IIE Transactions (2004) **36**, 155–167

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170490245450

A Bayesian stochastic programming approach to an employee scheduling problem

DAVID P. MORTON and ELMIRA POPOVA

Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin, Austin, TX 78712, USA
E-mail: morton@mail.utexas.edu and elmira@mail.utexas.edu

Received June 1999 and accepted May 2003

Bayesian forecasting models provide distributional estimates for random parameters, and relative to classical schemes, have the advantage that they can rapidly capture changes in nonstationary systems using limited historical data. Unlike deterministic optimization, stochastic programs explicitly incorporate distributions for random parameters in the model formulation, and thus have the advantage that the resulting solutions more fully hedge against future contingencies. In this paper, we exploit the strengths of Bayesian prediction and stochastic programming in a rolling-horizon approach that can be applied to solve real-world problems. We illustrate the methodology on an employee production scheduling problem with uncertain up-times of manufacturing equipment and uncertain production rates. Computational results indicate the value of our approach.

1. Introduction

We consider a production planning problem in which a nominal weekday employee schedule is published at the beginning of each month in the face of uncertain production rates and uncertain production equipment availability. During the course of the month, management follows the published weekday schedule as closely as possible. Overtime work for the weekends may be adaptively scheduled during the month. We use a Bayesian stochastic programming approach to this problem. Nominal weekday schedules are found by solving a two-stage stochastic program with recourse. The model's random parameters, i.e., the production rates and machine availabilities, are governed by Bayesian distributional forecasts. Based on observations made during the month, the distributions are then updated prior to reoptimizing to find the schedule for the coming month.

Optimal single-machine scheduling strategies, under deterministic processing times, are obtained by Smith (1956). Rothkopf (1966), Sevcik (1974) and Pinedo (1983) address stationary stochastic extensions of this problem. Gittins and Glazebrook (1977) and Hamada and Glazebrook (1993) solve Bayesian stochastic job scheduling problems on a single machine. The reliability literature includes Bayesian maintenance scheduling policies. Wilson and Popova (1996) provide a survey of these methods and Chen and Popova (2000) present a finite-horizon Bayesian maintenance scheduling problem.

Glover and McMillan (1986) survey various types of deterministic optimization models for employee scheduling problems. Süer (1996) considers a labor-intensive cellular manufacturing environment and uses a hierarchical approach to first determine alternative manpower levels and then determine product and employee assignments to cells via an integer program. Bailey *et al.* (1995) emphasize the importance of integrating project scheduling and manpower scheduling and develop a heuristic for an associated integer program. Lau (1996) considers an employee scheduling problem with logical constraints governing shift changes and shows that a restriction of this problem is NP-hard. Koop (1988) develops a network model for a shift-change problem that provides lower bounds on the required workforce size.

Birge and Dempster (1996) give a hierarchical framework for scheduling problems from a stochastic programming perspective, emphasizing the distinction between strategic, tactical and operational models. Easton and Rossin (1996) apply a tabu search heuristic to a stochastic integer program with simple recourse for an employee scheduling problem with uncertain demand. Kao and Queyranne (1985) use a two-stage stochastic program for a nurse scheduling problem with uncertain demand. The notion of recourse we use is similar in spirit to that of Kao and Queyranne in that we schedule regular-time work with first-stage variables and schedule overtime work with adaptive second-stage decisions.

An important and recurring theme in the stochastic programming literature concerns quantitatively establishing the importance of modeling uncertain parameters via distributional rather than point forecasts. This is usually done numerically, in the context of a specific application or test problem, by viewing the stochastic program as the "true" problem and by viewing optimization of the simpler point-forecast model as a heuristic solution procedure. In this paper, we extend this paradigm to the setting in which a distributional forecast obtained through a dynamic Bayesian update represents our best forecast for the stochastic parameters. We study the value of using this distributional forecast over a classical empirical distribution as well as Bayesian and empirical point estimates.

The stochastic program we develop uses a continuous distribution for production rates. It is usually impossible to solve such a problem exactly, and it is necessary to use an approximation. One common approach is to use Monte Carlo sampling to generate n observations of the random parameters and then solve the stochastic program obtained by replacing the original distribution with the sampled one. Under appropriate assumptions, accumulation points of the sequence of optimizers to these approximating problems under increasing sample size are optimal to the original problem, with probability one (w.p.1). For this and related results see Dupačová and Wets (1988), King and Wets (1991), Shapiro (1991), King and Rockafellar (1993) and Robinson (1996).

