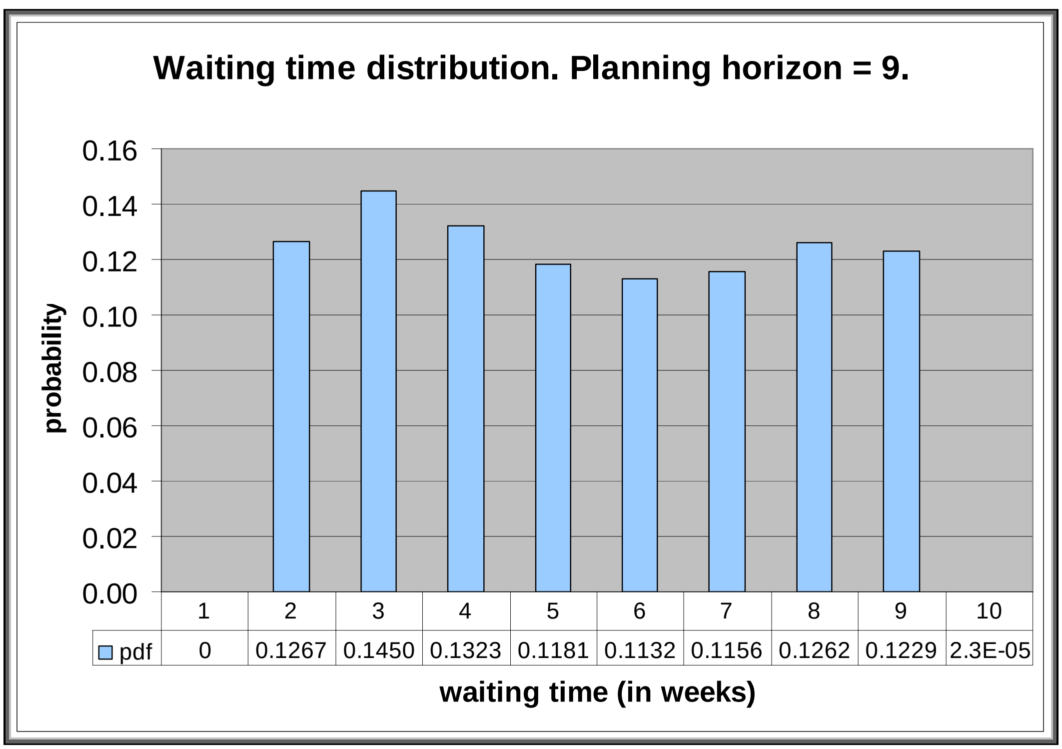
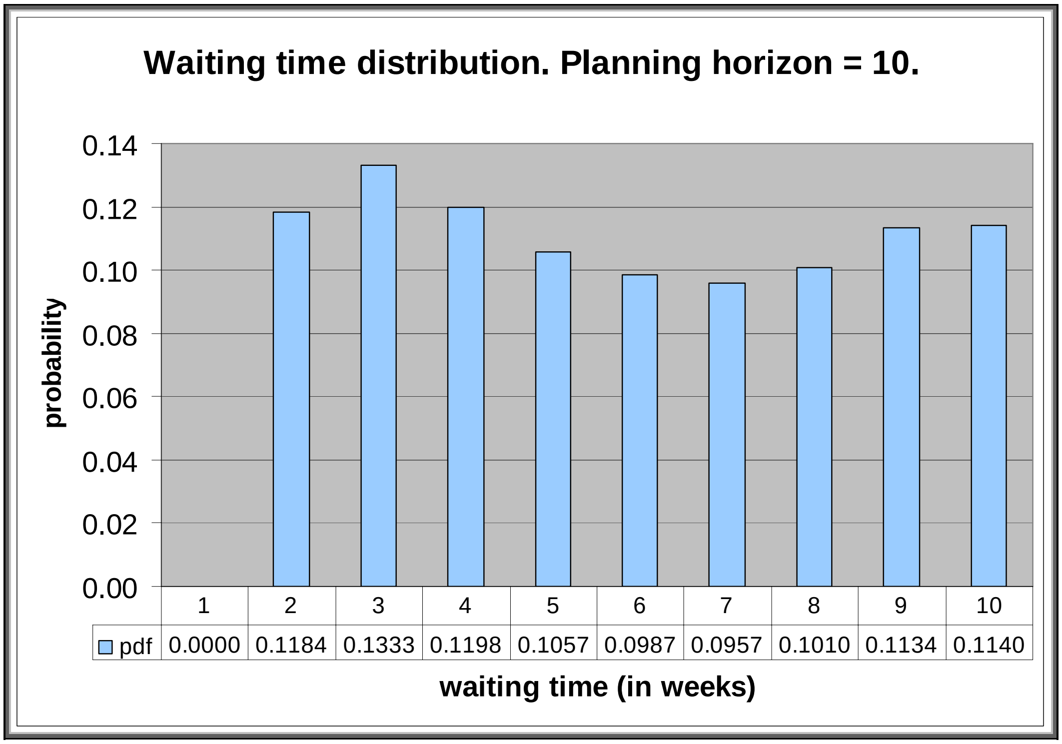
