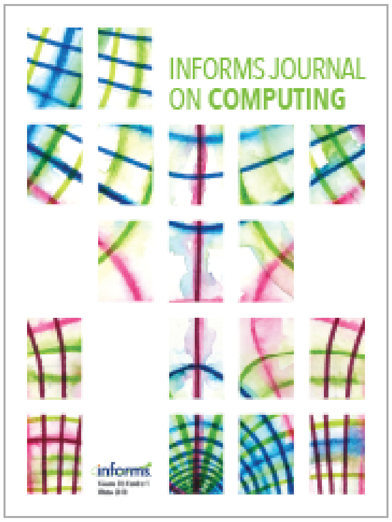

INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Progressive Hedging Approach for Surgery Planning Under Uncertainty

Serhat Gul, Brian T. Denton, John W. Fowler

INFORMS Journal on Computing
Vol. 27, No. 4, Fall 2015, pp. 755–772
ISSN 1091-9856 (print) | ISSN 1526-5528 (online)

http://dx.doi.org/10.1287/ijoc.2015.0658

A Progressive Hedging Approach for Surgery Planning Under Uncertainty

Serhat Gul
Department of Industrial Engineering, TED University, 06420, Kolej Çankaya, Ankara, Turkey,
serhat.gul@tedu.edu.tr

Brian T. Denton
Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109,
btdenton@umich.edu

John W. Fowler
Department of Supply Chain Management, Arizona State University, Tempe, Arizona 85287,
john.fowler@asu.edu

We propose a multistage stochastic mixed-integer programming formulation for the assignment of surgeries to operating rooms over a finite planning horizon. We consider the demand for and the duration of surgery to be random variables. The objective is to minimize three competing criteria: expected cost of surgery cancellations, patient waiting time, and operating room overtime. We discuss properties of the model and an implementation of the *progressive hedging* algorithm to find near-optimal surgery schedules. We conduct numerical experiments using data from a large hospital to identify managerial insights related to surgery planning and the avoidance of surgery cancellations. We compare the progressive hedging algorithm to an easy-to-implement heuristic for practical problem instances to estimate the value of the stochastic solution. Finally, we discuss an implementation of the progressive hedging algorithm within a rolling horizon framework for extended planning periods.

Keywords: surgery planning; scheduling; stochastic programming; progressive hedging; heuristics
History: Accepted by Allen Holder, Area Editor for Applications in Biology, Medicine, and Health Care; received August 2013; revised June 2014, April 2015; accepted May 2015. Published online December 9, 2015.

1. Introduction

Operating rooms (ORs) are responsible for a large portion of total hospital revenues (HFMA 2003) and costs. Therefore, substantial revenue increases and cost reductions might be achieved through better management of ORs. In many cases, a two-phase process is followed to plan for a day of surgery. In the first phase, surgeries are assigned to days. This is often done weeks prior to the day of surgery. In the second phase, surgeries are sequenced and scheduled within ORs, often days prior to the day of surgery. Surgery assignment, sequencing, and scheduling decisions have the potential to influence the cost of overtime and cancellations.

Surgery cancellations result in prolonged hospital stays, delayed perioperative treatments, and repeated preoperative tests and treatments, as well as anxiety and emotional hardship for the patient and the patient's family (for other causes of surgery cancellations, see Dexter et al. 2014). A recent study indicates that as many as 50% of cancellations can be prevented (Gillen et al. 2009). One way to prevent them is to create surgery plans that carefully consider future uncertainty in demand and duration of surgeries.

Designing surgery plans is a complicated task because of the uncertainty in demand for and duration of surgery. The occurrence of urgent and emergent cases is one of the reasons that uncertainty in demand is a significant factor (Gerchak et al. 1996, Zonderland et al. 2010). However, McManus et al. (2003) found that the principal source of uncertainty in demand is not the unscheduled cases. There also exists a considerable amount of uncertainty in demand for elective surgeries. Thus, the mix of surgeries requested varies from day to day. Combining this with uncertainty in the duration of individual surgeries (Gul et al. 2011) makes the task of creating surgery plans challenging.

In this article, we study the problem of assignment of surgeries into future days and ORs over a finite planning horizon. Decisions in our model include scheduling and rescheduling of surgeries, where the latter results from cancellations that may occur on the day of surgery. Cancellations are an important consideration, because they are commonly observed and they significantly influence efficiency and quality of patient care. For example, one study found that the percentage of cancelled surgeries ranges between 5% and 20% across institutions in the United States (Argo et al. 2009), and the cancellation rates on the day of surgery are reported to be greater than 10% in many facilities (Schuster et al. 2011). Evidence suggests that careful rescheduling of cancelled surgeries may avoid increases in variability in surgical workload (Epstein and Dexter 2013).

We formulate a multistage stochastic mixed-integer program for surgery planning. We consider three competing criteria in the objective function: expected cost of surgery cancellations, patient waiting (the number of days between when the surgery is requested and the day it is performed), and OR overtime. We implement a customized version of the progressive hedging algorithm (PHA), which exploits the underlying problem structure to find near-optimal surgery plans. In particular, we propose a Lagrangian multiplier update method that is motivated by having binary variables in the relaxed constraints. We also test a penalty update method that uses the information about the convergence pattern of the primal and dual variables. We compare the PHA with a deterministic heuristic that is similar to planning rules likely to be used in practice. We also discuss an implementation of the PHA within a rolling horizon framework for extended planning periods. We use our model to solve instances of the problem based on data from a large medical center. Our results provide insight regarding answers to the following three questions:

1. Which factors in the model have significant impact on surgery cancellations?
2. What is the value of considering the randomness in demand and total daily surgery durations when planning surgeries?
3. Which PHA parameters have the greatest impact on the performance and solution quality of the PHA?

The remainder of this article is organized as follows. In the next section, a brief literature review of surgery planning studies is presented. In §3, the decision-making process is described and a multistage stochastic mixed-integer programming model is formulated. In §4, our implementation of the PHA is discussed. In §5, an experimental study is presented. Concluding remarks are given in §6.

2. Literature Review

The literature review is divided into three categories. The first category is deterministic models for OR planning. The second category includes articles that consider uncertainties related to the surgery durations, but not demand uncertainty for elective surgeries. Since the demand for elective surgeries over the planning period is assumed to be known in these studies, the models are static; i.e., all decisions are given at the beginning of the planning period. The third category of articles considers uncertain elective surgery demands in the context of dynamic planning.

Among articles in the first category of research, Guinet and Chaabane (2003) used a two-phase approach based on weekly OR planning. Their integer-programming model assigns surgeries to ORs and particular time blocks of each day over a finite planning horizon. The objective is to minimize patients' indirect waiting time, i.e., the time between the procedure and hospitalization date, and OR overtime. Their model also considers equipment constraints and availability of surgeons. Fei et al. (2008, 2009, 2010) proposed an integer-programming model for optimal assignment of surgeries to ORs and days to minimize OR overtime and maximize OR utilization. They formulated the problem as a set partitioning model and applied a column generation-based heuristic to solve the model.

In the second category, Min and Yih (2010) modeled the problem of allocating surgeries to the blocks reserved for different surgery specialties. They formulated the problem as a two-stage stochastic mixed-integer program and used a sample average approximation method to solve the problem. Their model also considers the availability of intensive care unit (ICU) beds during the block assignment phase. The length of stay in an ICU bed and surgery durations are the stochastic parameters in their model. The objective function minimizes patient priority-based waiting costs and OR overtime costs.

Lamiri et al. (2008a) solved the problem of assigning elective surgeries to periods over a planning horizon while considering the impact of uncertainty related to emergency case arrivals. They first modeled the problem as a stochastic combinatorial optimization problem and then provided a reformulation in the form of a sample average approximation problem. The authors considered expected overtime costs and patient-related costs as the performance measures. The surgery durations are assumed to be deterministic. Lamiri et al. (2008b) extended the model in another study (Lamiri et al. 2008a) by considering the allocation of surgeries to ORs. Lamiri et al. (2009) proposed several heuristics to solve the same problem in that study (Lamiri et al. 2008a) and compared their performance with the performance of a Monte Carlo optimization method. Hans et al. (2008) also solved a stochastic OR-to-day allocation problem, where the stochasticity exists because of the uncertainty of the surgery durations. Their objective was to minimize the planned slack time reserved in the ORs each day that could be used by surgeries running longer than expected. The authors considered the trade-off between the OR utilization and OR overtime. They found that the surgeries having similar duration variability should be clustered together and assigned to the same OR-day.

In the third category of articles, Gerchak et al. (1996) modeled a surgery planning problem as a stochastic dynamic program. The decision process in their study was as follows: each day new requests for elective and emergency surgeries arise. Surgeries were scheduled for the current or future days and previously scheduled surgeries may be cancelled. The objectives included maximizing the expected profit gained by scheduling elective cases and minimizing the expected overtime and surgery cancellation costs. Zonderland et al. (2010) also considered a dynamic decision process where the days were assigned to blocks of surgeries at the beginning of every week for a variety of urgency levels. The different urgency levels included elective surgeries as well as semi-urgent surgeries that must be scheduled within one or two weeks. Based on a Markov decision process model, the authors provided a planning guideline by taking the costs related to the OR idle time, OR overtime, and cancellation of elective surgeries into consideration.

Our work differs from the studies in the first and second category in the stochastic dynamic setting for planning the surgeries. The articles most similar to ours are those by Gerchak et al. (1996) and Zonderland et al. (2010), which also consider a dynamic decision-making process. This article differs in the following ways. First, Gerchak et al. (1996) allow surgeries to be scheduled to the same day that they are requested; however, this is not a very realistic representation of many surgery practices. Second, the surgery durations generated in their model are independent from each other and identically distributed. Third, they do not consider OR allocation decisions and other scheduling complexities included in our model.

This article also differs from Zonderland et al. (2010) in a number of other ways. First, those authors do not consider the assignment of individual surgeries to days, but rather reserve time slots for elective or semi-urgent surgeries each day. Thus, for example, they do not make a distinction between different types of elective surgeries. Furthermore, they make strict assumptions about the nature of uncertainty, including that surgery requests arise according to a Poisson process, and surgery durations are assumed to be exponentially distributed. In contrast, our study makes no special assumptions about the random model parameters, and our numerical results are based on real data from a large medical center.

3. Problem Description

The model formulated and discussed in the remainder of this article considers daily decisions for the dynamic allocation of surgeries to ORs over a finite planning horizon under uncertainty (see Figure 1). The problem is formulated as a multistage stochastic mixed-integer program (MSSMIP). At each stage—i.e., day—newly requested surgeries are scheduled for future days; furthermore, some previously scheduled surgeries may be cancelled and subsequently rescheduled. In addition to assigning each surgery a day, an available OR is also assigned.