Mak *et al.* (1999) propose a method for testing the quality of a candidate solution by estimating the optimality gap, and we use this procedure to show that our approximate solutions are of high quality. This methodology has also been applied in: a bond portfolio optimization model (Bertocchi *et al.*, 2000); a stochastic integer knapsack problem (Morton and Wood, 1998); a stochastic water resources model to contain groundwater contaminants (Watkins *et al.*, 1998); and, stochastic vehicle routing problems (Kenyon and Morton, 2001; Verweij *et al.*, 2001). In addition this procedure is available on the Web through the VERIFY software as part of the MetaNEOS parallel computing project at Argonne National Laboratory (Zakeri, 2000).

In most of the stochastic programming literature, the distribution governing the random parameters does not depend on the decision taken. Only a few authors consider more general cases. Futschik and Pflug (1997) and Jonsbråten *et al.* (1998) consider problems in which there are a finite number of probability distributions that may arise. Decision-dependent randomness arises in our setting for the following reason; at the end of the month we construct predictive distributions for production rates based on observations which, in turn, depend on the production schedule we selected. For reasons of computational tractability we then solve the subsequent month's problem as a separate stochastic optimization problem, in a rolling-horizon fashion. A more sophisticated model would be required to capture the potential advantage of making decisions now that could "gather information" for the future. Artstein and Wets (1993) describe a framework for modeling the gathering of information in stochastic optimization problems, and Jonsbråten (1997) describes a stochastic programming approach for optimizing the sequence in which oil wells are drilled with Bayesian updating of well characteristics.

The remainder of the paper is organized as follows. Section 2 gives a description of the problem. Section 3 develops the Bayesian forecasting model and describes the deterministic and stochastic optimization models. Section 4 integrates the forecasting and optimization models. Computational results comparing four solution strategies based on Bayesian versus classical forecasting and deterministic versus stochastic optimization are detailed in Section 5.

2. Problem statement

This section describes an employee scheduling problem in a manufacturing plant. The system of interest is a production line in a major manufacturer of automobile parts. The line consists of several cells, each of which produces, or assembles, different parts of the line's products. The production equipment we consider is a set of lathe machines located in the line's initial cell. These machines produce different types of shafts. Due to the key role that these machines play in the production process, they operate up to 24 hours a day, on three shifts, each 8 hours long.

At the end of each month the production line's manager receives the demand forecast for the next month, which specifies the required number of shafts per type with due dates. Because the plant does not have significant storage space, the line operates almost as a just-in-time system, i.e., finished products are shipped to customers as they are produced. The nature of the industry is such that finished products may be viewed as commodities; there is substantial competition, and the manufacturer wants to maintain a high level of customer satisfaction. As a result, if a demand cannot be met on time, it is sent via a special express delivery service (which costs more than standard shipping) when production is complete. If the delay is too long, then the manufacturer must also pay a contractual penalty cost.

Scheduled production equipment down-times allow for preventive maintenance as well as crew training, meetings, and breaks. In addition, lathe machines fail at random times and then require corrective maintenance. Another source of uncertainty involves crew production rates. There is considerable variability in the actual production rates, both within a crew over time as well as between crews.

Each month, upper management allocates a budget to the production-line manager. The budget is a portfolio of different accounts. For instance, there is an account for workers' production wages (i.e., pay for time spent producing shafts), and other accounts to pay for training, for preventive or corrective maintenance, etc. At the end of each day, every worker fills out a time card describing what was done during the 8-hour shift. A typical example is: a 6-hour operating (production) time, 1-hour corrective maintenance, and 1-hour meeting time. The production-line manager, and senior management, are primarily interested in properly managing the account for production costs.

One of the line manager's tasks is to construct an employee schedule for shaft production for the coming month. The main goal is to deliver the requested production on time. In practice, the manager decides on an initial schedule, which is then adjusted, in a limited manner, over the course of the month. For instance, if, due to low production rates and machine availabilities, insufficient shifts are scheduled to meet demand, then overtime production shifts are scheduled for the weekends with associated higher overtime wages. If the production budget is exceeded by more than 2% in a month, then the manager receives a negative review. Because accurate production budgeting is important to senior management, this measure is one of the criteria used to decide wage increases for production managers and workers as well as to decide how bonuses are distributed.