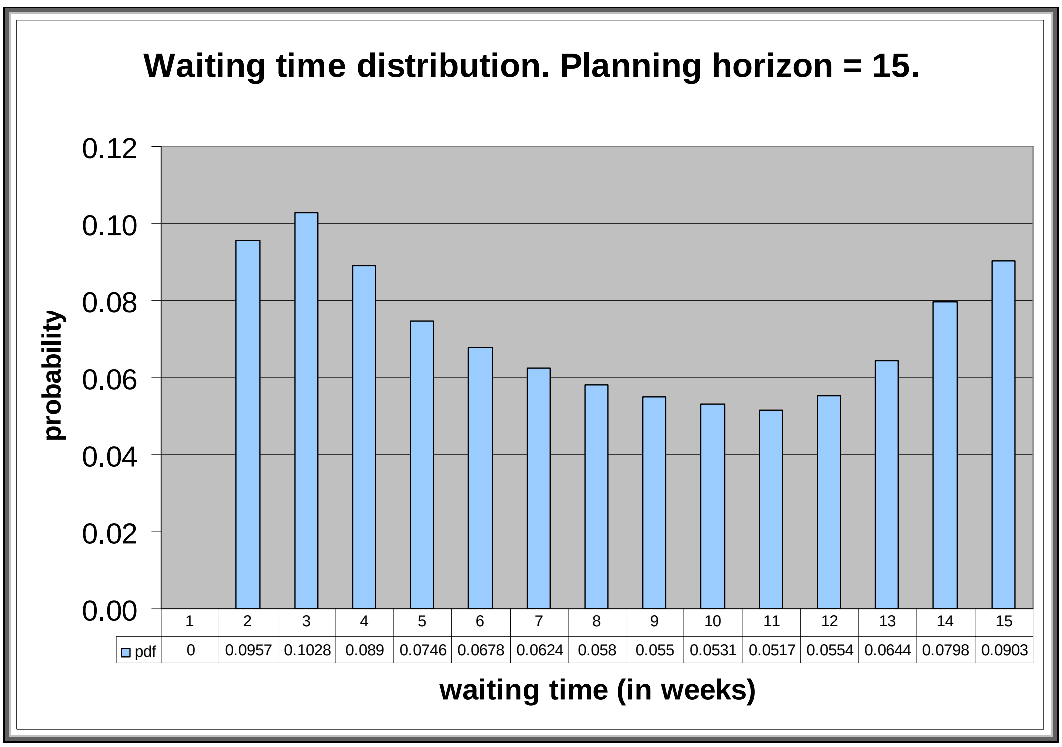


