

Improve OR-Schedule to Reduce Number of Required Beds

**J. Theresia van Essen · Joël M. Bosch · Erwin
W. Hans · Mark van Houdenhoven · Johann L.
Hurink**

Received: date / Accepted: date

Abstract After surgery most of the surgical patients have to be admitted in a ward in the hospital. Due to financial reasons and an decreasing number of available nurses in the Netherlands over the years, it is important to reduce the bed usage as much as possible. One possible way to achieve this is to create an operating room (OR) schedule that spreads the usage of beds nicely over time, and thereby minimizes the number of required beds. An OR-schedule is given by an assignment of OR-blocks to specific days in the planning horizon and has to fulfill several resource constraints. Due to the stochastic nature of the length of stay of patients, the analytic calculation of the number of required beds for a given OR-schedule is a complex task involving the convolution of discrete distributions. In this paper, two approaches to deal with this complexity are presented. First, a heuristic approach based on local search is given, which takes into account the detailed formulation of the objective. A second approach reduces the complexity by simplifying the objective function. This allows modeling and solving the resulting problem as an ILP. Both approaches are tested on data provided by HagaZiekenhuis in the Netherlands. Furthermore, several what-if scenarios are evaluated. The computational results show that the approach that uses the simplified objective function provides better solutions to the original problem. By using this approach, the number of required beds for the considered instance of HagaZiekenhuis can be reduced by almost 20%.

J. Theresia van Essen · Joël M. Bosch
Center of Healthcare Operations Improvement and Research (CHOIR),
University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
HagaZiekenhuis, P.O. Box 40551, 2504 LN Den Haag, The Netherlands
E-mail: j.t.vanessen@utwente.nl

Erwin W. Hans · Johann L. Hurink
Center of Healthcare Operations Improvement and Research (CHOIR),
University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Mark van Houdenhoven
HagaZiekenhuis, P.O. Box 40551, 2504 LN Den Haag, The Netherlands

Keywords Operating Room Scheduling · Ward Occupancy · Simulated Annealing · Integer Programming

1 Introduction

Due to an ageing population and increased health care costs, hospitals are forced to use their resources more efficiently, meaning that the same amount of patients has to be treated with less resources or more patients with the same amount of resources. One of the resources used in hospitals are the beds on the nursing wards. The cost for acquiring these beds is not significant, however, the costs for maintaining and cleaning the beds, and the labour costs for treating the admitted patients are significantly high. Also, more and more costly technical appliances, such as interactive screens, are available at each bed. In addition, the number of available nurses in the Netherlands has been decreasing significantly over the years and will further decrease the coming years. Therefore, it is important to reduce the number of required beds as much as possible.

The starting point of this research was a request from HagaZiekenhuis, a hospital in the Netherlands, to get more insight in the factors that influence the bed occupancy. The operating room (OR) schedule is one of the most important factors that influence the bed occupancy, since most of the surgical patients have to be admitted at one of the wards after surgery. Therefore, it is important to consider the required bed capacity when creating the OR-schedule, which is the topic of this paper. First, we analyze this problem and investigate several approaches to solve it, and second, we show what improvements can be made in HagaZiekenhuis.

There is a vast amount of literature on OR planning and scheduling. Hulshof et al. [12] provide an overview of papers that consider this topic. Several of these papers address the issue of considering the wards when creating an OR-schedule. The first paper that addressed this topic is the work of Beliën and Demeulemeester [4]. They schedule blocks of elective surgeries of the same type by assigning them to a day in the planning horizon while minimizing the number of required beds. They assume that the length of stay (LOS) is given by a multinomial distribution which differs per surgery type. The number of required beds resulting from an OR-schedule is approximated in several ways, however, no exact formulation is used. Beliën et al. [5] extend this approach by including multiple wards instead of one, allowing different block lengths and by scheduling individual surgeons instead of surgeon groups. In addition, they develop a decision support system which visualizes the OR-schedule and the resulting bed occupancy.

Van Oostrum et al. [14] schedule surgical procedures instead of OR-blocks by assigning the procedures to an OR and to a day in the planning horizon. The LOS of the patients is assumed to be deterministic and by using this deterministic LOS, the number of required beds is minimized. As Van Houdenhoven et al. [11] and Van Oostrum et al. [14], Adan et al. [1], [2] schedule surgical procedures by assigning them to a day in the planning horizon. But opposite to [11] and [14], they assume a fixed amount of beds is available at the hospital and minimize the over- and underuti-

lization of these beds. Thus, the number of required beds is not minimized, but their use is optimized.

Chow et al. [8] develop an integer linear programming model to generate improved OR-schedules in terms of the maximum expected bed occupancy. This expected bed occupancy is calculated by using the expected LOS of surgery types and after this, the average bed occupancy per day is determined by means of simulation.

As in [4], Vanberkel et al. [15], [16] schedule blocks of surgeries of the same type and assume a multinomial distribution of the LOS which differs per surgery type. However, in contrast to [4], they analytically determine the complete probability distribution of the number of occupied beds for each day in the planning horizon. The goal of this approach is not to minimize the number of required beds, but to develop a model that can be used as an evaluation tool for the OR-schedule.

Bekker and Koeleman [3] apply a time-dependent analysis to determine the mean bed occupancy per day. In addition, they use a quadratic programming model to determine the optimal number of elective admissions per day such that an average desired bed occupancy per day is achieved.

For the operational planning level, Cardoen et al. [6], [7] propose a mixed integer linear programming approach and a column generation approach to determine the sequence in which patients must have surgery on a given day such that the peak use of recovery beds is minimized. They assume the LOS on the recovery to be deterministic and thus, no stochasticity is included. Fei et al. [9] also focus on the sequence in which patients must have surgery, however, they consider the number of beds available at the recovery to be fixed and therefore, do not focus on minimizing the peak use.