At the beginning of each day, it is assumed that random durations for surgeries are observed for the current day. In other words, we assume that the duration of surgeries of the current day can be accurately estimated at the start of the day. After the final schedule is determined for each day, the cumulative duration of the surgeries assigned to the ORs, total amount of OR overtime, and cancellations for that day are determined. In the case of a surgery that is rescheduled to a future day, its duration will be updated at the beginning of that new surgery day.

Total expected OR overtime and waiting and cancellation costs are the performance measures considered. The amount of patient waiting for a surgery is equal to the number of days between the day surgery is performed and the day it is requested. Thus, our model favors scheduling surgeries earlier rather than later. To reduce overtime on a particular day, surgeries might be cancelled and rescheduled. However, it is preferable to limit cancellations because of surgery cancellation and waiting costs and the hardship on patients. To reflect this, the model includes a penalty per cancellation that represents the "cost" of cancellation. Once a surgery is cancelled, it is scheduled either to another day within the planning horizon, or the end of the planning horizon, implying it will be scheduled for beyond the planning horizon. Note that a surgery can be cancelled more than once over the planning horizon; however, we assume there exists a time window within which each surgery must be completed.

Following is a detailed description of the MSSMIP model.

Indices:

i: Surgery index
l, t, u: Day index
j: OR index
ω^t: Scenario realization at day t (i.e., surgeries requested at day t, and duration of surgeries scheduled on day t)
$\omega^{[t]} := (\omega^1, \ldots, \omega^t)$: History of scenario realization up to day t

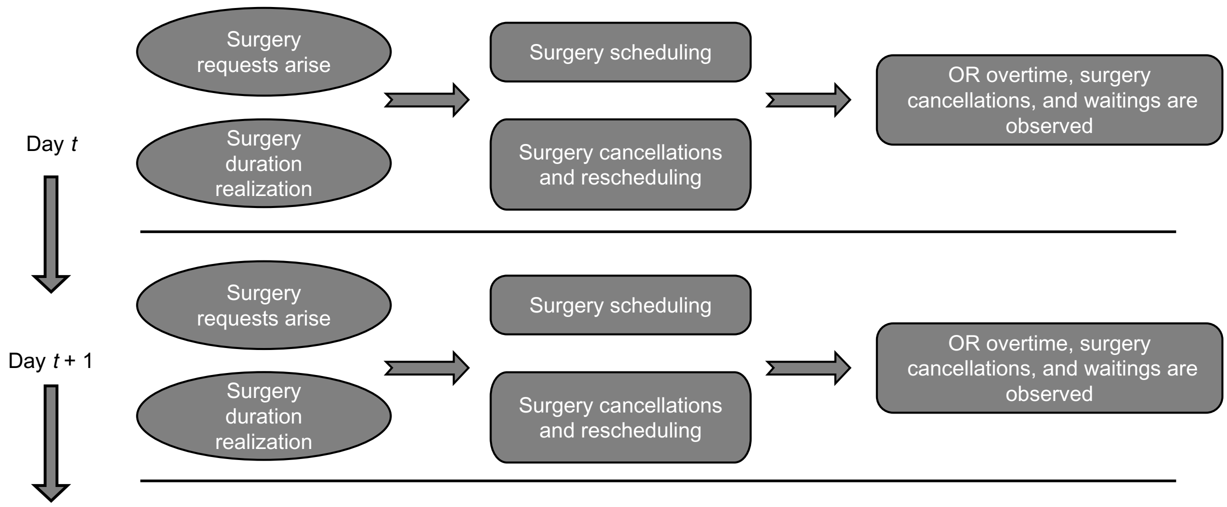

Figure 1 Decision-Making Process Throughout the Planning Period

Deterministic Parameters:

$$\lambda_{ij} = \begin{cases} 1 & \text{if there is no equipment constraint restricting the assignment of surgery } i \text{ to OR } j, \\ 0 & \text{otherwise} \end{cases}$$

$$p_{iu} = \begin{cases} 1 & \text{if surgery } i \text{ can be assigned to day } u, \\ 0 & \text{otherwise} \end{cases}$$

$g_i =$ Lead time (number of days between the earliest day the surgery can be assigned and the day the request arises) for scheduling surgery i

$h_i =$ Length of time window (number of days between the earliest day and the latest day that the surgery can be assigned) for scheduling surgery i

$$a_{ijt} = \begin{cases} 1 & \text{if surgery } i \text{ was already assigned to day } t \text{ and OR } j \text{ before the planning horizon starts,} \\ 0 & \text{otherwise} \end{cases}$$

$P_j^t =$ Capacity (in terms of minutes) of OR j at day t

$c^i =$ Cost per cancellation of surgery i

$l^i =$ Waiting cost per day for surgery i

$c^o =$ OR overtime cost per minute

$O =$ Number of ORs

$H =$ Length of planning horizon for scheduling surgeries

Random Parameters and Sets:

$s(\omega^t) =$ Set of surgeries requested at day t according to realization ω^t

$s(\omega^{[t]}) =$ Set of surgeries requested before and at day t under scenario $\omega^{[t]}$

$d_i^t(\omega^{[t]}) =$ Random duration of surgery i at day t under scenario $\omega^{[t]}$

$D(\omega^{[t]}) =$ Set of random durations of all surgeries requested before and at day t under scenario $\omega^{[t]}$

$\xi^{[t]} = (s(\omega^{[t]}), D(\omega^{[t]}))$: Set of realization history of random parameters up to day t

*t*th *Stage Decision Variables*:

$$x_{iju}^t(\omega^{[t]}) = \begin{cases} 1 & \text{if surgery } i \text{ is assigned to OR } j \text{ and day } u \text{ at day } t \text{ under scenario } \omega^{[t]}, \\ 0 & \text{otherwise} \end{cases}$$

$$\sigma_{ij}^t(\omega^{[t]}) = \begin{cases} 1 & \text{if surgery } i \text{ from OR } j \text{ is cancelled at day } t \text{ under scenario } \omega^{[t]}, \\ 0 & \text{otherwise} \end{cases}$$

$o_j^t(\omega^{[t]}) =$ Overtime for OR j observed at day t under scenario $\omega^{[t]}$

$\mathbf{y}^t(\omega^{[t]}) =$ Vector of values for all decision variables defined for and before day t under scenario $\omega^{[t]}$

The constraint set in the formulation of our problem has a block diagonal structure. There are H blocks of constraints as well as nonnegativity and binary restrictions on the decision variables. Each of the first $H-1$ blocks contains seven types of constraints; the last block has only four types. At day H, there are no new surgery requests since this is the last day of the planning horizon. For the last day, we have the following formulation for the recourse function:

$$Q^H(\mathbf{y}^{H-1}(\omega^{[H-1]}), \xi^{[H]}(\omega^{[H]})) = \min \sum_{j=1}^{O} \left(\sum_{i \in s(\omega^{[H-1]})} c^i \sigma_{ij}^H(\omega^{[H]}) + c^o o_j^H(\omega^{[H]}) + l^i x_{ijH+1}^H(\omega^{[H]}) \right) \tag{1}$$

$$\text{s.t.}\quad \sigma_{ij}^H(\omega^{[H]}) - a_{ijH} - \sum_{l=1}^{H-1} x_{ijH}^l(\omega^{[H]}) \le 0 \quad \forall\, i \in s(\omega^{[H-1]}),\, j \tag{2}$$

$$\sum_{j=1}^{O} \sigma_{ij}^H(\omega^{[H]}) - p_{iH+1} \le 0 \quad \forall\, i \in s(\omega^{[H-1]}) \tag{3}$$

$$\sum_{i\in s(\omega^{[H-1]})} d_i^H(\omega^{[H]})\left(a_{ijH} + \sum_{l=1}^{H-1} x_{ijH}^l(\omega^{[H]}) - \sigma_{ij}^H(\omega^{[H]})\right) - o_j^H(\omega^{[H]}) \le P_j^H \quad \forall\, j \tag{4}$$

$$x_{ijH+1}^H(\omega^{[H]}),\, \sigma_{ij}^H(\omega^{[H]}) \in \{0,1\} \quad \forall\, i,j \tag{5}$$

$$o_j^H(\omega^{[H]}) \ge 0 \quad \forall\, j. \tag{6}$$

The objective function minimizes the waiting, cancellation, and overtime costs at stage H. Constraints (2) and (3) require that a surgery in an OR can only be cancelled on day H if it was previously assigned to this day and OR, and if it is possible to reschedule the surgery to a day beyond the planning horizon. Constraint (4) determines the overtime for each OR. Constraints (5) and (6) define the nonnegativity and integrality restrictions.

Assuming interstage independence (i.e., $\xi^{[t+1]}$ is independent of the past outcomes) and letting

$$\mathcal{Q}^{t+1}(\mathbf{y}^t(\omega^{[t]})) = E_{\xi^{[t+1]}}[Q^{t+1}(\mathbf{y}^t(\omega^{[t]}),\quad \xi^{[t+1]}(\omega^{[t+1]}))],$$

we obtain the following recursion for $Q^t(\mathbf{y}^{t-1}(\omega^{[t-1]}), \xi^{[t]}(\omega^{[t]}))$ defined for $t = 2, \ldots, H-1$:

$$\min \sum_{j=1}^{O}\left(\sum_{i\in s(\omega^{[t-1]})} c^i \sigma_{ij}^t(\omega^{[t]}) + c^o o_j^t(\omega^{[t]}) + \sum_{i\in s(\omega^{[t]})}\sum_{u=t+1}^{H+1} l^i(u-t)x_{iju}^t(\omega^{[t]})\right) + \mathcal{Q}^{t+1}(\mathbf{y}^t(\omega^{[t]})) \tag{7}$$

$$\text{s.t.}\quad \sum_{u=t+1}^{H+1}\sum_{j=1}^{O} x_{iju}^t(\omega^{[t]}) = 1 \quad \forall\, i \in s(\omega^{[t]}) \tag{8}$$

$$\sum_{u=t+1}^{H+1}\sum_{j=1}^{O} x_{iju}^t(\omega^{[t]}) = \sum_{j=1}^{O} \sigma_{ij}^t(\omega^{[t]}) \quad \forall\, i \in s(\omega^{[t-1]}) \tag{9}$$

$$x_{iju}^t(\omega^{[t]}) \le \lambda_{ij} p_{iu} \quad \forall\, i \in s(\omega^{[t]}),\, j,\, u = t+1, \ldots, H+1 \tag{10}$$

$$\sigma_{ij}^t(\omega^{[t]}) - a_{ijt} - \sum_{l=1}^{t-1} x_{ijt}^l(\omega^{[t]}) \le 0 \quad \forall\, i \in s(\omega^{[t-1]}),\, j \tag{11}$$

$$\sum_{i\in s(\omega^{[t-1]})} d_i^t(\omega^{[t]})\left(a_{ijt} + \sum_{l=1}^{t-1} x_{ijt}^l(\omega^{[t]}) - \sigma_{ij}^t(\omega^{[t]})\right) - o_j^t(\omega^{[t]}) \le P_j^t \quad \forall\, j \tag{12}$$

$$x_{iju}^t(\omega^{[t]}),\, \sigma_{ij}^t(\omega^{[t]}) \in \{0,1\} \quad \forall\, i;\, \forall\, j;\, t = 1, \ldots, H;\, u = 2, \ldots, H+1 \tag{13}$$

$$o_j^t(\omega^{[t]}) \ge 0 \quad \forall\, j;\, t = 1, \ldots, H. \tag{14}$$

Constraint (8) ensures that requests for surgery on day t must be assigned to an OR in one of the days after day t. Constraint (9) enforces the assignment of a cancelled surgery to a future day and OR. Constraint (10) imposes restrictions on the particular day and OR to which a surgery may be assigned. When a request arises for a surgery, it must be scheduled within the allowable time window (h_i) for performing the surgery and at least g_i stages into the future. Note that parameters g_i and h_i are used implicitly to calculate parameters p_{iu} in constraint (10). A restriction on the assignment of a surgery to an OR might also exist, defined by the same constraint (e.g., if a particular type of OR is required for the surgery). Constraint (11) (which is equivalent to (2)) ensures that only surgeries that are scheduled for day t can be cancelled on day t. Constraint (12) determines the amount of OR overtime on day t. Constraints (13) and (14) define the nonnegativity and integrality restrictions on the variables.