Lathe machines have different production rates for each shaft type and for each crew as well as different down-time rates. Based on the demand schedule, and these relative efficiencies, the production-line manager makes work assignments to each shift crew for production on each lathe machine in an attempt to meet demands on time and to stay within the production budget.

3. Forecasting and optimization models

The purpose of the proposed forecasting and optimization models, and the associated interface, is two-fold. First, it is to generate production schedules that hedge against failing to meet demand requirements due to unscheduled machine down-times and random variations in production rates. Second, these schedules should be robust in that they minimize up-side deviations from a specified target for production costs.

The Bayesian estimation model of Section 3.1 provides point and distributional estimates for the hourly production rates by shift and shaft type and for up-times of the production equipment. These point estimates and predictive distributions are, respectively, used as inputs for the deterministic optimization model described in Section 3.2 and the stochastic optimization model of Section 3.3.

The deterministic optimization model constructs an "optimal" production schedule with regular-time and overtime work to minimize a weighted sum of penalties for late and nondelivered shipments plus a penalty for exceeding the target budget. However, this approach assumes that machine up-times and production rates are known with certainty, and as a result, its schedule may fail to hedge against the full range of machine-availability and production-rate contingencies. In Section 3.3, we extend the deterministic model to a stochastic program, which explicitly incorporates uncertainty in production rates and machine availabilities within the optimization model.

3.1. *Bayesian model for production rates and equipment up-time*

This section introduces Bayesian statistical models for machine up-time and production rates. If we had a large amount of historical data from a system believed to be relatively stationary, then classical point estimates and empirical frequency distributions could be used. However, when we do not have enough reliable data, or we have a new type of shaft to produce or a new lathe machine to operate or the equipment does not fail very often, then the use of classical estimates may yield either too optimistic or pessimistic results. In this section, we provide empirical evidence that the stochastic behavior of our system is nonstationary, and we propose a Bayesian time-dynamic model to capture this nonstationarity.

Five months (January–May) of data were collected. For each 8-hour shift we have the number of machine up-hours and the type of shaft(s) and number(s) produced. Most of the reliability literature provides models for the down-time of equipment. Here, we instead model the up-time because of the nature of the data. As described above, lathe operators account for each of the 8 hours of their shift on a time card. Because of management's emphasis on production accounting, a premium is placed on accurately documenting the time spent producing shafts, i.e., the up-time data is very reliable. In contrast, from the data cards it was clear that the same care was not taken in distinguishing the different reasons (e.g., meeting time versus corrective maintenance) for down-time. In Section 3.3, we show how preplanned down-times for reasons such as preventative maintenance, training meetings for operators, etc. are incorporated in the stochastic optimization model.

Table 1 contains descriptive statistics for the actual up-hours of a lathe machine. These statistics indicate that while the means do not fluctuate much from month to month, the variances do fluctuate. This observation is confirmed by the statistical tests we performed: t-tests for comparing the means and F-tests for the variances. Four t-tests were conducted for January–February, February–March, etc., under the null hypotheses that the mean up-times in adjacent months are equal. None of the four tests rejected the null hypothesis, each at a 0.05 level of significance. The second set of tests was analogous but for the variances. Each of these rejected the null hypothesis that adjacent variances are

Table 1. Descriptive statistics for the up-time of the lathe for each of the considered months

Month	*Mean*	*Median*	*Variance*
January	6.01	7.00	5.73
February	6.11	7.00	4.05
March	6.07	6.75	4.67
April	6.15	6.00	3.38
May	6.37	8.00	5.42