Figure 22: Waiting time distribution for all types of surgery. Planning horizon: 9 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 13) < 2.0 \cdot 10^{-6}$. $P(\text{waiting time} = t, t > 13) = 0$.



	1	2	3	4	5	6	7	8	9	10
pdf	0.0000	0.1184	0.1333	0.1198	0.1057	0.0987	0.0957	0.1010	0.1134	0.1140

Figure 23: Waiting time distribution for all types of surgery. Planning horizon: 10 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. $0 < P(\text{waiting time} = t, 10 < t \leq 13) < 6.4 \cdot 10^{-6}$. P(waiting time = t, t >13) = 0.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
pdf	0	0.0957	0.1028	0.089	0.0746	0.0678	0.0624	0.058	0.055	0.0531	0.0517	0.0554	0.0644	0.0798	0.0903

Figure 24: Waiting time distribution for all types of surgery. Planning horizon: 15 weeks. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. P(waiting time = t, t > 15) = 0.

B.5 Distribution of waiting time per surgery type

(for different surgery types with planning horizon 50)

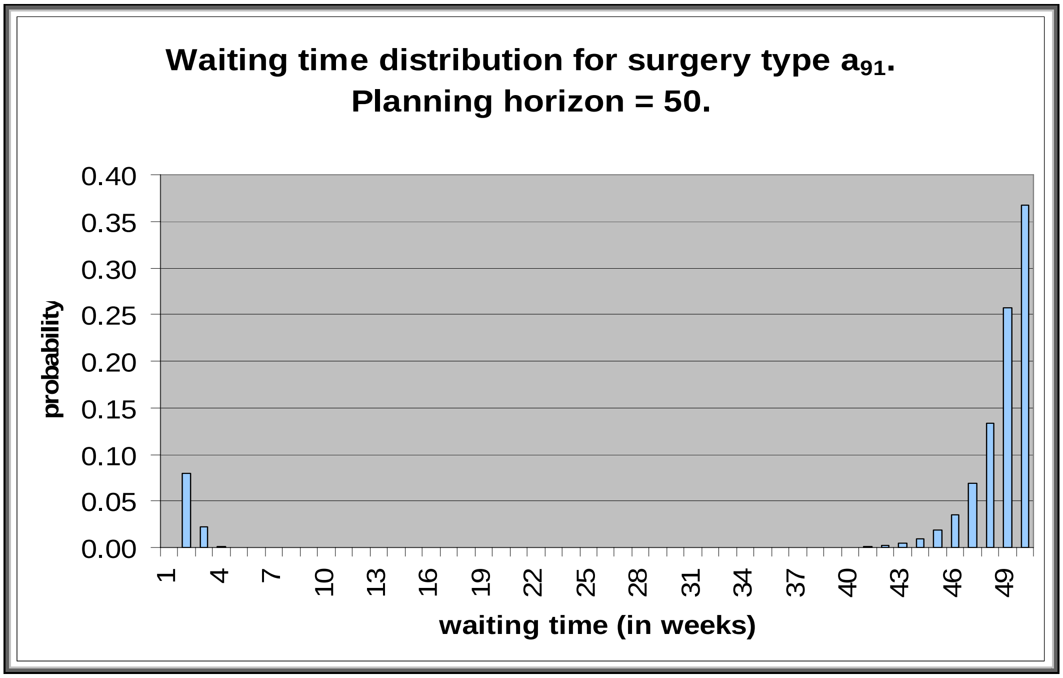


Figure 25: Waiting time distribution for surgery type a_{91}. $n_{91} = 3$. $\lambda_{91} = 3{,}39$. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. Planning horizon: 50 weeks.