Many of the discussed papers consider the expected LOS of patients instead of the LOS probability distribution or focus on minimizing the maximum expected bed occupancy without considering the bed occupancy probability distribution. However, in practice, the LOS of patients is stochastic, and thus, it is important to also consider the variance in the bed occupancy. In this paper, we both incorporate the stochasticity of the LOS and of the bed occupancy to account for these variances. As in practice most hospitals use a cyclic OR-schedule, we develop an OR-schedule by assigning OR-blocks to a day in the planning horizon. We assume an OR-block consists of not only one but several surgical procedure types to make the problem more suitable for application. Because we schedule surgery types and not individual patients, this scheduling problem is considered to be on the tactical level. As in [4] and [16], we assume the LOS to be multinomial distributed. This distribution can easily be obtained from historical data. We use the analytical formulation of Vanberkel et al. [16] to determine the number of required beds, and minimize this number while taking into account several restrictions on the OR-schedule such as OR, surgeon, and instrument availability. As the problem originated in HagaZiekenhuis, we focus on resource constraints that are relevant in the setting of this hospital. However, it is possible to add additional constraints without destroying the structure of the developed model. Note that we only consider the scheduling of elective surgeries, but it is quite easy to also include emergency surgeries when determining the number of required beds.

The developed model is discussed in Section 2 and it consists of linear constraints and a complex non-linear objective function which involves the convolution of discrete distributions. To deal with this complexity, we introduce in Section 3 two dif-

ferent approaches to approximate the optimal solution. The first approach is a local search approach, which takes into account the complex formulation of the objective function. We have chosen to use Simulated Annealing (SA) since this approach is easy to implement and has proven to be successful for other combinatorial optimization problems. The second approach reduces the complexity of the problem by linearizing the objective function. Although we prove the resulting problem to be NP-hard, we model and solve the resulting problem as an ILP, because our considered instances are small enough to be solved within a reasonable amount of time. By comparing these two different approaches, we can determine whether it is better to not fully search the solution space with a complete evaluation of the objective function or to approximate the objective function and search the complete solution space. In fact, it is investigated if it is necessary to model the problem in full detail to be able to achieve a good solution.

The comparison is performed on data provided by HagaZiekenhuis. However, we use this data not just for comparing the two contrasting approaches, but we also aim to support HagaZiekenhuis by determining which resources are a bottleneck for minimizing the number of required beds. We do this by considering several what-if scenarios that relax some or all of the resource constraints. The computation results of the comparison and the what-if scenarios are given in Section 4. Section 5 presents conclusions and gives recommendations for further research.

2 Problem Formulation

Hospitals aim to use as few beds as possible. When less beds are used, as a consequence less personal is needed and less money is spent on cleaning and maintaining these beds. Another effect of using less beds is that also the bed occupancy during the week is better levelled and this reduces stress on the wards.

In hospitals, the number of beds occupied during the week is mostly determined by the OR-schedule. In general, a patient is admitted on the day of surgery, and after surgery, the patient must stay in the hospital for a few extra days. Thus, in order to influence the number of beds used, we should create an OR-schedule which minimizes the number of required beds, and thereby levels the amount of occupied beds as much as possible. HagaZiekenhuis, like many other hospitals, uses a cyclic OR-schedule which repeats every T days. This means that we have to develop such a cyclic OR-schedule for T days and not an OR-schedule for a whole year.

An OR-schedule consists of OR-blocks which are assigned to days of the planning cycle. Each OR-block is dedicated to a specific specialism or specialist and is filled with several surgery types chosen by this specialism or specialist. Thus, each specialism or specialist provides a list containing as many OR-blocks as this specialism or specialist gets during a period of T days. It only remains to assign these OR-blocks to a specific day in the planning horizon to create an OR-schedule.

2.1 Restrictions

In this section, we discuss several restrictions on the OR-schedule that are relevant for HagaZiekenhuis. Although we only provide these specific constraints, it is possible to add additional constraints without destroying the structure of the chosen approach.

Let K be the given set of OR-blocks. To each OR-block $k \in K$ we have to assign to a specific day $t \in \mathcal{T} = \{1, \dots, T\}$. For this, we define binary decision variables X_{kt} which are one when OR-block $k \in K$ is assigned to day $t \in \mathcal{T}$, and zero otherwise. Then, the following constraints ensure that all OR-blocks $k \in K$ are assigned to a day in the OR-schedule:

$$\sum_{t \in \mathcal{T}} X_{kt} = 1, \quad \forall k \in K. \quad (1)$$

The assignment of OR-blocks to days is limited by several constraints. First, some OR-blocks can only be performed in a subset of the available ORs, because, for example, special equipment is needed which is not available in all ORs. To model this, we define a set J of different OR types and for OR type $j \in J$, we denote by the subset $K_j \subseteq K$ the OR-blocks that can be performed in OR type $j \in J$. In addition, the number of available ORs of type $j \in J$ on day $t \in \mathcal{T}$ is limited and denoted by a_{jt} . The following constraints ensure that the assignment of OR-blocks to days fulfils these limitations:

$$\sum_{k \in K_j} X_{kt} \leq a_{jt}, \quad \forall j \in J, t \in \mathcal{T}. \quad (2)$$

Each OR-block is allocated to a specific surgeon type, because most surgeons in HagaZiekenhuis are specialized in a certain set of surgery types. The surgeon types are given by set S and the OR-blocks that have to be performed by surgeon type $s \in S$ are given by subset $K_s \subseteq K$. The number of available surgeons of type $s \in S$ on day $t \in \mathcal{T}$ is limited and denoted by b_{st} . The following constraints ensure that the assignment of OR-blocks to days fulfils these restrictions:

$$\sum_{k \in K_s} X_{kt} \leq b_{st}, \quad \forall s \in S, t \in \mathcal{T}. \quad (3)$$

Each OR-block consists of several surgeries which must be performed consecutively. The total set of possible surgery types is defined by I and the number of surgeries of a specific type $i \in I$ performed in OR-block $k \in K$ is denoted by o_{ik} . For each surgery type $i \in I$ a specific set of instruments is needed to perform the surgery. The set of all available instrument sets is given by set R , and w_{kr} denotes how many instrument sets $r \in R$ are needed for OR-block $k \in K$. Because a limited number of instrument sets is available and the instrument sets have to be sterilized after surgery, the number of surgeries which need instrument set $r \in R$ scheduled per day is limited by q_r . This is ensured by the following constraints:

$$\sum_{k \in K} X_{kt} w_{kr} \leq q_r, \quad \forall r \in R, \forall t \in \mathcal{T}. \quad (4)$$

Note that the o_{ik} values are not used explicitly, but are covered in the w_{kr} values. However, we have introduced the values since they are needed in the next section.