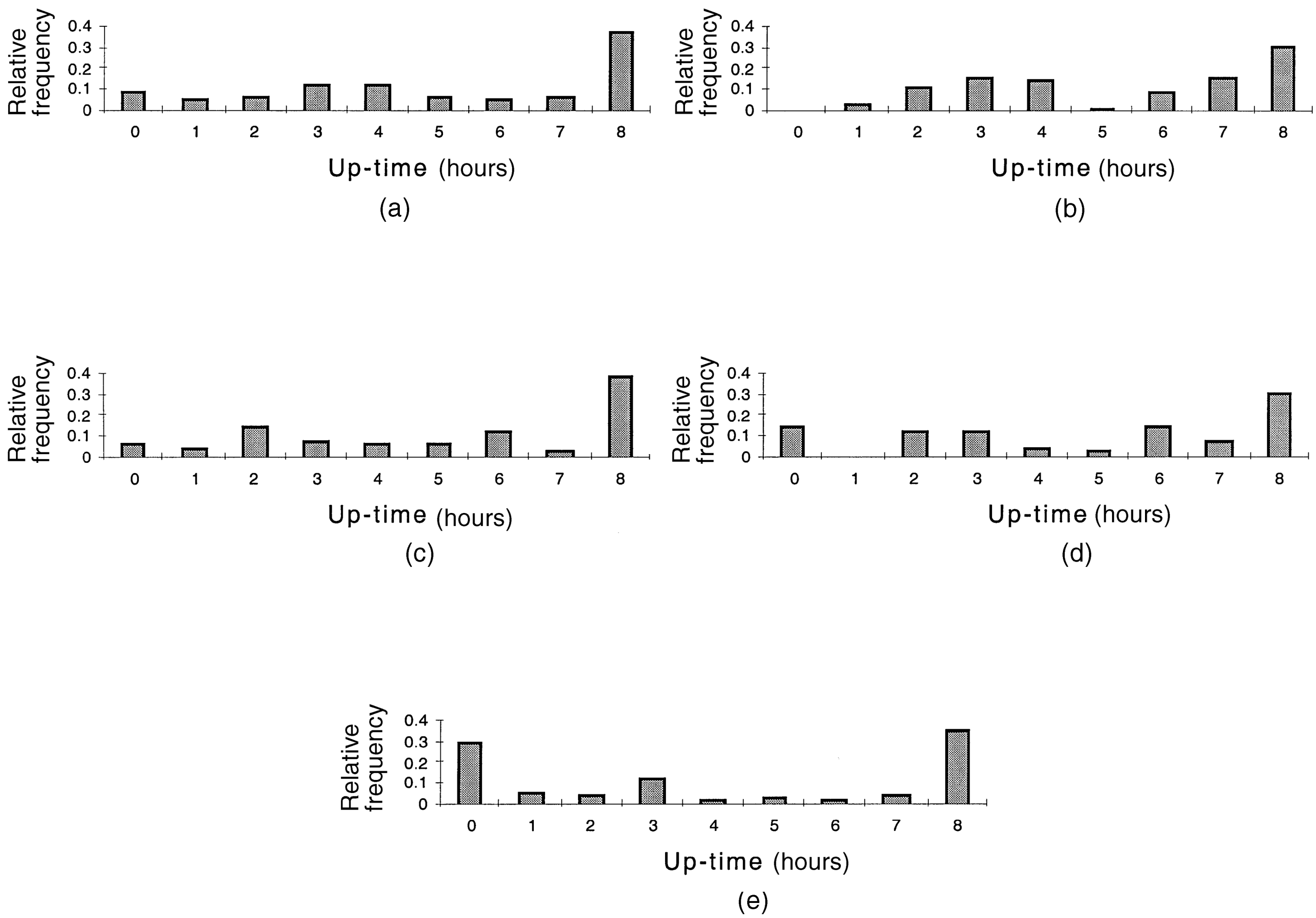

Fig. 1. Relative frequency histograms for the up-time of the lathe for each of the considered months: (a) January; (b) February; (c) March; (d) April; and (e) May.

equal. The tests' p-values ranged from 0.01 to 0.05. Schedules from a deterministic optimization model using such mean values would vary little from month to month. However, the changing variability of equipment availability may require different types of production schedules. Figure 1 depicts the relative frequency histograms of the up-time for each of the considered months, and further confirms the system's nonstationary behavior.

In addition to the empirical evidence of nonstationarity for up-times presented here, data for the production rates indicated their distributions also varied over time. As a result, it is clear that we require forecasting models that are nonstationary and learn from the past.