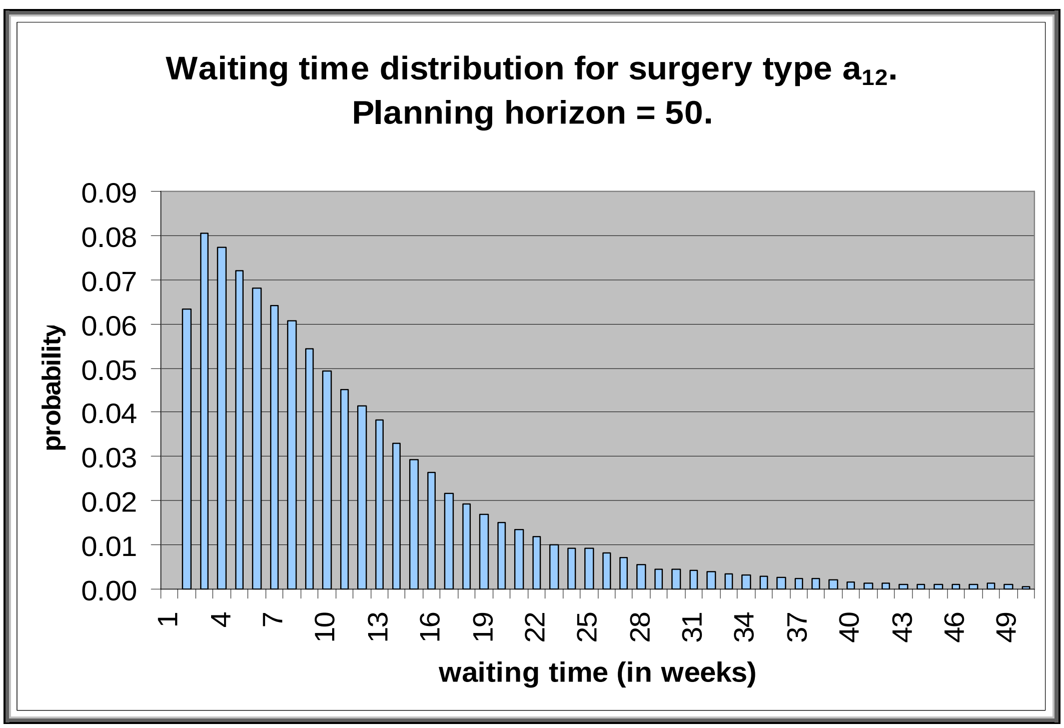


Figure 26: Waiting time distribution for surgery type a_{12}. $n_{12} = 3$. $\lambda_{12} = 2{,}98$. Probabilities obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks. Planning horizon: 50 weeks.

B.6 Distribution of bed occupation within cycle
(for planning horizons 2 and 50)

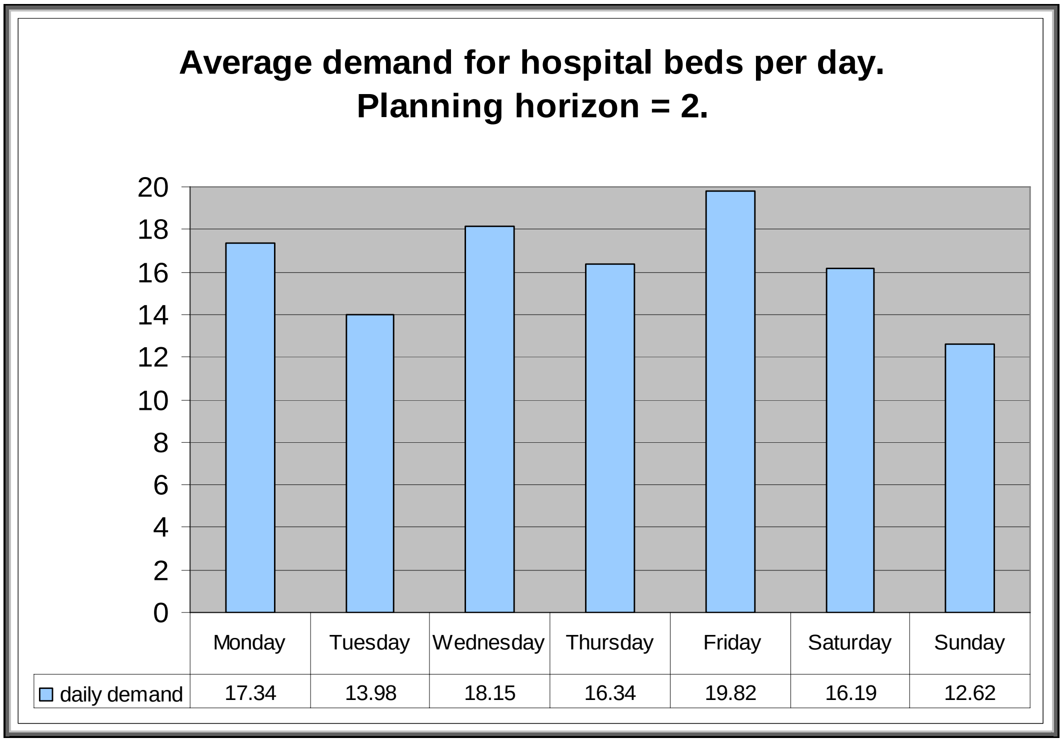


Figure 27:Average demand for hospital beds per day for planning horizon 2 weeks. Demand levels obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks.