2.2 Objective Function

The constraints (1)-(4) are the restrictions on the decision variables X_{kt} , and therefore, describe the set \mathcal{S} of feasible solutions. In this subsection, we specify the quality of a feasible solution $S \in \mathcal{S}$ given by the maximum number of beds needed during the entire planning horizon. To determine this number for a proposed OR-schedule, we have to determine the bed occupancy for each day. If we would specify the bed occupancy by a deterministic measure (e.g. maximum or expected number of used beds), we do not take the stochastic nature of the LOS into account. Using the expected bed occupancy per day results in canceling patients for surgery because quite often not enough beds are available to admit them after surgery. Using the maximum number of beds needed leads to a solution for which almost always too much beds are available. Therefore, we choose to calculate the complete bed occupancy probability distribution per day and afterwards take the p -percentile of these probability distributions to ensure that sufficient beds are available with p percent chance. Since these percentiles represent the number of beds needed on day $t \in \mathcal{T}$ of the planning horizon we obtain the number of beds needed in the wards by taking the maximum over all days.

For a given OR-schedule, the probability distribution of the bed occupancy can be obtained by using the LOS distribution of all surgery types scheduled in the OR-blocks. The LOS distribution of each surgery type $i \in I$ is given by a multinomial distribution which can be obtained from historical data. By taking discrete convolutions of these LOS distributions, we compute the probability distribution of the bed occupancy for each day as in [16]. In the following paragraphs, we shortly explain this method. For a more detailed description of this method, we refer to Vanberkel et al. [16].

The probability distribution of the LOS of surgery type $i \in I$ is given by values l_n^i , which denote the probability that the LOS of a surgery type $i \in I$ is exactly n days ($n \in \{1, \dots, L_i\}$), where L_i is the maximum LOS of surgery type $i \in I$. From this, we can determine the probability that a patient who is still admitted on day n is discharged that day, which is denoted by d_n^i . Note that d_1^i denotes the probability that a patient is discharged on the day of surgery (i.e., an outpatient surgery) and $d_{L_i}^i = 1$. The value of d_n^i is given by:

$$d_n^i = \frac{l_n^i}{\sum_{m=n}^{L_i} l_m^i}. \quad (5)$$

From these values, we can calculate the probability distribution $h_n^{ik}(x)$ that n days after carrying out OR-block $k \in K$, x patients of surgery type $i \in I$ are still in recovery. Recall that o_{ik} denotes the number of patients of type $i \in I$ assigned to OR-block $k \in K$. Therefore, these probabilities are computed recursively as follows:

For $n = 1$:

$$h_1^{ik}(x) = \begin{cases} 1 & \text{when } x = o_{ik}, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

For $n > 1$:

$$h_n^{ik}(x) = \sum_{y=x}^{o_{ik}} \binom{y}{x} (d_{n-1}^i)^{y-x} (1 - d_{n-1}^i)^x h_{n-1}^{ik}(y). \quad (7)$$

Next, we take discrete convolutions of $h_n^{ik}(x)$ over all $i \in I$ to determine the bed occupancy caused by OR-block $k \in K$. This gives the probability $\tilde{h}_n^k(x)$ that n days after carrying out OR-block $k \in K$, x patients are still in recovery:

$$\tilde{h}_n^k(x) = h_n^{1k}(x) * h_n^{2k}(x) * \dots * h_n^{Ik}(x). \quad (8)$$

Because we use a cyclic OR-schedule, which repeats every T days, patients who had surgery in one cycle may still be admitted in the next cycle. Therefore, we must take into account $\lceil N_k/T \rceil$ consecutive cycles, where N_k denotes the maximum LOS of the surgeries scheduled in OR-block $k \in K$, i.e., $N_k = \max_{i \in I | o_{ik} \geq 1} L_i$. In other words, N_k represents the range of one cycle of the OR-schedule. Now, by again using discrete convolutions, we can compute the probability distribution $H_t^k(x)$ of recovering patients on day $t \in \mathcal{T}$ of the cycle caused by OR-block $k \in K$ as follows:

$$H_t^k(x) = \tilde{h}_t^k(x) * \tilde{h}_{t+T}^k(x) * \tilde{h}_{t+2T}^k(x) * \dots * \tilde{h}_{t+\lceil N_k/T \rceil T}^k(x). \quad (9)$$

The last step in calculating the probability distribution of the bed occupancy is to combine the probability distributions $H_t^k(x)$ for all OR-blocks. To do this, we first have to shift the distribution $H_t^k(x)$ such that the patients who have surgery in OR-block $k \in K$ are admitted on the day they have surgery, i.e., the day $t \in \mathcal{T}$ for which $X_{kt} = 1$. The shifted probability distribution is denoted by $\bar{H}_t^k(x)$ and is defined as follows:

$$\bar{H}_t^k(x) = \begin{cases} H_{t-\hat{t}+1}^k(x) & \text{for } \hat{t} \text{ with } X_{k\hat{t}} = 1 \text{ and } \hat{t} \leq t, \\ H_{t-\hat{t}+T+1}^k(x) & \text{otherwise.} \end{cases} \quad (10)$$

By taking the discrete convolutions of $\bar{H}_t^k(x)$ over $k \in K$, we now determine the probability distribution of the bed occupancy for each day $t \in \mathcal{T}$ denoted by H_t , which is computed by:

$$H_t(x) = \bar{H}_t^1(x) * \bar{H}_t^2(x) * \dots * \bar{H}_t^K(x). \quad (11)$$