This MSSMIP model is NP-hard. This follows from the fact that an instance of this problem, where the model has only one scenario, corresponds to the well-known bin packing problem.

4. Solution Methodology

The above model can be approximately solved using the progressive hedging algorithm (PHA) proposed by Rockafellar and Wets (1991). The PHA proceeds by applying scenario decomposition iteratively, solving the resulting individual scenario subproblems, and finally aggregating individual scenario solutions. Although the PHA is guaranteed to converge to a global optimal solution asymptotically in the convex case (Rockafellar and Wets 1991), it may converge to only a local optimal solution in this case, because the problem is nonconvex, because of the binary decision variables.

Many authors of PHA-based studies have analyzed the algorithm and proposed ways to improve its performance by taking advantage of the special structure of the problem of interest (Mulvey and Vladimirou 1991a, b; Wallace and Helgason 1991; Hvattum and Lokketangen 2009; Watson and Woodruff 2011; Crainic et al. 2011). The PHA has also been applied in several application areas (i.e., see Mulvey and Vladimirou 1992 for a financial planning application; Helgason and Wallace 1991 for fisheries management application; Santos et al. 2009 for hydrothermal systems operation planning application). The reader is referred to Wallace and Helgason (1991) and Watson and Woodruff (2011) for detailed discussions about the algorithm implementation. Background on our implementation of the PHA is given in §4.2.

4.1. Problem Reformulation

To apply PHA, we begin by reformulating the model to put it in the standard form appropriate for scenario decomposition. In the new formulation, which we refer to as the PHA deterministic equivalent model (PHA-DEM), a new parameter η that represents a sequence of consecutive scenarios aggregated over all days (i.e., $\omega^{[1]}, \omega^{[2]}, \ldots, \omega^{[H]}$) is defined and introduced. Figure 2 illustrates how the reformulation impacts the scenario tree. Figures 2(a) and 2(b) compare scenario trees for the MSSMIP and PHA-DEM, respectively. Each oval node in the scenario tree represents a particular scenario realization, ω^t, at a particular stage t. The accumulation of oval nodes until stage t (i.e., $\omega^1, \omega^2, \ldots, \omega^t$) defines a particular scenario at day t (i.e., $\omega^{[t]}$). The circle nodes within the oval nodes indicate the surgeries requested on a particular day under the realization that the oval node represents. Note that, for simplicity, the example in Figure 2 assumes that the uncertainty is based only on the surgery requests (i.e., the surgery durations are deterministic).

Figure 2(a) illustrates that $\omega^{[4]}$ varies based on the scenario represented by $\omega^{[3]}$. The same relation exists also for ($\omega^{[1]}$, $\omega^{[2]}$) and ($\omega^{[2]}$, $\omega^{[3]}$). In contrast, Figure 2(b) illustrates an alternative representation of the scenario tree given in Figure 2(a), where the individual scenarios observed in the particular stages are aggregated over all days to form three scenario sequences, $\eta = 1, 2, 3$. However, the above redefinition of the scenario tree is not permissible since the solutions found might not be feasible for the overall problem, because they imply decisions that anticipate future uncertain events. Therefore, *nonanticipativity constraints* are required in the PHA-DEM. These constraints enforce the following property: If two scenario sequences (i.e., $\eta = a, b$) share the same history up to day t, the surgery schedules created progressively over the planning period should always have the same content until day t. In other words, if a decision is given for a surgery at some day l, where $l \le t$ under scenario sequence a, the same decision holds under scenario sequence b.

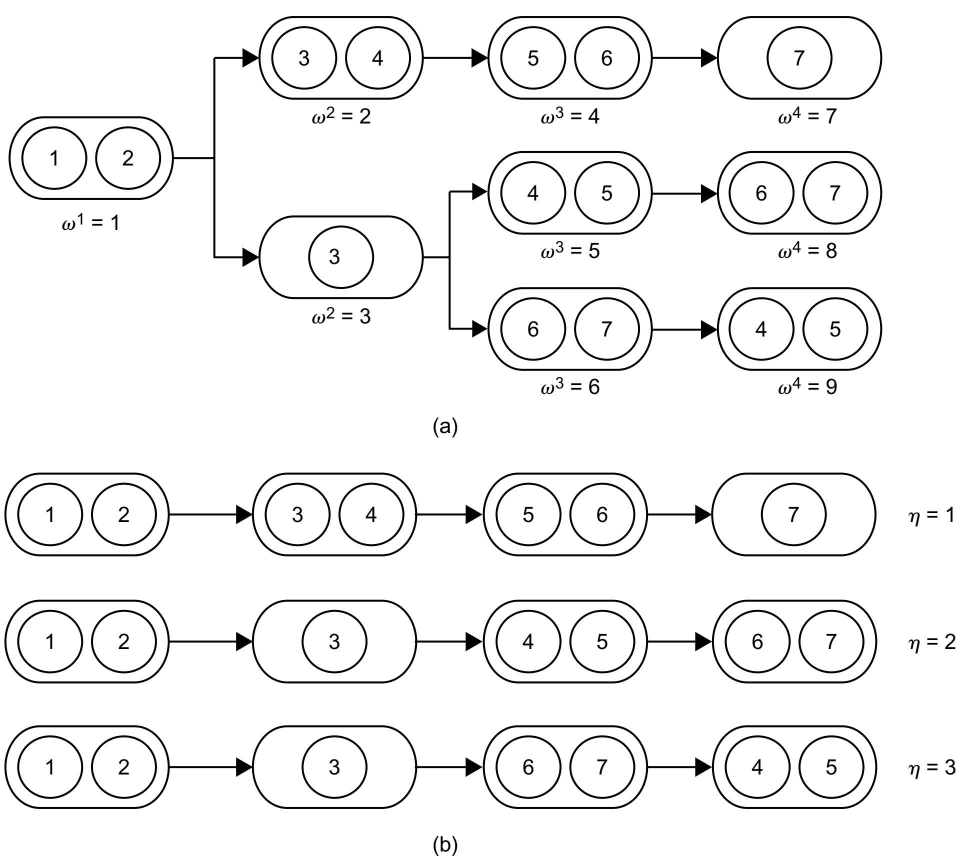

Figure 2 Reformulation Impact on the Scenario Tree

Note. (a) A scenario tree example illustrating the surgeries requested each day over a four-day planning period. (b) The example from (a) shown in terms of individual scenario sequences.

Following is the additional notation used to formulate the PHA-DEM.

Additional Indices:

η: Scenario sequence index

$B(\eta, t)$: Scenario bundle index of the surgeries considered for scheduling at stage t under scenario sequence η

Additional Parameters:

Z: Number of scenario sequences

N: Number of surgeries requested in a sample of scenario tree

$$s_{i\eta}^t = \begin{cases} 1 & \text{if surgery } i \text{ is requested at day } t \text{ under scenario sequence } \eta, \\ 0 & \text{otherwise} \end{cases}$$

$d_{i\eta}^t =$ Duration of surgery i at day t under scenario sequence η

$\Pr_\eta =$ Probability of the occurrence of scenario sequence η

Revised Decision Variables:

$$x_{i\eta ju}^t = \begin{cases} 1 & \text{if surgery } i \text{ is assigned to OR } j \text{ and day } u \text{ at day } t \text{ under scenario sequence } \eta, \\ 0 & \text{otherwise} \end{cases}$$

$$\sigma_{i\eta j}^t = \begin{cases} 1 & \text{if surgery } i \text{ from OR } j \text{ is cancelled at day } t \text{ under scenario sequence } \eta, \\ 0 & \text{otherwise} \end{cases}$$

$o_{\eta j}^t =$ Resulting overtime amount for OR j on day t under scenario sequence η

Additional Decision Variables:

$$x_{iju}^{B(\eta,t)} = \begin{cases} 1 & \text{if surgery } i \text{ is assigned to day } u \text{ and OR } j \text{ at all day-scenario sequence combinations in the bundle, } B(\eta, t), \text{ that day } t\text{-scenario } \eta \text{ belongs to,} \\ 0 & \text{otherwise.} \end{cases}$$

The nonanticipativity constraints are also referred to as *bundle constraints*. If the scenario sequences a and b share the same history up to day t, then this indicates that they also share the same *scenario bundle* on day t: $B(a, t) = B(b, t)$. Thus, the scheduling decisions given on this day are the same among all scenario sequences placed in the same scenario bundle.

Figure 3 illustrates the scenario bundle concept using the example given in Figure 2. The rectangles covering the oval nodes represent the particular scenario bundles that exist in the example. Because all three scenario sequences have the same realization (e.g., $\omega^1 = 1$) at day 1, $\eta = 1, 2, 3$ share the same bundle on this day and $B(1, 1) = B(2, 1) = B(3, 1) = 1$. The second day also contains one scenario bundle, because $\eta = 2$ and $\eta = 3$ share the same history by day 2.