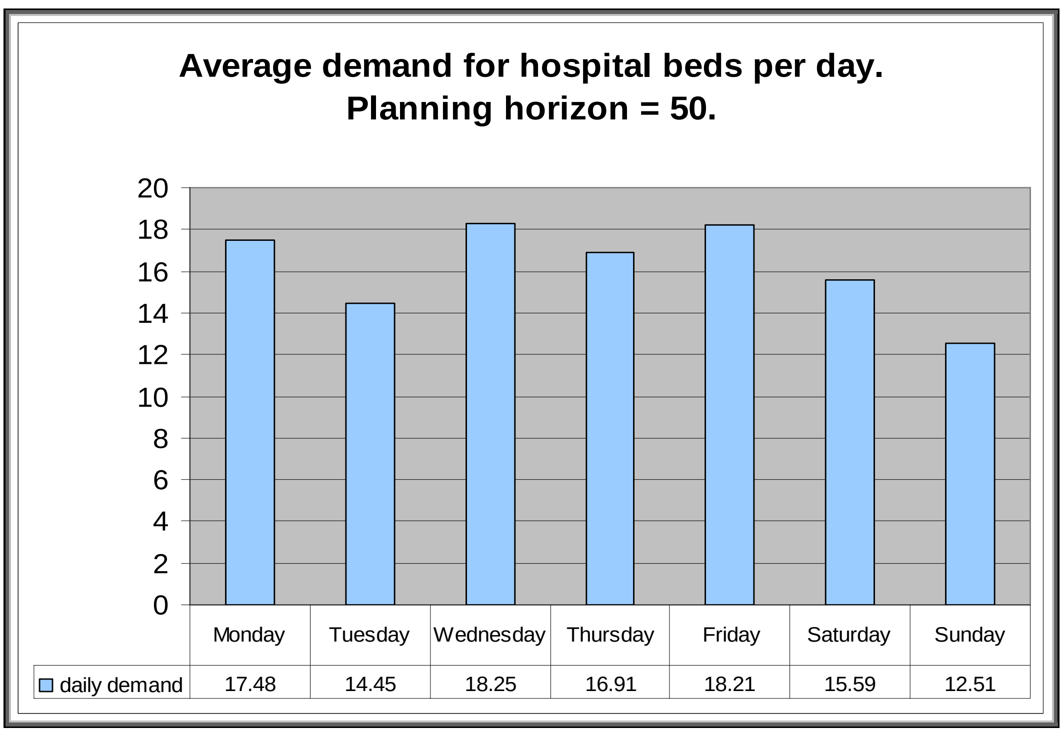


Figure 28: Average demand for hospital beds per day for planning horizon 50 weeks. Demand levels obtained over 100 simulations. Run length: 2000 weeks, warm-up period: 800 weeks.