The stochastic optimization model defined in Section 3.3 is solved at the end of each month to obtain a schedule for the coming month. At these decision points we: (i) perform a statistical analysis of the past month's data and update our distributional forecast of the stochastic up-times and production rates for the coming month; and (ii) simulate the sample paths for these random parameters over the next month that are needed in our Monte Carlo approximation of the stochastic programming model (details on this approximation are in Sections 4.2 and 5.2).

First we describe the Bayesian model for up-hours. From the data cards, it was clear that up-time values entered by the operators were typically estimated to within an hour. So, we divide the 8-hour shift into eight 1-hour subintervals. Denote by $\boldsymbol{AH}_{cmt}$ the up-time of the equipment (available hours) during the 8-hour shift of crew c on machine m on day t. We use a hierarchical Bayesian structure, and model the (random) probability mass function (pmf) of $\boldsymbol{AH}_{cmt}$. Let $\boldsymbol{Z}$ be a random vector with realizations of the form $(r_1, \ldots, r_8)$, where $r_i = 0, 1, 2, \ldots$ and $\sum_{i=1}^{8} r_i = n$. Here, n is the number of observations over a month and r_i is the number of instances in the interval $(i-1, i]$, $i = 1, \ldots, 8$. We assume that $\boldsymbol{Z}$ is distributed as a multinomial random vector with parameters n and $\boldsymbol{W} = (\boldsymbol{W}_1, \ldots, \boldsymbol{W}_8)$, where $0 \leq \boldsymbol{W}_i \leq 1$, $\sum_{i=1}^{8} \boldsymbol{W}_i = 1$, and the value of $\boldsymbol{W}$ is unknown. The pmf for $\boldsymbol{AH}_{cmt}$ is obtained by dividing each of the components of $\boldsymbol{Z}$ by n. While the pmf for $\boldsymbol{AH}_{cmt}$ depends on the shift, machine, and time, we suppress the dependency of $\boldsymbol{Z}$ and $\boldsymbol{W}$ on c, m and t for clarity of the presentation.

Suppose that $\boldsymbol{W}$ is a Dirichlet random variable with parametric vector $\boldsymbol{\alpha} \equiv 0$ (which is the improper prior density, (DeGroot, 1970, p. 222). When prior knowledge is available a proper prior distribution is used instead. At the end of a month we have n observations of up-times for a particular crew-machine (*c-m*) combination with r_i, $i = 1, \ldots, 8$, observations in each interval. The Dirichlet distribution is a conjugate prior (e.g., DeGroot, 1970, p. 174) for the multinomial distribution, and as a result, the posterior distribution of $\boldsymbol{W}$ given r_i, $i = 1, \ldots, 8$, is a Dirichlet distribution with parametric vector $(\alpha_1 + r_1, \ldots, \alpha_8 + r_8)$. The predictive distribution for $\boldsymbol{Z}$ is a multi-nomial-Dirichlet distribution with parameters $(\alpha_1 + r_1, \ldots, \alpha_8 + r_8)$ (see Bernardo and Smith, 1994, p. 441). Thus, generating an observation of $\boldsymbol{AH}_{cmt}$ in a "fully Bayesian" manner involves two steps. First, we generate an observation of the pmf of $\boldsymbol{AH}_{cmt}$, i.e., $(1/n)\,\boldsymbol{Z}$, given $r_1, \ldots, r_8$, from the predictive multinomial-Dirichlet distribution. Then, given this instance of the pmf, we sample an observation of $\boldsymbol{AH}_{cmt}$. Our implementation takes a simpler approach; using the conditional mean of $(1/n)\,\boldsymbol{Z}$ given the previous month's observations, yields the following time-dynamic distributional forecast for $\boldsymbol{AH}_{cmt}$ (see Bernardo and Smith, 1994, p. 136):

$$P[\boldsymbol{AH}_{cmt} = i] = \frac{\alpha_i + r_i}{\sum_{j=1}^{8}(\alpha_j + r_j)}, \quad i = 1, \ldots, 8. \tag{1}$$

In generating observations of the up-hours we sample according to the distribution defined in Equation (1).