We now show how decisions are synchronized using the bundle constraints. First, recall that the model includes decisions for two different cases: (i) a request arises for a new surgery; (ii) one of the surgeries is cancelled. At day 1, for all η, surgeries 1 and 2 are scheduled into the future, corresponding to case (i). One can enforce nonanticipativity at the first stage using the following constraints:

$$x_{i1ju}^1 = x_{i2ju}^1 = x_{i3ju}^1 \quad \forall\, j, u = 2, 3, 4, 5 \text{ and } i = 1, 2.$$

Similarly, nonanticipativity related to surgery 3 at the second stage can be generated using:

$$x_{32ju}^2 = x_{33ju}^2 \quad \forall\, j, u = 3, 4, 5.$$

For case (ii), the decisions to reschedule cancelled surgeries under scenario sequences $\eta = 2, 3$ are bundled as follows:

$$x_{i2ju}^2 = x_{i3ju}^2 \quad \forall\, j, u = 3, 4, 5 \text{ and } i = 1, 2.$$

To facilitate the generation of a separable stochastic program, a new decision variable, called the *consensus*

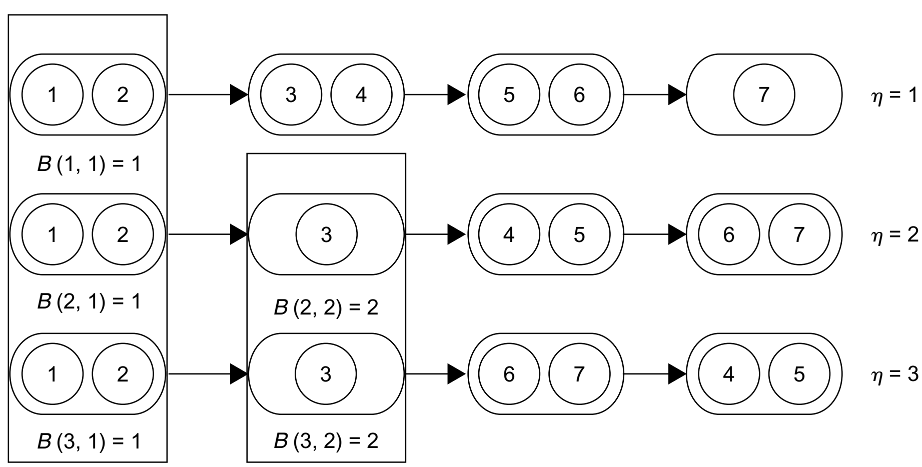

Figure 3 **Representation of Scenario Bundles by Rectangles Covering the Scenario Realizations at a Particular Day**

variable, $x_{iju}^{B(\eta,t)}$, is defined. The PHA-DEM is formulated as follows:

$$\min \sum_{\eta=1}^{Z}\Pr_\eta\left(\sum_{t=1}^{H}\sum_{j=1}^{O}\left(c^o o_{\eta j}^t+\sum_{i=1}^{N}c^i\sigma_{i\eta j}^t+\sum_{i=1}^{N}\sum_{u=t+1}^{H+1}l^i(u-t)x_{i\eta ju}^t\right)\right) \quad (15)$$

$$\text{s.t.}\quad x_{i\eta ju}^t=x_{iju}^{B(\eta,t)} \quad \forall i,\eta,j,t,u>t \quad (16)$$

$$\sum_{u=t+1}^{H+1}\sum_{j=1}^{O}x_{i\eta ju}^t=s_{i\eta}^t+\sum_{j=1}^{O}\sigma_{i\eta j}^t \quad \forall i,\eta,t \quad (17)$$

$$x_{i\eta ju}^t\le\lambda_{ij}p_{iu} \quad \forall i,\eta,j,t,u>t \quad (18)$$

$$\sigma_{i\eta j}^t-a_{ijt}-\sum_{l=1}^{t-1}x_{i\eta jt}^l\le 0 \quad \forall i,\eta,j,t \quad (19)$$

$$\sum_{i=1}^{N}d_{i\eta}^t\left(a_{ijt}+\sum_{l=1}^{t-1}x_{i\eta jt}^l-\sigma_{i\eta j}^t\right)-o_{\eta j}^t\le P_j^t \quad \forall j,\eta,t \quad (20)$$

$$x_{i\eta ju}^t,x_{i\eta ju}^{B(\eta,t)},\sigma_{i\eta j}^t\in\{0,1\} \quad o_{\eta j}^t\ge 0 \quad \forall i,\eta,j,t,u>t. \quad (21)$$

The objective function (15) is the expectation of the total scenario costs over all scenarios. The total scenario cost is weighted by the probability associated with scenario $\Pr_\eta$. The total cost for a scenario includes the total OR overtime cost and surgery cancellation and waiting costs over all days.

Constraint (16) is the bundle constraint. Constraints (17)–(21) have the same meaning as (8)–(14) in MSSMIP but use one less constraint. The number of constraints is one less because we are now able to define the parameter, $s_{i\eta}^t$, that indicates whether a surgery is requested on a particular day. This parameter allows formulating both scheduling and rescheduling decisions in one constraint instead of two. Constraint (17) sets the conditions to be satisfied to give a scheduling decision at a particular day. Constraint (18) defines the allowable days and ORs for the assignment of a particular surgery. Constraint (19) ensures that the cancellation decision for a surgery on a particular day can be made if the surgery was assigned to that day in the past. Constraint (20) measures overtime values for each OR, each day. Constraint (21) defines the nonnegativity and binary restrictions on the decision variables.

Using the PHA-DEM formulation, an augmented Lagrangian relaxation technique is applied by dualizing the bundle constraint. The relaxed formulation still includes the constraints (17)–(21) in the constraint set. However, the objective function (15) is now

$$\min\sum_{\eta=1}^{Z}\Pr_\eta\Bigg(\sum_{t=1}^{H}\sum_{j=1}^{O}\left(c^o o_{\eta j}^t+\sum_{i=1}^{N}c^i\sigma_{i\eta j}^t+\sum_{i=1}^{N}\sum_{u=t+1}^{H+1}l^i(u-t)x_{i\eta ju}^t\right)+\sum_{i=1}^{N}\sum_{t=1}^{H}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}\mu_{i\eta ju}^t(x_{i\eta ju}^t-x_{iju}^{B(\eta,t)})+\frac{\rho}{2}\sum_{i=1}^{N}\sum_{t=1}^{H}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}\|x_{i\eta ju}^t-x_{iju}^{B(\eta,t)}\|^2\Bigg), \quad (22)$$

where $\mu_{i\eta ju}^t$, $\forall i,\eta,t,j,u$ denote the Lagrangian multipliers; ρ is the penalty parameter; and $\|\cdot\|$ is the ordinary Euclidean norm. The additional components in the function (22) penalize the violation of the bundle constraint. Since $x_{i\eta ju}^t, x_{iju}^{B(\eta,t)}\in\{0,1\}$, the penalty component in (22) is rewritten as follows:

$$\|x_{i\eta ju}^t-x_{iju}^{B(\eta,t)}\|^2=x_{i\eta ju}^t-2x_{i\eta ju}^t x_{iju}^{B(\eta,t)}+x_{iju}^{B(\eta,t)}. \quad (23)$$

To make the deterministic equivalent formulation scenario separable requires fixing the consensus variable $x_{iju}^{B(\eta,t)}$. Using a *proximal point method* (Rockafellar 1976), this value can be estimated using the following weighted sum calculation:

$$\hat{x}_{iju}^{B(\eta,t)}=\sum_{\eta\in B(\eta,t)}\frac{\Pr_\eta}{\sum_{\eta\in B(\eta,t)}\Pr_\eta}x_{i\eta ju}^t \quad \forall i,\eta,t,j,u. \quad (24)$$

Note that (23) no longer contains a quadratic term after replacing $x_{iju}^{B(\eta,t)}$ with $\hat{x}_{iju}^{B(\eta,t)}$, which facilitates the solution of the subproblems following the scenario decomposition.

Equation (24) determines the weighted sum of the individual scheduling decision variables within a decision bundle. The weights are set by normalizing the probability of the scenario associated with a decision variable. The objective function of the resulting separable formulation, which we refer to as the separable deterministic equivalent model (PHA-SDEM), is

$$\min \sum_{\eta=1}^{Z}\Pr_\eta\Bigg(\sum_{t=1}^{H}\sum_{j=1}^{O}\left(c^o o_{\eta j}^t+\sum_{i=1}^{N}c^i\sigma_{i\eta j}^t+\sum_{i=1}^{N}\sum_{u=t+1}^{H+1}l^i(u-t)x_{i\eta ju}^t\right)+\sum_{i=1}^{N}\sum_{t=1}^{H}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}\mu_{i\eta ju}^t(x_{i\eta ju}^t-\hat{x}_{iju}^{B(\eta,t)})+\frac{\rho}{2}\sum_{i=1}^{N}\sum_{t=1}^{H}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}(x_{i\eta ju}^t-2x_{i\eta ju}^t\hat{x}_{iju}^{B(\eta,t)})\Bigg) \quad (25)$$

s.t. (17)–(21).

Note that the last term of (23) is ignored in (25), because it is fixed.

The estimate of $\hat{x}_{iju}^{B(\eta,t)}$ is referred to as the *consensus parameter*. The consensus parameter is an estimation of the implementable solution at a given iteration of the PHA. If this solution is also feasible in PHA-SDEM, then it is called an *admissible solution*. The goal of the PHA is to identify a good solution (ideally the optimal solution) among all admissible solutions.

The mixed-integer programming formulation for a particular scenario subproblem model (PHA-SSM) is obtained by decomposing PHA-SDEM into scenario subproblems. Hence, the constraint set for PHA-SSM is a subset of that of PHA-SDEM, which should be satisfied for all scenarios rather than only for one scenario.

4.2. Progressive Hedging Algorithm

In this section we describe our implementation of PHA. Let ρ^0 denote the initial value of the penalty parameter, k the index for the iteration number of the PHA, $\mu^{t(k)}_{i\eta ju}$ $\forall i, \eta, t, j, u$ the Lagrangian multipliers, and $\rho^{(k)}$ the penalty parameter at iteration k. Then the general steps of the PHA are stated as follows:

PHA

Set algorithm terminates = false, $k = 1$, $\rho^{(k)} = \rho^0$, $\mu^{t(k)}_{i\eta ju} = 0$ $\forall i, \eta, t, j, u$
while algorithm terminates = false
  for $\eta = 1$ to Z
    Solve the PHA-SSM to obtain $x^{t(k)}_{i\eta ju}$ $\forall i, \eta, t, j, u$
  end for
  Calculate the consensus parameter: $\hat{x}^{B(\eta,t)}_{iju}$ $\forall i, \eta, t, j, u$
  if $k > 1$
    Update the penalty parameter according to the following scheme: $\rho^{(k+1)} = \alpha\rho^{(k)}$, where $\alpha > 0$
  end if
  Update the Lagrangian multipliers according to the following scheme: $\mu^{t(k+1)}_{i\eta ju} = \mu^{t(k)}_{i\eta ju} + \rho^{(k)}(x^{t(k)}_{i\eta ju} - \hat{x}^{B(\eta,t)}_{iju})$
  if $x^{t(k)}_{i\eta ju} = \hat{x}^{B(\eta,t)}_{iju}$ $\forall i, \eta, t, j, u$
    Set algorithm terminates = true
  end if
  else
    Set $k = k + 1$
  end else
end while

4.3. Enhanced Progressive Hedging Algorithm

In our implementation, we took advantage of the special structure of the model formulation to accelerate the computational performance of the PHA and improve the quality of the PHA solutions. We refer to this algorithm as the *enhanced progressive hedging algorithm* (EPHA). The degree of violation of the bundle constraints and decisions taken by the majority of the variables in the decision bundles motivates the Lagrangian multiplier update method. We also analyze if a penalty update method utilizing the information about the convergence pattern of the primal and dual variables may enhance the solutions.