B.7 QQ-plots: normality of demand levels for hospital beds

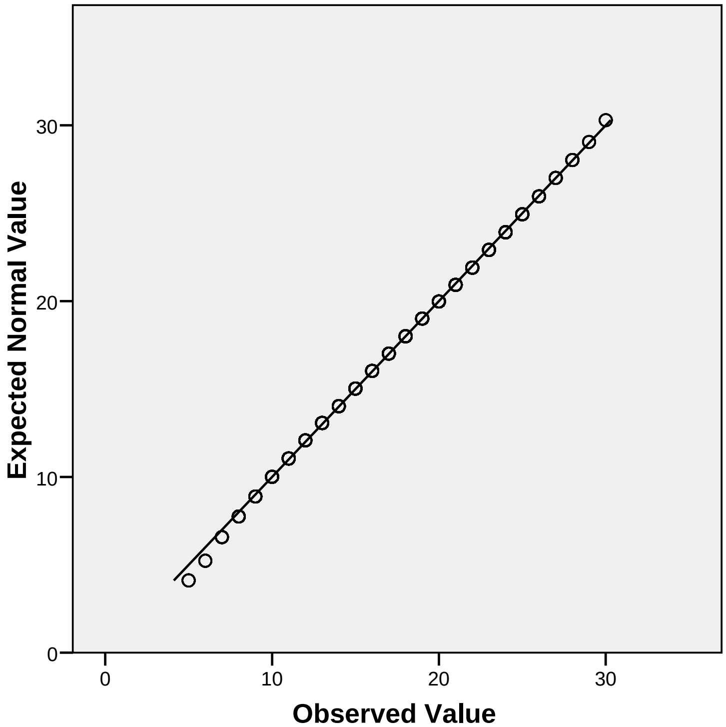


Figure 29: QQ-plot of demand levels for hospital beds from Monday to Friday

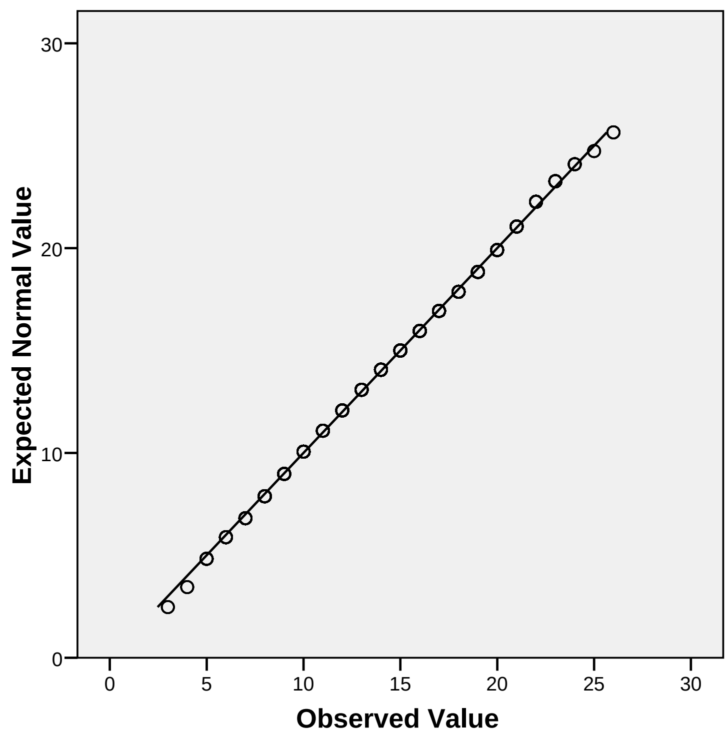


Figure 30: QQ-plot of demand levels for hospital beds during weekends

B.8 Time series identification of bed occupation levels

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.346	0.346	1674.2	0.000
		2	0.388	0.305	3782.3	0.000
		3	0.159	-0.050	4135.3	0.000
		4	0.119	-0.034	4333.2	0.000
		5	0.248	0.244	5196.3	0.000
		6	0.089	-0.059	5306.8	0.000
		7	0.554	0.528	9599.9	0.000
		8	0.066	-0.400	9660.4	0.000
		9	0.195	0.013	10192.	0.000
		10	0.027	0.004	10202.	0.000
		11	0.018	0.005	10207.	0.000
		12	0.161	0.059	10568.	0.000
		13	0.008	0.042	10569.	0.000
		14	0.469	0.223	13649.	0.000
		15	-0.005	-0.240	13650.	0.000
		16	0.135	-0.002	13904.	0.000
		17	-0.028	-0.007	13915.	0.000
		18	-0.031	0.001	13928.	0.000
		19	0.116	0.042	14119.	0.000
		20	-0.034	0.034	14135.	0.000
		21	0.425	0.141	16667.	0.000
		22	-0.046	-0.180	16697.	0.000
		23	0.096	-0.011	16826.	0.000
		24	-0.066	-0.012	16888.	0.000
		25	-0.067	-0.000	16951.	0.000
		26	0.087	0.042	17057.	0.000
		27	-0.063	0.024	17113.	0.000
		28	0.398	0.109	19338.	0.000
		29	-0.072	-0.148	19411.	0.000
		30	0.075	-0.002	19490.	0.000