Thus, the number of required beds $\gamma(S)$ for a given solution $S \in \mathcal{S}$ is given by:

$$\gamma(S) = \max_{t \in \mathcal{T}} \min \left\{ x \left| \sum_{y=0}^x H_t(y) \geq \frac{p}{100} \right. \right\}. \quad (12)$$

The above derivation shows that it is not straightforward to quantify or predict the effect in the objective function when changing an OR-schedule and that it is hard to approximate the objective function. Moreover, calculating the objective function

takes a lot of computational time. To reduce this computation time, we can either choose to not fully search the solution space, or to approximate the objective function. This leads to the following two solution approaches: (1) use a local search heuristic based on the given constraints and objective function, and (2) approximate the objective function and incorporate this approximation in an Integer Linear Program (ILP), which includes the given constraints of the OR-schedule. For the second approach, the original objective value is determined afterwards to determine the number of beds needed in practice and to make a fair comparison between the two approaches. The comparison is used to determine whether it is better to not fully search the solution space with a complete evaluation of the objective function or to approximate the objective function and search the complete solution space. The two approaches are discussed in more detail in the following section.

3 Solution Approaches

Because of the complex objective function, we cannot solve reasonable instances of the problem to optimality. Therefore, we have to make a choice between using a heuristic procedure to solve the original problem and using a global approach to solve a simplified version of the problem. Both approaches do not guarantee to find the optimal solution, however, we want to investigate which of these two methods leads to better solutions. The first approach is based on Simulated Annealing, which is a local search method. The second approach is an Integer Linear Program (ILP), which uses an approximation of the objective function. In the following, these two approaches are discussed in more detail.

3.1 Local Search Approach: Simulated Annealing

The first approach we have chosen to solve our problem is Simulated Annealing (SA) [13]. SA is a local search procedure that in each step moves from the current solution, denoted by S_c , to a randomly selected neighbor solution, denoted by S_n . A solution is represented by the assignment of OR-blocks to a day in the planning horizon and is considered to be feasible when it satisfies constraints (1)-(4). As neighbor solutions, we consider all feasible solutions that can be obtained by swapping two OR-blocks that are assigned to two different days. We do not consider swapping two OR-blocks assigned to the same day, because this does not affect the objective value. If the randomly selected neighbor solution has a lower objective function value than the current solution, i.e., $\gamma(S_n) \leq \gamma(S_c)$, the neighbor solution is accepted as the new current solution. Otherwise, the neighbor solution is accepted with a probability that depends on the objective value of the current and neighbor solution and on a temperature parameter. This temperature parameter, denoted by T , gradually decreases during the search process, and therefore, also the acceptance probability of a worse solution decreases. The allowance of moving to worse solutions makes it possible to escape from a (poor) local minimum. For each temperature value, we perform ω iterations which together form a Markov chain, because the next state only depends on the current

state. Also, during the entire process of SA, we keep track of the best solution found this far. A more detailed description of this method is given by Kirkpatrick et al. [13].

Summarizing, our implementation of SA is as follows, where \bar{S} denotes the current best solution:

- Step 1. Start with the initial solution S_c given by the OR-schedule currently used at the hospital. Set $\bar{S} := S_c$ and determine the objective function $\gamma(S_c)$. Set the initial temperature, i.e., $\Gamma := \Gamma_s$, and a reduction factor α .
- Step 2. Repeat ω times:
 - Step a) Randomly select a neighbor solution S_n of the current solution and determine $\gamma(S_n)$.
 - Step b) If $\gamma(S_n) \leq \gamma(S_c)$, set $S_c := S_n$, and if $\gamma(S_n) \leq \gamma(\bar{S})$, set $\bar{S} := S_n$.
Otherwise, set $S_c := S_n$ with probability $e^{\frac{\gamma(S_c) - \gamma(S_n)}{\Gamma}}$.
- Step 3. Set $\Gamma = \alpha\Gamma$. If $\Gamma < \Gamma_f$, the final temperature, then stop; else, go to 2.

We choose the initial temperature Γ_s such that an increase of the objective value at the beginning of the procedure is accepted with a relatively high probability. This is needed to easily escape from a local minimum. We observe that the maximum increase of the objective value equals the maximum over the number of surgeries assigned to an OR-block minus the minimum over the number of surgeries assigned to an OR-block, i.e., $\max_k \sum_i o_{ik} - \min_k \sum_i o_{ik}$, because all patients are admitted on the day of surgery. We want to accept this maximum increase at the start of the procedure with probability 0.95, thus the initial temperature is given by:

$$\Gamma_s = \frac{\max_k \sum_i o_{ik} - \min_k \sum_i o_{ik}}{\ln 0.95} \quad (13)$$

Using the same approach, we determine the final temperature Γ_f . This temperature is chosen such that the probability of accepting the minimum increase of the objective value is very low. This means that at the end of the procedure almost no worse solution is accepted, and thus, the procedure converts to a local minimum. Since our objective function returns an integer amount of beds, the minimum increase is one bed. Thus, we set the threshold temperature Γ_f such that an increase of one bed is accepted with probability 0.001, i.e.,

$$\Gamma_f = \frac{-1}{\ln 0.001} \quad (14)$$

We set the number of iterations for each temperature value equal to the number of neighbour solutions that can be achieved by one swap of the current solution. This number is equal to the total amount of OR-blocks, i.e., ω equals the cardinality of set K , because in theory each OR-block can be swapped with one of the other OR-blocks. However, due to the restrictions described in Section 2.1, some swaps are prohibited. The reduction factor α is set to 0.95.

During preliminary runs, we tested the effect of increasing the length of the Markov chain and increasing the reduction factor α . Only increasing both factors at the same time improves the quality of the solution slightly, but also increases the computation time significantly. As the resulting computation time exceeds the computation time of the global approach, we use the length of the Markov chain and reduction factor α as described in this section.