In the PHA literature, many other studies propose enhancements on the PHA based on the special structures of the models. We first present a brief review of enhancements proposed for various problems. Then we discuss the methods of our EPHA.

4.3.1. PHA Enhancements. Mulvey and Vladimirou (1991a, b) discussed the trade-off between the selection of high and low values for the penalty parameters and the impact of the problem structure into this selection. They also discussed the benefits of the dynamic penalty adjustment methods. Helgason and Wallace (1991) and Listes and Dekker (2005) discussed the sensitivity of the convergence of the PHA to the choice of penalty parameter. Hvattum and Lokketangen (2009) proposed a method to set a direction of improvement while updating the penalty parameters. They tested the case where there exist parameters for individual nonanticipativity constraints in the model. Watson and Woodruff (2011) also proposed methods to set the penalty parameters for individual nonanticipativity constraints for a class of resource-allocation problems.

In our experiments, we observe that updating penalty parameters based on the information on the convergence pattern in the primal and dual space did not achieve significantly better results than keeping the penalty parameter constant. Our experiments suggest that the initial values of the Lagrangian multipliers have significant impact on the quality of the final solution. We selected the initial values after assessing the trade-off between the marginal improvement in solution quality and additional computational time needed as a result of a variation in the values. Mulvey and Vladimirou (1991a) and Santos et al. (2009) also discussed the importance of the initial estimates for the Lagrangian multipliers and tested warm start methods, including simple heuristics to find reasonable initial values.

It is well known that in the nonconvex case, the PHA is not guaranteed to converge (Takriti and Birge 2000). Watson and Woodruff (2011) defined some techniques to detect nonconvergence in the form of cyclic behaviors. Whenever they detect a cycle for a variable, they fix the variable value using a simple rule (the largest value of the variable across scenarios is selected). In our case, the Lagrangian multiplier update method prevents cycling. Since the Lagrangian multipliers are defined only for binary variables, the method aims to favor one feasible value over the other primarily based on the selection in the majority of the subproblems. When the majority is not achieved, the waiting and cancellation costs are also considered.

Because of the typically large number of subproblems to be solved following the scenario decomposition at each PHA iteration, computational efficiency in subproblems is important. Furthermore, it has been shown that the PHA is often a reasonable heuristic to use if there exists an efficient algorithm to solve the subproblems of a very large-scale stochastic mixed-integer problem (Watson and Woodruff 2011). Takriti et al. (1996) developed methods to solve the subproblems of their multistage stochastic production planning problem. Barro and Canestrelli (2005) further decomposed the subproblems of a dynamic portfolio management problem into stages to solve those efficiently. An important reason necessitating the implementation of an efficient solution method on the subproblems is that each subproblem has a quadratic objective function because of the penalty component. Haugen et al. (2001) relaxed the quadratic term in the subproblem objective function and applied a dynamic programming approach to find an optimal solution for the relaxed subproblems. Listes and Dekker (2005) solved the linear relaxation of the subproblems of a robust airline fleet composition problem, which contained integer variables, and used a simple rounding procedure to find a feasible solution for the overall problem. In our problem formulation, the quadratic component is linearized since the corresponding component includes binary variables.

4.3.2. Enhanced Progressive Hedging Algorithm Implementation. In this section, we present our penalty update and Lagrangian multiplier update methods and EPHA termination criterion.

We set a constant value for the penalty parameter after conducting some experimental analysis. The experimental analysis was based on the observation of the trade-off between fast convergence to a suboptimal solution (when ρ is too large) and slow convergence to a near optimal solution in the primal feasible space (when ρ is too low).

Next, the method proposed in Hvattum and Lokketangen (2009) was used to compare the convergence rate at iterations k and $k-1$, increasing ρ if it appears that the convergence rate is decreasing. The parameter ρ is decreased if the current status is closer to consensus among variables at iteration $k-1$ than at iteration k. Let $\Delta_D^{(k)}$ and $\Delta_P^{(k)}$ be indicators of the convergence rates in the dual space and in the primal space, respectively. Let b index a unique bundle among the ones represented by all $B(\eta, t)$s, and B represent the total number of unique bundles. Then Equations (26)–(28) define the penalty update method as follows:

$$\Delta_P^{(k)} = \sum_{i=1}^{N}\sum_{b=1}^{B}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}(\hat{x}_{iju}^{b(k)} - \hat{x}_{iju}^{b(k-1)})^2 \tag{26}$$

$$\Delta_D^{(k)} = \sum_{i=1}^{N}\sum_{\eta=1}^{Z}\sum_{t=1}^{H}\sum_{j=1}^{O}\sum_{u=t+1}^{H+1}(x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)})^2 \tag{27}$$

$$\rho^{(k+1)} = \begin{cases} \delta_D \rho^{(k)} & \text{if } \Delta_D^{(k)} - \Delta_D^{(k-1)} > 0 \\ \dfrac{1}{\delta_P}\rho^{(k)} & \text{else if } \Delta_P^{(k)} - \Delta_P^{(k-1)} > 0, \end{cases} \tag{28}$$

where $\delta_P > 1$ and $\delta_D > 1$ in (28) are fixed multipliers.

We use a Lagrangian multiplier update method that ensures convergence of the algorithm. The method aims to achieve convergence of the consensus parameter value to one of the two feasible values: 0 or 1 (see Crainic et al. 2011 for a similar approach). The selection between these two values as the convergence point is made according to the majority of the variable values in a bundle. We define a threshold parameter called θ to help define the majority condition. Once the value of the consensus parameter is greater than θ, the majority is assumed to be achieved. In the case that no value is favored by the majority (i.e., consensus parameter value is equal to θ), the value to which the consensus parameter should converge is determined according to the cancellation and waiting costs.

The approach is based on the following observation. If $\hat{x}_{iju}^{B(\eta,t)(k)}$ is greater than θ, then this indicates that the majority of the scenario subproblem solutions within the associated bundle dictate the assignment of surgery i to OR j on day u. Our method decreases the values of the Lagrangian multipliers in the subproblems where the relevant variable value is 0. The expectation here is that the surgery is assigned to the same day and OR in the following iterations. On the other hand, if $\hat{x}_{iju}^{B(\eta,t)(k)}$ is less than θ, this shows that surgery i is not assigned to OR j on day u in the majority of the subproblems. Then the Lagrangian multiplier values are increased in the subproblems where the relevant variable is equal to 1. If $\hat{x}_{iju}^{B(\eta,t)(k)}$ is equal to θ, this means no particular day-OR couple is favored for the assignment of surgery i among subproblems. If the same day is selected in all subproblems, then any OR is favored for the assignment. Otherwise, the day to be favored is selected according to the cancellation and waiting costs. If the cancellation cost is greater than the waiting cost, then the latest feasible day is preferred. Therefore, the values of the Lagrangian multipliers are increased for the relevant variables that assign surgery i to the earlier days. If the waiting cost is greater than the cancellation cost, then the earliest feasible day is favored to reduce waiting. This requires an increase in the values of the Lagrangian multipliers for the relevant variables that assign surgery i to the later days. Updates are computed as follows:

$$\mu_{i\eta ju}^{t(k+1)} = \begin{cases} \mu_{i\eta ju}^{t(k)} + \rho^{(k)}(x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)}) \\ \quad \text{if } \hat{x}_{iju}^{B(\eta,t)(k)} < \theta;\, x_{i\eta ju}^{t(k)} = 1 \\ \mu_{i\eta ju}^{t(k)} - \rho^{(k)}|(x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)})| \\ \quad \text{if } \hat{x}_{iju}^{B(\eta,t)(k)} > \theta;\, x_{i\eta ju}^{t(k)} = 0 \\ \mu_{i\eta ju}^{t(k)} + \rho^{(k)}(x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)}) \\ \quad \text{if } \hat{x}_{iju}^{B(\eta,t)(k)} = \theta;\, x_{i\eta ju}^{t(k)} = 1 \text{ and } c^i \geq l_i; \\ \quad\quad u \neq \max\{u\colon x_{i\eta ju}^{t(k)} = 1, (\eta,t) \in B(\eta,t)\} \\ \mu_{i\eta ju}^{t(k)} + \rho^{(k)}(x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)}) \\ \quad \text{if } \hat{x}_{iju}^{B(\eta,t)(k)} = \theta;\, x_{i\eta ju}^{t(k)} = 1 \text{ and } c^i < l_i; \\ \quad\quad u \neq \min\{u\colon x_{i\eta ju}^{t(k)} = 1, (\eta,t) \in B(\eta,t)\} \\ \mu_{i\eta ju}^{t(k)} \quad \text{otherwise.} \end{cases} \tag{29}$$

4.3.2.1. *Termination Criteria.* The EPHA terminates when the following condition is satisfied:

$$\sum_{\eta=1}^{Z} \Pr_\eta \sum_{i=1}^{N} \sum_{t=1}^{H} \sum_{j=1}^{O} \sum_{u=t+1}^{H+1} |x_{i\eta ju}^{t(k)} - \hat{x}_{iju}^{B(\eta,t)(k)}| \leq \epsilon. \tag{30}$$

This can be interpreted as a measure of the bundle constraint violation being sufficiently small.

It is possible that the objective function coefficients of a subproblem may not change from one iteration to the next because of the method we propose for updating Lagrangian multipliers. The Lagrangian multiplier for a decision variable is not updated at an iteration in case the value of the variable is equal to the value taken by the sufficient majority of the variables in the bundle. Therefore, the objective function coefficients of a subproblem may remain the same if none of its variables require an update for the Lagrangian multipliers. We detect those subproblems at each iteration to minimize the number of times that the subproblem solution routine is called.

5. Experimental Study

We tested our methods using data from an outpatient procedure center at Mayo Clinic from the year 2006 for 4,034 patients (Gul et al. 2011). We generated scenario trees representing arrivals of surgery requests over a planning period. Each problem instance is based on a particular service including urology, ophthalmology, pain medicine, and oral maxillofacial. Surgeries of a service are grouped into acuity levels. Urology and pain medicine have five acuity levels, and ophthalmology and oral maxillofacial have two. The probability distribution for the duration of a surgery is based on its acuity level.