Now we turn to the Bayesian model for production rates. Let $\boldsymbol{PD}_{scmt}$ be the hourly production rate for a shaft of type s by crew shift c on machine m on day t. We follow the model of DeGroot (1970, p. 169), and assume that $\boldsymbol{PD}_{scmt}$ has a normal distribution with unknown mean, $\boldsymbol{M}$ and precision $\boldsymbol{R}$. (Precision is the reciprocal of variance.) Let $\boldsymbol{M}$ and $\boldsymbol{R}$ have an improper joint prior distribution $\pi(M, R) = 1/R$, $R > 0$. Then (see DeGroot, 1970, p. 170), given the observed production rates for a month $pd_{scmt} = (pd^1_{scmt}, pd^2_{scmt}, \ldots, pd^n_{scmt})$, the posterior distribution for $\boldsymbol{M}$ is a Student-*t* with parameters $\bar{x}$, $(n-1)/s^2$, and $(n-1)$ degrees of freedom, where $\bar{x}$ and s^2 are the sample mean and variance of the observed rates. The posterior distribution for $\boldsymbol{R}$ is gamma with parameters $(n-1)/2$ and $ns^2/2$. The predictive distribution for $\boldsymbol{PD}_{scmt}$ is a Student-*t* distribution with parameters $\bar{x}$, $[(n-1)(n+1)s^2]^{-1}$ and $(n-1)$ degrees of freedom. In building the Monte Carlo approximation of the stochastic program, we draw observations from a truncated version of the predictive Student-*t* to ensure non-negative rates, i.e.:

$$(\boldsymbol{PD}_{scmt} \mid pd_{scmt}) = \max\{t(\bar{x}, [(n-1)(n+1)s^2]^{-1}, n-1), 0\}. \tag{2}$$

To verify that we are sampling up-times from a representative distribution we performed 12 Kolmogorov–Smirnov tests (one for each of the three shifts in February–May). The null and alternative hypothesis are as follows:

H_0 The distribution of the lathe up-hours follows the mathematical model defined in Equation (1).
H_a The observed up-times do not coincide with those predicted by Equation (1).

None of the 12 tests rejected the null hypothesis at a 0.05 level of significance. Therefore, we regard the probability distribution of Equation (1) that is assumed to govern the up-times as appropriate. Similar tests can be performed for the production rates. However, the 5-month data set has 35 different types of shafts and three crews, and so we do not report such results here.

The random parameters $\boldsymbol{AH}_{cmt}$ and $\boldsymbol{PD}_{cmt}$ are assumed to be independent over crews, machines and time. We sample each such scalar random parameter according to the marginal distributions defined by Equations (1) and (2). While this form allows us to sample observations of the random parameters in a very straightforward fashion, a statistical model with a specified dependency structure could also be employed in the methodology we develop.

Prior to describing the stochastic optimization model we introduce a simpler model in which the up-time and production rate parameters are assumed to be deterministic. In order to compare the quality of the scheduling solutions obtained by distributional and point forecasts (Section 5), we use a Bayesian point forecast that is the mean of the distributions defined by Equations (1) and (2).

3.2. *Deterministic optimization model*

In this section we formulate a deterministic model for scheduling shaft production at the level of daily crew shifts for a month to minimize a weighted sum of penalties for late and nondelivered shipments plus a penalty for exceeding the target budget. The data for this model include deterministic demand for the number of shafts of each type with required delivery dates, as well as known production rates and equipment availabilities. After explaining the details of the deterministic model in this section, we then develop an extension to the case where machine up-times and production rates are random in Section 3.3.

Indices

t : time periods (days);
s : shaft type;
c : crew shift;
m : lathe machine.

Data

B : target production budget;
D_{st} : demand for shaft s on day t (during its delivery window);

AH_{cmt} : machine m availability: hours on day t for crew shift c;
PD_{scmt} : production rate: shafts per hour of type s by crew shift c on machine m on day t;
WR_c : hourly regular-time production wages for crew shift c;
WO_c : hourly overtime production wages for crew shift c.

Objective function terms

$g_{st}(\cdot)$: piecewise-linear penalty functions for late deliveries;
$h(\cdot)$: piecewise-linear penalty function for exceeding target budget;
λ : weight on penalty for exceeding target budget.