3.2 Global Approach: Linearization of Objective Function

The local approach described in the previous subsection incorporates the complete evaluation of the objective function but only searches a part of the solution space. The global approach described in this subsection searches the entire solution space, but for such an approach the relation between a solution and the objective function must be evident. However, Section 2 shows that there is no straightforward and direct relation between a given OR-schedule and the resulting required number of beds. Therefore, we choose to linearize the objective function by replacing it with the maximum over the expected number of occupied beds per day. For calculating the expected number of occupied beds per day, we follow the approach of Beliën and Demeulemeester [4]. However, their formula does not work properly when the LOS of a patient exceeds the planning horizon. When the maximum LOS L_i of a surgery type $i \in I$ exceeds the length T of the planning horizon, patients of two cycles of the OR-schedule may be admitted simultaneously, i.e., patients from different cycles may overlap. In [4], it is assumed that this holds for all days in the planning horizon, however, this only holds for a few days as shown in Figure 1, where a situation is sketched with $T = 5$ and $l_i = 7$. This deficiency can be accounted for by a small modification in the weight factor used in the formula defined in [4]. This modification ensures that patients are only counted multiple times on the days of the planning horizon that the LOS of several cycles overlap.

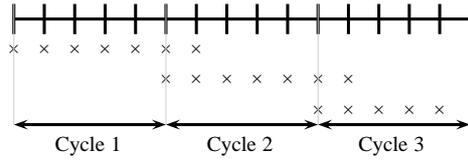


Fig. 1 Overlap patients multiple cycles

As a result, the expected number of occupied beds $\gamma_t(S)$ on day $t \in \mathcal{T}$ of the planning horizon is given by:

$$\begin{aligned} \gamma_t(S) = & \sum_{i \in I} \sum_{k \in K} \sum_{\tau \leq t} \left(\sum_{n=t-\tau+1}^{L_i} l_{in} o_{ik} \left\lceil \frac{n-t+\tau}{T} \right\rceil \right) X_{k\tau} \\ & + \sum_{i \in I} \sum_{k \in K} \sum_{\tau > t} \left(\sum_{n=T+t-\tau+1}^{L_i} l_{in} o_{ik} \left\lceil \frac{n-T-t+\tau}{T} \right\rceil \right) X_{k\tau} \end{aligned} \quad (15)$$

Equation (15) determines for each day t in the planning horizon, the impact of all OR-blocks on the bed occupancy. Thus, for all OR-blocks it is determined whether patients operated on in this OR-block are still admitted in the hospital while taking into account overlapping cycles. Note that the expected value associated with the probability distribution of the bed occupancy as given in Section 2.2 corresponds to

$\gamma_t(S)$. We can incorporate this linearized objective function in an ILP which includes the constraints given in Section 2.1. Then, the resulting ILP is:

$$\begin{aligned} \min_S \bar{\gamma}(S) & \quad (16) \\ \text{s.t. } & (1) - (4), (15) \\ & \bar{\gamma}(S) \geq \gamma_t(S), \forall t \in T \\ & X_{kt} \in \{0, 1\} \end{aligned}$$

This resulting problem is strongly NP-hard. For this, consider an instance with 3 ORs, a planning horizon of T days and thus $3T$ OR-blocks. Each of the $3T$ OR-blocks consists of a_k patients with $k = 1, \dots, 3T$ and $\sum_k a_k = Tb$, and each of the patients has a LOS of exactly one day. For this instance, determining whether there exists an OR-schedule which requires b beds is equivalent to determining whether there are T pairwise disjoint subsets $R_l \subset \{1, \dots, 3T\}$ such that $\sum_{R_l} a_k = b$ for $l = 1, \dots, T$, which is known as the 3-partition problem [10].

The ILP given by (16) consists of $|K|T$ binary variables and $(|J| + |S| + |R| + 2)T + |K|$ constraints. After solving the ILP, each OR-block is assigned to a day in the planning horizon leading to a solution $S \in \mathcal{S}$. For this solution, the real objective value, i.e., the maximum over the p -percentiles of the resulting bed occupancy probability distribution, can be determined using the method described in Section 2.2.

4 Results

The purpose of this section is twofold. First, we compare the local and global approach in Section 4.1. We do this by generating 100 random instances and solving these instances with the two approaches. For the global approach, the original objective value for the resulting OR-schedule is determined afterwards such that a fair comparison can be made between the local and global approach. The results are used to determine whether it is better to not fully search the solution space with a complete evaluation of the objective function or to approximate the objective function and search the complete solution space. Second, we consider several what-if scenarios for HagaZiekenhuis with the solution approach that performed best in Section 4.1. We use these scenarios to determine whether the resource availability in HagaZiekenhuis limits the reduction of the number of required beds.

The data used in the following subsections is based on data of HagaZiekenhuis. HagaZiekenhuis provided us with an OR-schedule of the orthopedics department with a planning horizon of 28 days, where up to three ORs and nine surgeons are available. The exact availability of the ORs and surgeons is given for each day in the planning horizon. The OR-schedule consists of 49 unique OR-blocks which have to be scheduled exactly once during the planning horizon. In total, 43 different surgery types are scheduled and the LOS varies from 1 to 59 days with an average LOS of 3.7 days. For each surgery type it is denoted which instrument sets are needed and for each of the ten available instrument sets it is given how many are available each

day. As the number of required beds, we take the maximum of the 95-percentile of the probability distribution of the bed occupancy over the 28 days.