We conducted experiments with moderate-size test cases that can be solved to optimality and compared the optimal solutions to those of the EPHA. We also conducted experiments to estimate the value of the stochastic solution (VSS) by comparing solutions obtained with the EPHA and a deterministic heuristic for a set of large-size test cases. We also used large-size test cases for our evaluation of the EPHA within a rolling horizon framework. In our test cases, two ORs are open every day for each surgical service. Each OR operates for eight hours daily. The moderate-size cases included, on average, 27 surgeries to be scheduled during a six-day planning period. These cases represent the surgery scheduling process during a typical week (i.e., five days). The surgeries that cannot be scheduled for the current week are assigned to a dummy day at the end of the planning period. The large-size test cases considered 50 surgeries, on average, to be scheduled during an 11-day planning period. These cases represent a biweekly surgery scheduling process (i.e., 10 days) with a dummy period included for the surgeries that cannot be performed within the two-week period.

Each experiment was performed on an instance set that consisted of 20 different scenario trees of the same size. Each tree consisted of 1,024 scenarios. We used 56 different moderate-size instance sets of 20 instances for performing sensitivity analysis. We used a single large-size instance set to estimate the value of the stochastic solution, and a second one to test the EPHA within a rolling horizon framework. We report the average and worst case performance for each experiment. The EPHA algorithm was implemented in Microsoft Visual C++ 2008 using CPLEX 12 Concert Technology. The experiments were conducted on an Intel Core i5 PC with processors running at 2.27 GHz and 4GB memory under Windows XP.

5.1. Generating Problem Instances

We generated a scenario tree for each problem instance. There are two parameters that determine the size of the tree: (1) n_s, *number of stages* and (2) n_o, *number of different outcomes observed at each stage except the first stage*. The realization at stage 1 is assumed to be known, thus there is only one outcome at this stage. There exist $n_o^{n_s-1}$ scenario sequences in total in a scenario tree. The generated scenario outcome at stage t does not depend on the past outcomes (i.e., interstage independence is assumed). The interstage independence allows us to use the *common samples* approach (Chiralaksanakul and Morton 2004). The common samples approach indicates that the generated outcomes may exist more than once and the same number of times at a particular stage of a scenario tree. In particular, among the n_o^{t-1} outcomes at stage t, where $t > 1$, in our scenario tree, only n_o of them are unique.

A scenario outcome is characterized by the number of surgeries requested from each acuity level. For example, assume that we generate a scenario tree for the case where $n_s = 4$ and $n_o = 2$ for the surgical service of pain medicine. Let the number of surgery requests for acuity level i, where $i = 1, 2, 3, 4, 5$, be represented by $acu(i)$. The value of each $acu(i)$ is independently sampled according to the probability distribution fit for the surgery request frequency in the data. A sample tree with its $acu(i)$ set is illustrated in Figure 4. Each rectangle in the tree denotes a scenario outcome at the corresponding stage. The numbers in the rectangles represent a vector of $acu(i)$ values for $i = 1, 2, 3, 4, 5$.

We sampled individual surgery durations from independent distributions to obtain a sample for the sum of surgery durations. We generated these samples so the average sum of surgery durations over the samples and the standard deviation of the sum of durations are consistent with the range of values observed in practice (e.g., mean is nearly nine hours, and the standard deviation is either 78 or 234 minutes).

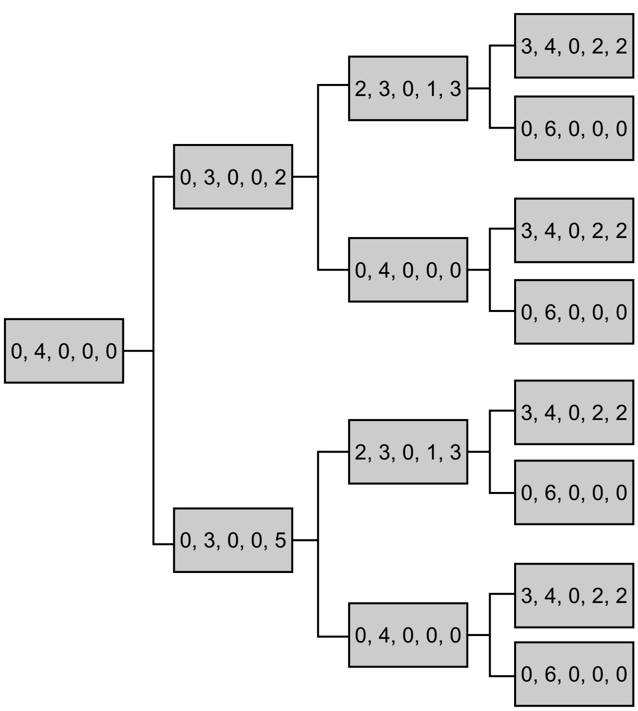

Figure 4 Sample Scenario Tree with $n_s = 4$ and $n_o = 2$ Represents the Requests for Urology Surgeries over a Four-Day Planning Period

Note. Each index, separated by comma, denotes the number of surgeries requested from a particular acuity level.

5.2. Analysis of the Enhanced Progressive Hedging Algorithm Solutions

We tested the EPHA to identify the factors that most affect the solution quality and algorithm performance. We investigated whether having the penalty update method achieves improvements over the solutions found by keeping ρ constant (i.e., $\delta_P = 1$, $\delta_D = 1$) through the iterations. We varied the values of δ_P and δ_D, keeping everything else constant, and then observed the changes in the average objective function value and computational time. Table 1 reports the resulting average objective function value and computational time for cases where varying values of δ_P and δ_D are used. Note that the penalty update method improves the computation time by 8% (i.e., when $(\delta_P, \delta_D) = (4, 1)$) while deteriorating the objective value only by 0.4%. However, in our experiments, since we used large-size test cases to estimate the VSS, we set $(\delta_P, \delta_D) = (1, 1)$.

The Lagrangian multiplier update method presented in Equation (29) also provides significant improvements to computation time, as it prevents cycling and ensures that the algorithm converges.

For the purpose of assessing the optimality gap, we tested the EPHA on four different instance sets (see Table 2). As indicated, each instance set comprises 20 different instances. Each set is characterized by the values of two different attributes: cancellation cost and expected standard deviation of the cumulative duration of the surgeries to be performed each day. The overtime cost was \$13/minute based on the average cost for Mayo Clinic reported by Batun et al. (2011). Note that the overtime cost includes more

Table 1 Sensitivity of the Objective Function Value and CPU Time to the Variations in the Penalty Update Multipliers

(δ_P, δ_D)	Expected cost (in dollars)	CPU time (in seconds)
(1, 1)	116,708	443.73
(2, 2)	117,173	443.51
(2, 1.5)	117,334	467.25
(2, 1)	117,339	450.95
(4, 1)	117,185	408.81
(16, 1)	117,268	422.30

Table 2 Characterization of the Instance Sets

Instance set #	Cancellation cost (in dollars)	Standard deviation of cumulative durations
1	1,700	78
2	2,000	78
3	1,700	234
4	2,000	234

Table 3 Comparison of the PHA Solutions with the CPLEX Solutions of the PHA Deterministic Equivalent Model (PHA-DEM) for Moderate-Size Test Cases in Terms of Objective Function Values

	Expected cost (in dollars)			Worst-case cost (in dollars)			Standard deviation (in dollars)	
Instance set #	CPLEX	PHA	Difference (%)	CPLEX	PHA	Difference (%)	CPLEX	PHA
1	23,988	24,039	0.21	30,270	30,270	0.00	1,789	1,817
2	23,988	24,032	0.18	30,270	30,270	0.00	1,789	1,814
3	26,828	27,012	0.68	42,388	42,388	0.00	3,643	3,696
4	26,828	27,004	0.65	42,388	42,388	0.00	3,643	3,692

than the direct exact financial cost of overtime for staff. In the context of ambulatory surgery, overtime is associated with an additional *goodwill cost* from staff who prefer not to have substantial overtime, even in light of the additional pay. It is also worth noting that reported cost of overtime varies significantly in the literature, and for this reason we present results of sensitivity analysis on overtime costs. It may be impossible to know the true costs of cancellation and waiting, since they are highly dependent on individual patients; therefore, we treat these "costs" as penalties in our experiments. The waiting cost was fixed at $600 based on studies by Stepaniak et al. (2009) and Tessler et al. (1997). The cancellation cost was set to either low ($1,700/surgery) or high ($2,000/surgery), based on the lower and upper bounds on the estimation of the average cost reported by Argo et al. (2009) (in §5.3, we present sensitivity analysis for waiting and cancellation costs). The mean sum of surgery durations was set to 540 minutes (nine hours) for the surgeries to be scheduled to an OR that is open for eight hours during regular time. The standard deviation of cumulative surgery durations for each service having two ORs was set to either a low of 78 or a high of 234. These values are consistent with the Mayo Clinic data and represent services observing lower and higher variability in surgery requests.

In Table 3, we compare the objective function values for the PHA solutions with the solutions found after directly solving the PHA-DEM using CPLEX 12 for the moderate-size test cases. Note that the gap between the EPHA and CPLEX solutions is 0.43% on average. CPLEX found the optimal solutions for all instances (taking about 1 second on average). EPHA also did not consume a substantial amount of time (about five minutes on average). Note that we compared the EPHA and CPLEX solutions only for the moderate-size test cases, because the use of CPLEX for large instances is not feasible because of memory and space limitations. The model initialization on CPLEX alone took three hours for the large-size test cases, whereas our EPHA obtained very good solutions within this time frame.

5.3. Model Sensitivity Analysis

5.3.1. Sensitivity to Cost Coefficients. We analyzed the sensitivity of optimal solutions to the changes in the cost coefficients for cancellation, waiting, and OR overtime costs. We emphasize the impact of the changes on the number of surgery cancellations over a planning period. The average number of cancellations for different choices of cancellation and waiting costs are given in Table 4. The comparison was conducted on instance set #2 in Table 2 and 29 new instance sets that were generated by varying the cancellation and waiting costs while keeping the overtime cost constant. Together these comprise the 30 instance sets used to generate Table 4.

Table 4 reveals that the main factor affecting the number of cancellations is the ratio of waiting cost to cancellation cost. This factor is important, because if the waiting cost is low, then the optimal solution assigns surgery to later days in the time window to prevent future cancellations from occurring. Since the low waiting cost leads to a lower number of cancellations, the services performing surgeries of lower urgency are likely to observe fewer cancellations. Figure 5 illustrates the trade-off between total expected number of cancellations and total expected number of waiting days. The trade-off for a given solution depends on the ratio of waiting cost to cancellation cost.