Table 7: correlogram containing the (partial) autocorrelations of 14000 demand levels resulting from one simulation run with planning horizon 2.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.326	0.326	1492.6	0.000
		2	0.389	0.316	3609.9	0.000
		3	0.122	-0.084	3820.0	0.000
		4	0.085	-0.061	3920.2	0.000
		5	0.269	0.316	4930.6	0.000
		6	0.108	-0.030	5095.1	0.000
		7	0.597	0.535	10087.	0.000
		8	0.069	-0.415	10154.	0.000
		9	0.194	-0.032	10680.	0.000
		10	-0.028	-0.037	10691.	0.000
		11	-0.034	0.021	10707.	0.000
		12	0.170	0.083	11111.	0.000
		13	0.020	0.044	11116.	0.000
		14	0.521	0.264	14916.	0.000
		15	0.010	-0.236	14918.	0.000
		16	0.143	-0.031	15205.	0.000
		17	-0.073	-0.024	15279.	0.000
		18	-0.080	-0.003	15368.	0.000
		19	0.124	0.039	15583.	0.000
		20	-0.021	0.050	15589.	0.000
		21	0.477	0.158	18774.	0.000
		22	-0.035	-0.188	18792.	0.000
		23	0.102	-0.021	18937.	0.000
		24	-0.111	-0.017	19111.	0.000
		25	-0.115	0.000	19295.	0.000
		26	0.096	0.040	19424.	0.000
		27	-0.045	0.042	19453.	0.000
		28	0.457	0.130	22389.	0.000
		29	-0.050	-0.140	22424.	0.000
		30	0.092	-0.010	22543.	0.000

Table 8: correlogram containing the (partial) autocorrelations of 14000 demand levels resulting from one simulation run with planning horizon 50.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.710	0.710	7056.9	0.000
		2	0.545	0.083	11221.	0.000
		3	0.427	0.027	13775.	0.000
		4	0.365	0.066	15637.	0.000
		5	0.326	0.053	17128.	0.000
		6	0.307	0.057	18450.	0.000
		7	0.300	0.060	19714.	0.000
		8	0.271	-0.002	20743.	0.000
		9	0.243	0.008	21567.	0.000
		10	0.221	0.018	22251.	0.000
		11	0.206	0.018	22848.	0.000
		12	0.189	0.003	23348.	0.000
		13	0.180	0.019	23804.	0.000
		14	0.168	0.003	24199.	0.000
		15	0.160	0.014	24558.	0.000
		16	0.149	0.003	24868.	0.000
		17	0.135	-0.003	25123.	0.000
		18	0.129	0.014	25357.	0.000
		19	0.120	-0.000	25559.	0.000
		20	0.114	0.006	25741.	0.000
		21	0.100	-0.013	25880.	0.000
		22	0.096	0.012	26009.	0.000
		23	0.088	-0.003	26117.	0.000
		24	0.074	-0.015	26194.	0.000
		25	0.073	0.014	26268.	0.000
		26	0.074	0.010	26344.	0.000
		27	0.069	-0.004	26411.	0.000
		28	0.058	-0.012	26458.	0.000
		29	0.054	0.006	26499.	0.000
		30	0.055	0.010	26542.	0.000

Table 9: correlogram containing the (partial) autocorrelations of 14000 deseasonalised values indicating the deviation from the daily average resulting from one simulation run with planning horizon 2.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DB2(-1)	0.709901	0.005953	119.2563	0.0000

R-squared	0.503970	Mean dependent var	-0.000552
Adjusted R-squared	0.503970	S.D. dependent var	3.005154
S.E. of regression	2.116513	Akaike info criterion	4.337488
Sum squared resid	62705.80	Schwarz criterion	4.338027
Log likelihood	-30359.25	Hannan-Quinn criter.	4.337667
Durbin-Watson stat	2.118294		