4.1 Comparing Local and Global Approach

To determine which of the two considered approaches performs better, we have generated 100 random instances based on the data of HagaZiekenhuis. To generate the random instances, we vary the number of times a certain OR-block has to be performed during the planning horizon. For the original data set, each OR-block has to be performed exactly once, but for the generated random instances, some OR-blocks are not performed at all and some are performed multiple times in one cycle. To make sure there exists a solution that satisfies the fixed OR, surgeon and instrument sets availability, we generate the instances as follows: first, we randomly select for each available OR for each day in the planning horizon a surgeon available on that day. After this, we randomly select one of the OR-blocks which can be performed in the considered OR by the selected surgeon. During this selection process, we also consider the instrument sets availability. Because the number of times an OR-block is performed varies for the generated instances, we create instances that vary among the number of surgeries and the average LOS of the patients. Therefore, we analyze a broad range of different instances. The number of surgeries varies between 201 and 236 and the average LOS varies between 3.75 and 4.11 days.

The SA approach is implemented in CodeGear Delphi and the ILP is solved with CPLEX 12.3. Both methods are executed on a Intel Core2 Duo CPU P8600 2.40 GHz with 3.45 GB RAM. Since proving the optimality of a solution by CPLEX takes quite some time, we interrupt the solver after 10 minutes or when the integrality gap is less than 1%. The maximum achieved integrality gap for the 100 random instances is 1.38%. The results for the 100 instances can be found in Figure 2 where the dashed lines denote the average of the objective values for the two approaches. Note that the random instances are sorted according to the objective function values to clarify the differences between the two approaches.

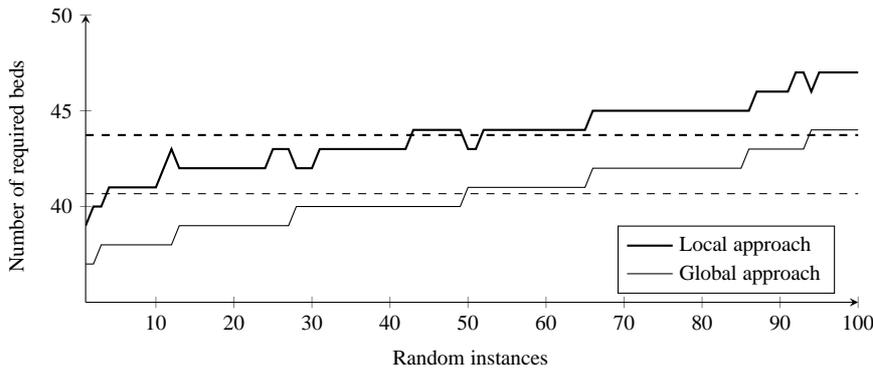


Fig. 2 Results 100 random instances

Figure 2 shows that the global approach performs better than the local approach for all of the 100 random instances. In addition, we see that the difference between the global and local objective value is almost everywhere the same. The difference in the objective values is two beds for 9% of the instances, three beds for 77% of the instances, four beds for 13% of the instances, and five beds for 1% of the instances. Note that both objective values represent the maximum of the 95-percentile of the probability distribution of the number of required beds over all days.

Figure 3 shows one of the random instances for which the difference between the two approaches can be explained nicely. The peaks for the global approach are flat, which results in a constant number of occupied beds. The peaks of the local approach fluctuate, which results in a higher number of required beds.

The solution time needed for the local approach varies between 32 and 74 seconds with an average of 42 seconds. The solution time needed for the global approach varies between 1.8 and 600 seconds with an average of 240 seconds. Therefore, as expected, the global approach takes longer than the local approach, but 4 minutes on average is still a reasonable amount of time.

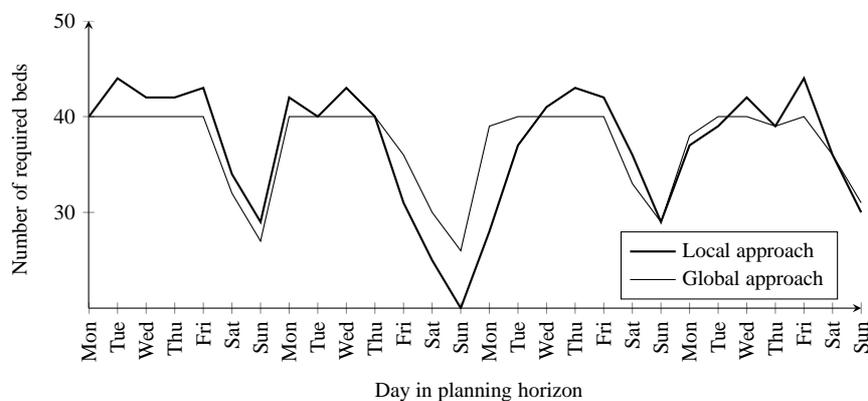


Fig. 3 Difference in levelled bed occupancy

The given results already indicate that for the considered instances it is not necessary to include the detailed objective function to determine a good OR-schedule. The correlation coefficients of the expected and the 95-percentile number of occupied beds per day for both the OR-schedules obtained by the local and global approach confirm this as it equals 0.998 for the local approach and 0.996 for the global approach. This high correlation is caused by the fact that the variance of the probability distributions of the bed occupancy is similar for each day in the planning horizon. Because we considered OR-schedules that vary in the number of surgeries and LOS probability distributions, we believe that this is often the case in practice. However, before using this approach on another instance, it should be verified that this also holds for the instance considered. Based on the instances considered in this research, we conclude that the global approach performs better although it does not incorporate the detailed objective function. Note that the detailed formulation of the objective

function is still used for comparing the solutions and is also needed to determine the number of beds required in practice.