In Figure 5, the ratio was varied by keeping the waiting cost fixed at $600 and changing the cancellation cost from $200 to $2,000. We note that the number of cancellations is more sensitive to the ratio than the total expected waiting time is. We also note from

Table 4 Sensitivity of the Expected Number of Cancellations to the Variations in the Cancellation and Waiting Costs

	Waiting cost (in dollars)				
Cancellation cost (in dollars)	200	400	600	800	1,000
200	13.60	16.35	17.00	18.03	18.32
400	0.55	5.75	12.98	14.06	16.07
600	0.20	0.40	6.80	12.51	11.96
800	0.00	0.33	0.33	7.13	10.66
1,000	0.00	0.00	0.33	0.33	2.72
2,000	0.00	0.00	0.00	0.00	0.00

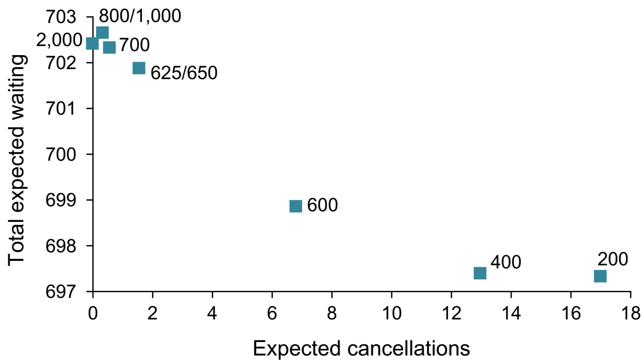

Figure 5 (Color online) Trade-Off Between Total Expected Number of Cancellations and Waiting Days Based on Different Waiting to Cancellation Cost Ratio

Note. Cancellation cost used (in dollars) for a test case is shown next to the corresponding data point.

Table 5 Expected Number of Cancellations as a Function of the Cancellation and Overtime Costs (Under Constant Waiting Cost)

Waiting cost = \$600	Overtime cost (\$)			
Cancellation cost (\$)	6	9	13	17
200	10.35	17.35	17	21.16
400	8.80	12.21	12.98	15.91
600	0.00	3.68	6.8	7.53
800	0.00	0.00	0.33	0.33
1,000	0.00	0.00	0.33	0.33
2,000	0.00	0.00	0.00	0.00

Figure 5 that there is a cancellation cost threshold at which the average number of cancellations becomes approximately zero. Table 5 illustrates the sensitivity of the average number of cancellations to the changes in the cancellation cost and overtime cost (waiting cost is constant in these cases). For Table 5, in addition to the six instance sets used for Table 4 (where overtime was set to \$13), 18 new sets were generated. Together these comprise the 24 instance sets used to create Table 5. The table also suggests that the ratio of waiting to cancellation cost has a significant influence on the number of cancellations. In the case that this ratio is less than or equal to one, the number of cancellations increases because of an increase in overtime cost.

5.3.2. Sensitivity to Variability in Demand. We evaluated the impact of uncertainty on the expected total cost and the number of surgery cancellations. Table 6 shows a comparison among six different moderate size instance sets. Besides the instance set used for Table 4 (where the standard deviation was set to 78), five new instance sets were included to generate Table 6. Note that the expected cumulative duration of the requested surgeries is held nearly constant (i.e., nine hours), whereas the standard deviation of the surgery durations is varied over different sets. The increase in standard deviation of cumulative durations is achieved by generating requests from various acuity levels (i.e., by increasing variability in requests). The values for the cost parameters are fixed for all instance sets ($c^o = 13, c^i = 200, l_i = 400$). Table 6 shows that the expected number of cancellations increases with respect to variance. The table also shows that when there is no uncertainty in demand, cancellations are not observed, and when the coefficient of variation is low, cancellations are still not observed. The results suggest that removing the variation in requests over a planning period will reduce the number of cancellations. It also suggests that in the surgical services whose surgery mixes may change significantly from day to day, cancellations are more likely to occur. Therefore, grouping similar types of surgeries and assigning them to the same day and OR may reduce cancellations.

Table 6 Variation in Expected Number of Cancellations and Total Cost (in Dollars) Due to Change in the Standard Deviation (σ^{ins}) of Daily Cumulative Surgery Durations in Minutes

	Standard deviation (σ^{ins})					
	0	39	78	117	156	234
Expected cancellations	0	0	3.44	4.31	8.03	23.21
Expected total cost	19,081	17,810	18,235	18,107	19,487	21,510

Note. Mean = 540 minutes for all instance sets.

5.4. Value of the Stochastic Solution

We compared the solutions for the stochastic model with the solutions of a deterministic heuristic to estimate the VSS. We used this estimate of VSS because it is not practical to find the exact VSS based on the mean value problem solution in our case. This is because the uncertainty in demand corresponds to discrete surgery requests, which lose their meaning if we take the mean and obtain a noninteger number of requests. The heuristic we used is an extension of the first fit decreasing heuristic, which is a well-known heuristic for bin-packing problems and representative of how scheduling is done in practice. The heuristic is myopic in the sense that it does not consider future outcomes when deciding on the assignment of surgeries into the future.

At each stage, the surgeries requested at that stage are ordered from longest to shortest expected duration. Next, the surgeries are assigned to a future day and OR, consecutively, according to their order on the surgery list. The heuristic attempts to assign a surgery to the earliest day available within the allowable time window. The availability of the day depends on the remaining capacities of the ORs that are appropriate for assignment in terms of equipment restrictions.

A capacity threshold is set for the ORs to prevent having high overtime. The thresholds are defined such that at most one surgery can be performed in an OR during the overtime period. The heuristic attempts to assign a surgery to the earliest available OR. If a surgery cannot be assigned to any open OR on a particular day, then a new OR is opened to assign the surgery. If there is no additional OR available on the same day, then the next day is considered. If the next day is outside of the allowable time window, the surgery is performed during the overtime hours in one of the day and OR combinations. The steps of the heuristic are summarized as follows:

Bin-Packing Heuristic

```
for t = 1 to H
  Sort the surgeries requested at stage t from
      longest to shortest duration to form the
      sorted list, L. Let L_i be the surgery in the ith
      order, and n^L be the size of the ordered list
  for i = 1 to n^L
    while surgery assigned = false
      for u = (t + g_{L_i}) to (t + g_{L_i} + h_{L_i})
        for j = 1 to O
          if Equipment constraint is not violated
            if Capacity constraint is not violated
              surgery assigned = true
            end if
          end if
          if surgery assigned = false
            if There is no more additional
                OR to open
              surgery assigned = true
            end if
          end if
        end for
      end for
    end while
  end for
end for
```

Table 7 compares the EPHA and the heuristic according to the solution quality and computation time based on 20 instances of a large-size instance set. We set the upper limit for an EPHA run as three hours. From Table 7, we conclude that the EPHA significantly improves the quality of solutions found by the heuristic. Note that the average and maximum gap between the objective values for the EPHA and heuristic solutions are 10.43% and 17.20%, respectively. The heuristic was very fast with a CPU time of less than one second. In contrast, on average, 2,304 CPU seconds were needed to obtain a solution using the EPHA.

5.5. Rolling Horizon Procedure

Our model can be used in practice for an extended length of planning periods (i.e., $P > H$) using a rolling horizon procedure (RHP) (Chand et al. 2002). This could be achieved as follows. Suppose that the surgery planning model is solved on day i for a planning period of H days. Once the schedule for day $i+1$ is put into practice, the model is solved for the next H days. The surgeries that are cancelled on day $i+1$ are then reconsidered for scheduling. The urgency levels of these surgeries are also increased by adjusting their cancellation and waiting costs before solving the new instance of the model. The schedule for day $i+2$ that the updated model yields is implemented next. After making the necessary updates on cancellation and waiting cost values, the model is again solved for the following H days. This pattern is repeated for the remaining periods of the extended horizon.

The RHP that we suggested represents an approximate solution procedure for a theoretically infinite horizon problem. Notably, Huang and Ahmed (2009) showed that solving a longer horizon problem with a shorter horizon stochastic programming model may provide poor results. To investigate this, we simulated the implementation of our model within a rolling horizon framework and compared the solutions with the heuristic solutions. We conducted the comparison based on 20 instances of a large-size instance set. The MSSMIP model was solved for a planning period of five days in each of the first five periods of the 11-period horizon. The solutions were averaged across periods and compared with the heuristic solutions obtained for the same 20 instances. Table 8 presents these results comparing RHP with the heuristic. From Table 8, we conclude that the

Table 7 Comparison of the EPHA with the First Fit Decreasing Heuristic

	Expected cost (in dollars)			CPU time (in seconds)	
Instance #	PHA	Heuristic	Difference (%)	PHA	Heuristic
1	22,179	24,322	8.81	1,246.15	0.00
2	21,138	25,397	16.77	834.98	0.00
3	17,941	18,708	4.09	1,115.56	0.00
4	17,332	19,500	11.11	580.00	0.00
5	19,197	20,888	8.09	1,803.31	0.00
6	16,277	17,234	5.55	530.56	0.00
7	22,681	25,236	10.12	10,800.00	0.00
8	18,055	21,728	16.90	781.57	0.00
9	15,104	16,010	5.65	537.83	0.00
10	22,898	25,900	11.59	785.59	0.00
11	15,860	17,119	7.35	1,051.96	0.00
12	17,480	19,066	8.31	1,493.44	0.00
13	21,391	23,425	8.68	570.20	0.00
14	20,399	22,798	10.52	10,800.00	0.00
15	16,783	18,781	10.63	340.50	0.00
16	19,880	22,514	11.69	540.95	0.00
17	20,548	24,182	15.02	728.26	0.00
18	20,307	24,527	17.20	309.18	0.00
19	19,911	21,456	7.20	10,800.00	0.00
20	27,241	31,427	13.31	421.78	0.00

Table 8 **Comparison of the RHP with the First Fit Decreasing Heuristic**

	Expected cost (in dollars)		
Instance #	RHP	Heuristic	Difference (%)
1	38,754	49,577	21.83
2	44,434	49,906	10.96
3	33,762	39,496	14.52
4	31,219	36,630	14.77
5	24,145	26,080	7.42
6	37,585	42,345	11.24
7	32,187	37,149	13.36
8	51,367	65,880	22.03
9	37,550	41,254	8.98
10	39,222	46,862	16.30
11	37,254	44,619	16.51
12	36,390	40,713	10.62
13	37,731	44,741	15.67
14	39,192	46,066	14.92
15	34,757	40,487	14.15
16	37,971	46,399	18.16
17	32,474	35,516	8.57
18	38,814	43,012	9.76
19	44,620	59,631	25.17
20	40,517	48,192	15.93

EPHA still outperforms the heuristic solutions significantly in the rolling horizon context. The average and maximum gaps between the objective values for the RHP and heuristic solutions are 14.54% and 25.17%, respectively.