Table 10: outcomes of the least squares regression of the seasonal AR(1) model, based on 14000 deseasonalised values resulting from one simulation run with planning horizon 2.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.713	0.713	7125.8	0.000
		2	0.576	0.137	11776.	0.000
		3	0.480	0.057	15005.	0.000
		4	0.416	0.051	17426.	0.000
		5	0.372	0.047	19363.	0.000
		6	0.343	0.047	21015.	0.000
		7	0.317	0.029	22420.	0.000
		8	0.277	-0.011	23493.	0.000
		9	0.245	0.004	24336.	0.000
		10	0.225	0.020	25048.	0.000
		11	0.215	0.026	25693.	0.000
		12	0.204	0.017	26277.	0.000
		13	0.193	0.012	26801.	0.000
		14	0.188	0.021	27294.	0.000
		15	0.177	0.007	27730.	0.000
		16	0.159	-0.008	28086.	0.000
		17	0.149	0.007	28397.	0.000
		18	0.137	-0.002	28658.	0.000
		19	0.126	0.002	28882.	0.000
		20	0.124	0.017	29099.	0.000
		21	0.113	-0.008	29279.	0.000
		22	0.099	-0.010	29417.	0.000
		23	0.089	-0.001	29528.	0.000
		24	0.083	0.005	29625.	0.000
		25	0.077	-0.000	29708.	0.000
		26	0.079	0.016	29797.	0.000
		27	0.083	0.015	29894.	0.000
		28	0.081	0.003	29987.	0.000
		29	0.075	-0.003	30066.	0.000
		30	0.073	0.007	30141.	0.000

Table 11: correlogram containing the (partial) autocorrelations of 14000 deseasonalized values indicating the deviation from the daily average resulting from one simulation run with planning horizon 50.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DB50(-1)	0.713360	0.005923	120.4340	0.0000

R-squared	0.508883	Mean dependent var	3.36E-05
Adjusted R-squared	0.508883	S.D. dependent var	2.388611
S.E. of regression	1.673933	Akaike info criterion	3.868301
Sum squared resid	39223.13	Schwarz criterion	3.868840
Log likelihood	-27075.17	Hannan-Quinn criter.	3.868480
Durbin-Watson stat	2.195513		

Table 12: outcomes of the least squares regression of the seasonal AR(1) model, based on 14000 deseasonalized values resulting from one simulation run with planning horizon 50.

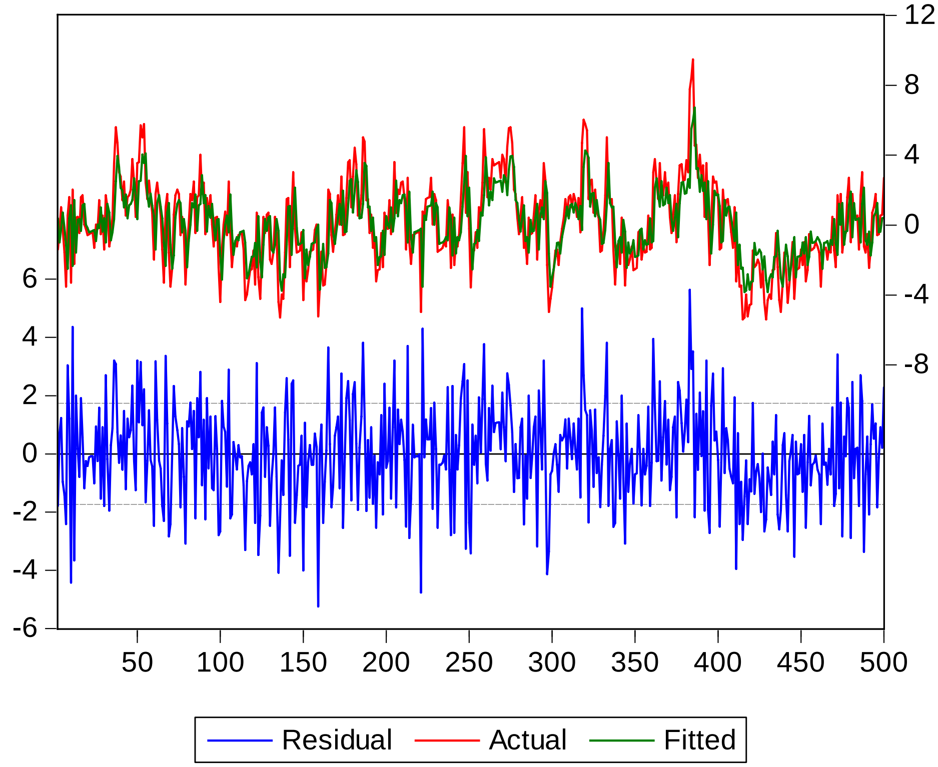


Figure 31: Actual, fitted and residual time series of the deseasonalised time series Y_t* for planning horizon 50, shown for the first 500 observations, after a warm-up period of 800 weeks.