4.2 What-if Scenarios

The starting point of this research was the request from HagaZiekenhuis to get more insight in the factors that influence the bed occupancy. Therefore, we use the global approach to show the reduction in the number of required beds when the OR-schedule is changed and we investigate whether the resource availability at HagaZiekenhuis limits this reduction. The hospital provided us an OR-schedule with a planning horizon of 28 days used by the orthopedics department. For the OR-schedule as provided by HagaZiekenhuis, 48 beds were needed to admit all surgical patients. We determined a new OR-schedule by solving the ILP and interrupting the solver after 10 minutes or when the integrality gap is less than 1%. To determine whether one or more of the constraints limit the improvement of the OR-schedule, we also consider the following scenarios:

- **Relax the number of available ORs per day:** The number of ORs of type $j \in J$ that are available each day is given by a_{jt} . By disregarding constraint (2), we do not restrict the model to schedule a fixed amount of OR-blocks per day. Since we relax the problem, we expect to come up with a schedule which requires less beds on the wards. We allow a maximum of 5 OR-blocks scheduled per day since 5 ORs are physically available at the operating department.
- **Relax the surgeon availability:** As with the previous scenario, the surgeon availability corresponds to a constraint in the model. To determine what restriction this constraint imposes on the resulting OR-schedule, and thus on the required number of beds, we solve the model without constraint (3).
- **Relax the instrument availability:** For each instrument set $r \in R$, q_r denotes the number of instrument sets available per day. By omitting constraint (4), we can determine the impact of this constraint on the number of required beds.
- **Relax all constraints:** By solving the model without all of the above mentioned constraints, we can determine the number of required beds when all resource capacities are unrestricted.
- **Relax all constraints including weekends:** Typically, no elective surgeries are performed during the weekends. However, it might be interesting to see which restriction this imposes on the objective function. Therefore, we also relaxed the availability of the ORs and surgeons during the weekends.

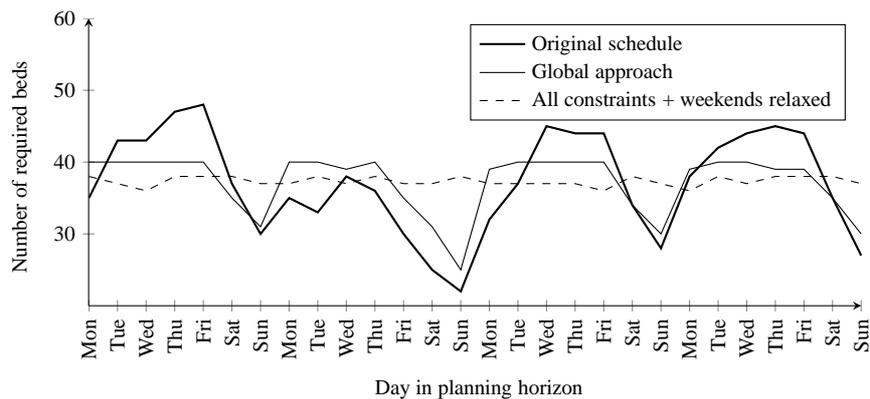
The results of the considered scenarios are given in Table 1. This table shows the expected number of required beds and the number of beds required when the 95-percentile is considered. In addition, we show the number of days in the planning horizon for which the maximum number of beds is achieved, which is denoted by ‘# Times Peak Achieved’. We also provide the integrality gap denoted by ‘Int. Gap’.

Table 1 shows that the global approach reduces the number of required beds from 48 to 40 by reassigning the OR-blocks while taking into account all resource constraints. The results for the other scenarios show that the OR and surgeon availability

Table 1 Results scenarios

	# Expected beds	# 95-perc. beds	# Times peak achieved	Int. gap (%)
Original	43.2	48	1	—
Global approach	34.7	40	14	1.06
Relax OR availability	34.4	40	5	1.93
Relax surgeon availability	34.1	40	5	1.08
Relax instrument availability	34.7	40	15	1.16
Relax all constraints	33.9	40	2	1.68
Relax all constraints including weekends	32.7	38	12	2.14

during the week and the number of available instrument sets do not influence the required number of beds, because for all considered scenarios, except the last one, the number of beds needed equals 40. However, the expected number of required beds and the number of times the peak bed occupancy is achieved indicates a slight improvement for the scenarios where the OR availability, the surgeon availability or all resource constraints are relaxed. The scenario where the constraints are also relaxed during the weekends decreases the number of required beds from 40 to 38 beds. The resulting bed occupancies over the entire planning horizon for the original OR-schedule used in HagaZiekenhuis, the OR-schedule obtained by the global approach and the OR-schedule for the last mentioned scenario are given in Figure 4.

**Fig. 4** Resulting bed occupancy three scenarios

The differences between the bed occupancy for the original OR-schedule and the one resulting from using the global approach might be explained by the differing number of surgeries scheduled per day. For the original OR-schedule, there is a peak in the number of surgeries scheduled per day at the start of the week and halfway through the week, while for the OR-schedule created by the global approach, there is only a peak at the start of the week. The global approach also schedules OR-blocks with a high average LOS at the end of the week. Note that the bed occupancy, shown

in Figure 4, for the global approach is rather flat during the week, however, during the weekends the bed occupancy is rather low. To flatten out these peaks, HagaZiekenhuis should consider to open the OR for elective surgeries during the weekends, because then, the number of beds needed can be reduced by two extra beds.

5 Conclusions

In this paper, we developed two approaches to improve the OR-schedule such that the number of required beds is reduced. The first approach incorporates the analytical formulation of the probability distribution of the bed occupancy and improves the OR-schedule by using a local search procedure. The second approach approximates the required number of beds by the expected bed occupancy, which enables us to solve the problem as an ILP. Both approaches are tested on 100 random instances to determine which of the two approaches provides the best solution to the original problem. The computational results show that the ILP with the simplified objective function performs the best. Note that after solving the ILP, the number of required beds is still determined by using the analytic formulation. The computational results show that the number of required beds at the orthopedic department of HagaZiekenhuis can be reduced by almost 20% when the ILP is used. None of the resources used at HagaZiekenhuis restrict the improvement that can be made to the OR-schedule, however, the number of required beds can be reduced slightly when the OR is also available for elective surgeries during the weekends.