The following review of articles where RHP has been used also justifies the RHP we suggested. Ovacik and Uzsoy (1994) used RHP for scheduling dynamic job arrivals on a single machine. They first solved a finite horizon problem, where a finite job set is considered. This set includes the jobs that arrive on the current day and some of the expected future job arrivals. This approach resembles ours, as we also considered some of the future surgery arrivals in our stochastic programming model. Among the future arrivals, Ovacik and Uzsoy (1994) considered the ones with earlier due dates. Next, they implemented the resulting finite horizon model solutions within an RHP. They showed that the RHP significantly outperforms myopic dispatching rules and a local search method. They also extended their results into the parallel machine problems in Ovacik and Uzsoy (1995), which has some similarities to the context of multi-OR surgery scheduling.

In another example, in which multistage stochastic programming was used, RHP was recommended for solving an infinite horizon problem to generate a master production schedule (MPS) (Korpeoglu et al. 2011). The authors also considered uncertainty in demand in their formulation for a finite planning horizon. They recommended that their stochastic programming model for MPS could be solved every day within an RHP. Finally, RHP was also suggested for healthcare scheduling problems. For example, Rohleder and Klassen (2002) tested appointment scheduling policies in a rolling horizon environment. They suggested that RHP-based rules would be an appropriate approach for practical settings.

6. Conclusions

In this article, we propose a multistage stochastic mixed-integer programming formulation for the allocation of surgeries to ORs under uncertainty over a finite planning horizon. We first implemented an extension of PHA, called EPHA, and compared EPHA solutions with the solutions found using CPLEX 12. We then analyzed the trade-offs between cancellation, waiting, and overtime costs with respect to their impact on total expected costs and surgery cancellations. We also assessed the impact of varying levels of uncertainty. We compared an easy-to-implement heuristic with the EPHA to estimate VSS to quantify the benefit of considering uncertainty in the surgery planning and scheduling process. We also discussed an implementation of the PHA within a rolling horizon framework for extended planning periods. The most significant findings of our study are as follows.

- The ratio of waiting cost to cancellation cost is one of the factors that most impacts the expected number of cancellations. If the waiting cost is higher than the cancellation cost, then cancellations are likely. A relatively low waiting cost allows a wider time window for scheduling surgeries, thus helping to reduce the cancellations. Cancellations also increase with an increase in overtime cost for the cases where the ratio of waiting to cancellation cost is less than or equal to one. Another factor to which the expected number of cancellations is highly sensitive is the level of uncertainty in demand and total daily surgery duration. Cancellations do not exist in the deterministic case but increase as the variation in demand increases.
- The EPHA outperforms a deterministic heuristic that may be representative of methods used in practice. The results suggest that a good heuristic must carefully handle the uncertainty in surgery requests, because the cancellations can be reduced when surgeries with similar durations are performed together. Thus, there may be benefits to studying either heuristics that minimize the maximum variance of total surgery durations over ORs and days, or ways to reduce variation in surgery requests.
- The initial values of Lagrangian multipliers and the Lagrangian multiplier and penalty update methods are the factors that most affect the performance and solution quality of the PHA. Our Lagrangian multiplier update method achieved fast convergence of the EPHA. Our penalty update method accelerated convergence but had a small negative effect on

the solution quality. The EPHA requires a short time to yield near optimal solutions for difficult problem instances that cannot be solved by CPLEX within a reasonable amount of time.

Acknowledgments
This material is based in part on work supported by the National Science Foundation [Grant no. CMMI 0844511]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Argo JL, Vick CC, Graham LA, Itani KMF, Bishop MJ, Hawn MT (2009) Elective surgical case cancellation in the Veterans Health Administration system: Identifying areas for improvement. *Amer. J. Surgery* 198:600–606.

Barro D, Canestrelli E (2005) Dynamic portfolio optimization: Time decomposition using the maximum principle with a scenario approach. *Eur. J. Oper. Res.* 163:217–229.

Batun S, Denton BT, Huschka TR, Schaefer AJ (2011) Operating room pooling and parallel surgery processing under uncertainty. *INFORMS J. Comput.* 23:220–237.

Chand S, Hsu VN, Sethi S (2002) Forecast, solution, and rolling horizons in operations management problems: A classified bibliography. *Manufacturing Service Oper. Management* 4:25–43.

Chiralaksanakul A, Morton DP (2004) Assessing policy quality in multi-stage stochastic programming. Higle JL, Römisch W, Sen S, eds. *Stochastic Programming E-Print Ser.* Accessed November 21, 2015, http://www.speps.org.

Crainic TG, Fu X, Gendreau M, Rei W, Wallace SW (2011) Progressive hedging-based metaheuristics for stochastic network design. *Networks* 58:114–124.

Dexter F, Maxbauer T, Stout C, Archbold L, Epstein R (2014) Relative influence on total cancelled operating room time from patients who are inpatients or outpatients preoperatively. *Anesthesia Analgesia* 118:1072–1080.

Epstein RH, Dexter F (2013) Rescheduling of previously cancelled surgical cases does not increase variability in operating room workload when cases are scheduled based on maximizing efficiency of use of operating room time. *Anesthesia Analgesia* 117:995–1002.

Fei H, Chu C, Meskens N (2009) Solving a tactical operating room planning problem by a column generation based heuristic procedure with four criteria. *Ann. Oper. Res.* 166:91–108.

Fei H, Meskens N, Chu C (2010) A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Comput. Indust. Engrg.* 58:221–230.

Fei H, Chu C, Meskens N, Artiba A (2008) Solving surgical cases assignment problem by a branch-and-price approach. *Internat. J. Production Econom.* 112:96–108.

Gerchak Y, Gupta D, Henig M (1996) Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Sci.* 42:321–334.

Gillen SMI, Catchings K, Edney L, Prescott R, Andrews SM (2009) What's all the fuss about? Day-of-surgery cancellations and the role of perianesthesia nurses in prevention. *J. Perianesthesia Nursing* 26:396–398.

Guinet A, Chaabane S (2003) Operating theatre planning. *Internat. J. Production Econom.* 85:69–81.

Gul S, Denton B, Fowler J, Huschka T (2011) Bi-criteria scheduling of surgical services for an outpatient procedure center. *Production Oper. Management* 20:406–417.

Hans E, Wullink G, van Houdenhoven M, Kazemier G (2008) Robust surgery loading. *Eur. J. Oper. Res.* 185:1038–1050.

Haugen KK, Lokketangen A, Woodruff DL (2001) Progressive hedging as a meta-heuristic applied to stochastic lot-sizing. *Eur. J. Oper. Res.* 132:116–122.

Healthcare Financial Management Association (HFMA) (2003) Achieving operating room efficiency through process integration. *Healthcare Financial Management* 57(3, Suppl.):1–8.

Helgason T, Wallace SW (1991) Approximate scenario solutions in the progressive hedging algorithm. *Ann. Oper. Res.* 31:425–444.

Huang K, Ahmed S (2009) The value of multistage stochastic programming in capacity planning under uncertainty. *Oper. Res.* 57:893–904.

Hvattum LM, Lokketangen A (2009) Using scenario trees and progressive hedging for stochastic inventory routing problems. *J. Heuristics* 15:527–557.

Korpeoglu E, Yaman H, Akturk MS (2011) A multi-stage stochastic programming approach in master production scheduling. *Eur. J. Oper. Res.* 213:166–179.

Lamiri M, Grimaud F, Xie X (2009) Optimization methods for a stochastic surgery planning problem. *Internat. J. Production Econom.* 120:400–410.

Lamiri M, Xie X, Zhang S (2008a) Column generation approach to operating theater planning with elective and emergency patients. *IIE Trans.* 40:838–852.

Lamiri M, Xie X, Dolgui A, Grimaud F (2008b) A stochastic model for operating room planning with elective and emergency demand for surgery. *Eur. J. Oper. Res.* 185:1026–1037.

Listes O, Dekker R (2005) A scenario aggregation-based approach for determining a robust airline fleet composition for dynamic capacity allocation. *Transportation Sci.* 39:367–382.

McManus ML, Long MC, Cooper A, Mandell J, Berwick DM, Pagano M, Litvak E (2003) Variability in surgical caseload and access to intensive care services. *Anesthesiology* 98:1491–1496.

Min D, Yih Y (2010) Scheduling elective surgery under uncertainty and downstream capacity constraints. *Eur. J. Oper. Res.* 206:642–652.

Mulvey JM, Vladimirou H (1991a) Applying the progressive hedging algorithm to stochastic generalized networks. *Ann. Oper. Res.* 31:399–424.

Mulvey JM, Vladimirou H (1991b) Solving multistage stochastic networks: An application of scenario aggregation. *Networks* 21:619–643.

Mulvey JM, Vladimirou H (1992) Stochastic network programming for financial planning problems. *Management Sci.* 38:1642–1664.

Ovacik IM, Uzsoy R (1994) Rolling horizon algorithms for a single machine dynamic scheduling problem with sequence dependent setup times. *Internat. J. Production Res.* 32:1243–1263.

Ovacik IM, Uzsoy R (1995) Rolling horizon procedures for dynamic parallel machine scheduling with sequence dependent setup times. *Internat. J. Production Res.* 33:3173–3192.

Rockafellar RT (1976) Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* 14:877–898.

Rockafellar RT, Wets RJB (1991) Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.* 16:119–147.

Rohleder TR, Klassen KJ (2002) Rolling horizon appointment scheduling: A simulation study. *Health Care Management Sci.* 5:201–209.

Santos MLLD, Silva ELD, Finardi EC, Goncalves REC (2009) Practical aspects in solving the medium-term operation planning problem of hydrothermal power systems by using the progressive hedging method. *Electrical Power Energy Systems* 31:546–552.

Schuster M, Neumann C, Neumann K, Braun J, Geldner G, Martin J, Spies C, Bauer M, CASCAES Study Group (2011) The effect of hospital size and surgical service on case cancellation in elective surgery: Results from a prospective multicenter study. *Anesthesia Analgesia* 113:578–585.

Stepaniak PS, Mannaerts GH, de Quelerij M, de Vries G (2009) The effect of the operating room coordinator's risk appreciation on operating room efficiency. *Anesthesia Analgesia* 108:1249–1256.

Takriti S, Birge JR (2000) Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Oper. Res.* 48:91–98.

Takriti S, Birge JR, Long E (1996) A stochastic model for the unit commitment problem. *IEEE Trans. Power Systems* 11:1497–1508.

Tessler MJ, Kleiman SJ, Huberman MM (1997) A zero tolerance for overtime increases surgical per case costs. *Canadian J. Anesthesia* 44:1036–1041.

Wallace SW, Helgason T (1991) Structural properties of the progressive hedging algorithm. *Ann. Oper. Res.* 31:445–456.

Watson JP, Woodruff DL (2011) Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Sci.* 8:355–370.

Zonderland ME, Boucherie RJ, Litvak N, Vleggeert-Lankamp CLAM (2010) Planning and scheduling of semi-urgent surgeries. *Health Care Management Sci.* 13:256–267.