Beliën and Demeulemeester [4] considered a similar problem as discussed in this paper, however, they focused on minimizing the total expected bed shortage instead of minimizing the number of required beds. They compared an SA approach, which considers the original objective function, and an ILP, which considers an approximation of the objective function. The approximation used in the ILP is given by the minimization of the maximum expected bed occupancy which is quite different from the original objective function that indirectly focuses on minimizing the expected bed occupancy for all days in the planning horizon. Opposite to our results, Beliën and Demeulemeester [4] conclude that the SA approach performs better than the ILP when compared based on the original objective function. This can be explained by the fact that the original and approximated objective function used by [4] differ significantly, while in our case, both objective functions are quite similar which is shown by the high correlation coefficient. Therefore, we conclude that approximating the objective function only provides good solutions to the original problem when the correlation between the approximated and original objective function is close to one. Therefore, when using the proposed solution approach in practice, it should be verified that the correlation between the approximated and original objective function for the considered instance is also close to one.

The approach developed in this paper only considers elective surgeries, because only these surgeries can be scheduled in advance. However, patients who have to undergo surgery immediately, and as a consequence, their surgery cannot be scheduled beforehand, also have to be admitted at one of the wards after surgery. By using the model of Vanberkel et al. [16], we can incorporate these emergency surgeries by in-

roducing dummy OR-blocks which are already fixed to a specific day in the planning horizon and contain the expected number of emergency surgeries. In this way, the arrival and admission of emergency patients is considered while determining a new OR-schedule for the elective surgeries, and thus, the total number of required beds is minimized and both elective and emergency patients can be admitted after surgery.

The developed approach can also be used to determine the admission schedule for non-surgical patients. To achieve this, we should schedule individual admissions instead of OR-blocks. This increases the complexity of the ILP as the number of variables increases. In addition, for the case of non-surgical patients, it is not defined how many admissions can be scheduled per day as this number may be unlimited. This also increases the complexity of the ILP due to the increasing solution space. Therefore, it might be needed to improve the solution approach to guarantee a reasonable computation time.

In the considered model, we assumed that the assignment of surgery types to OR-blocks is determined beforehand by the specialism of specialist. However, this assignment also influences the number of required beds on the wards. Therefore, it would be interesting to also incorporate this assignment when creating an OR-schedule such that the number of required beds can be reduced even further. Note that this also imposes some extra constraints on the model, because we also have to consider the stochastic duration of the surgeries such that the needed surgical time does not exceed the available surgical time. Thus, it would be interesting to investigate this problem in future research.

Another interesting topic for future research would be to take into account the available bed capacity at the wards when minimizing the number of required beds. When the available bed capacity at the wards equals 40 beds, it is not necessary to reduce the number of required beds to 38. In addition, it might be beneficial to free as many wards as possible during the weekends. This reduces costs as less staff is needed during the costly weekends.

Acknowledgements This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs.

References

1. Adan, I., Bekkers, J., Dellaert, N., Jeunet, J., Vissers, J.: Improving operational effectiveness of tactical master plans for emergency and elective patients under stochastic demand and capacitated resources. *European Journal of Operational Research* **213**(1), 290–308 (2011)
2. Adan, I., Bekkers, J., Dellaert, N., Vissers, J., Yu, X.: Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. *Health Care Management Science* **12**(2), 129–141 (2009)
3. Bekker, R., Koeleman, P.M.: Scheduling admissions and reducing variability in bed demand. *Health Care Management Science* **14**(3), 237–249 (2011)
4. Beliën, J., Demeulemeester, E.: Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research* **176**(2), 1185 – 1204 (2007)
5. Beliën, J., Demeulemeester, E., Cardoen, B.: A decision support system for cyclic master surgery scheduling with multiple objectives. *Journal of Scheduling* **12**(2), 147–161 (2009)
6. Cardoen, B., Demeulemeester, E., Beliën, J.: Optimizing a multiple objective surgical case sequencing problem. *International Journal of Production Economics* **119**(2), 354–366 (2009)

7. Cardoen, B., Demeulemeester, E., Beliën, J.: Sequencing surgical cases in a day-care environment: An exact branch-and-price approach. *Computers & Operations Research* **36**(9), 2660–2669 (2009)
8. Chow, V.S., Puterman, M.L., Salehirad, N., Huang, W., Atkins, D.: Reducing surgical ward congestion through improved surgical scheduling and uncapacitated simulation. *Production and Operations Management* **20**(3), 418–430 (2011)
9. Fei, H., Meskens, N., Chu, C.: A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Computers & Industrial Engineering* **58**(2), 221–230 (2010)
10. Garey, M.R., Johnson, D.S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co Ltd (1979)
11. van Houdenhoven, M., van Oostrum, J.M., Wullink, G., Hans, E., Hurink, J.L., Bakker, J., Kazemier, G.: Fewer intensive care unit refusals and a higher capacity utilization by using a cyclic surgical case schedule. *Journal of Critical Care* **23**(2), 222–226 (2008)
12. Hulshof, P.J.H., Boucherie, R.J., van Essen, J.T., Hans, E.W., Hurink, J.L., Kortbeek, N., Litvak, N., Vanberkel, P.T., van der Veen, E., Veltman, B., Vliegen, I.M.H., Zonderland, M.E.: Orchestra: an online reference database of or/ms literature in health care. *Health Care Management Science* **14**(4), 383–384 (2011)
13. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. *Science* **220**(4598), 671–680 (1983)
14. van Oostrum, J.M., van Houdenhoven, M., Hurink, J.L., Hans, E.W., Wullink, G., Kazemier, G.: A master surgical scheduling approach for cyclic scheduling in operating room departments. *OR Spectrum* **30**(2), 355 – 374 (2008)
15. Vanberkel, P.T., Boucherie, R.J., Hans, E.W., Hurink, J.L., van Lent, W.A.M., van Harten, W.H.: Accounting for inpatient wards when developing master surgical schedules. *Anesthesia & Analgesia* **112**(6), 1472–1479 (2011)
16. Vanberkel, P.T., Boucherie, R.J., Hans, E.W., Hurink, J.L., van Lent, W.A.M., van Harten, W.H.: An exact approach for relating recovering surgical patient workload to the master surgical schedule. *Journal of the Operational Research Society* **62**(10), 1851 – 1860 (2011)