Decision variables

x_{scmt} : regular-time hours producing shaft type s by crew shift c on machine m on day t;
y_{scmt} : overtime hours producing shaft type s by crew shift c on machine m on day t;
u_{st} : unmet demand for shaft s as of day t;
v : total production costs.

$$\min_{x,y,u,v\geq 0} \sum_{s,t} g_{st}(u_{st}) + \lambda h(v), \tag{3}$$

subject to

$$\sum_{c,m,\,t'\leq t} PD_{scmt'}\, x_{scmt'} + \sum_{c,m,\,t'\leq t} PD_{scmt'}\, y_{scmt'} + u_{st} \geq D_{st} \quad \forall\, s,\, t, \tag{4}$$

$$\sum_{s} x_{scmt} \leq AH_{cmt} \quad \forall\, c, m, t, \tag{5}$$

$$\sum_{s} y_{scmt} \leq AH_{cmt} \quad \forall\, c, m, t, \tag{6}$$

$$\sum_{s,c,m,t} WR_c\, x_{scmt} + \sum_{s,c,m,t} WO_c\, y_{scmt} - v = 0. \tag{7}$$

The primary decision variables, x_{scmt} and y_{scmt}, are the number of regular-time and overtime hours that crew shift c should devote to producing shaft type s on machine m on day t. These decisions are driven by the demand schedule, D_{st}, and the relative efficiencies, by shaft type, of the different crew shifts on each lathe machine. Each day has three 8-hour crew shifts and each crew has its own lathe machine. So, the "c" index effectively acts as a finer-step time index. For each shipment requirement, there is a time-window during which production may be feasibly delivered, and D_{st} takes on the required number of shafts when the time index t is in this window and has a value of zero otherwise. If total shaft production of type s through day t, i.e., the first two terms on the left-hand side of constraint (4), meets the demand then the corresponding shortage variable, u_{st}, is zero and no penalty is incurred. Otherwise, $u_{st} > 0$ and this shortage as of period t (for t in the delivery window) is penalized at an increasing rate with the amount of shortfall via $g_{st}(u_{st})$. Increasingly strident penalties are incurred as the delivery becomes more delayed. The penalty structure also incorporates weights representing priorities for each s-t shipment combination. Shipments with flexible due dates are modeled with larger time windows.

If an order for 47 shafts of type s is due on day 5 but delivery will be accepted as late as day 9 then D_{st} has a value of 47 for $t = 5, \ldots, 9$ and has a value of zero otherwise. Separate shipment orders for the same shaft type are effectively separate commodities and can be modeled as such by introducing artificial shaft types. Alternatively, fewer decision variables are required if it is possible to aggregate such orders by using a *cumulative* requirements schedule, D_{st}, and defining appropriate penalty functions g_{st}. Our system has a modest number of separate orders for the same shaft type within a month and so we use the former approach.

The formulation is defined with appropriate restrictions on the allowed combinations of s, c, m, and t. For example, in producing shafts for a particular shipment, limited storage space is modeled by allowing the corresponding production variables, x_{scmt} and y_{scmt}, to exist only for an appropriate time window prior to the due date. Also, regular-time production is limited to weekdays and overtime production is limited to weekends, and certain shaft types can only be produced by certain crews. Constraints (5) and (6) limit the production time to be within the machine availability during each crew shift for each machine on each day of the month; constraint (5) is enforced only on weekday t's and constraint (6) only on weekend t's.

The penalty function $h(v)$ is convex and piecewise-linear with three pieces: (i) a penalty is incurred at a mild rate when production costs fall in $(0, B)$; (ii) up-side deviations from the target budget B are penalized at a modest rate up to $1.02B$; and (iii) the penalty rate is significantly higher when costs exceed $1.02B$. If the budget can be met, then (i) encourages lower-cost schedules. The differing penalty rates below and above 102% of the target budget represent the manager's differing preferences for these outcomes. The value of the weight λ is selected in such a way that minimizing unmet demand and late shipments is the primary objective and minimizing up-side deviations from the target budget is the secondary objective. So, the manager's goal of staying below the target budget will not be achieved, for example, by intentionally delaying or missing shipment deadlines in order to save wages. We note that the model has continuous decision variables because set-up times on the lathe machines are not significant and partial shipments are allowed.

3.3. *Stochastic optimization model*

We now extend the deterministic model of models (3)–(7) to handle uncertainty in machine availabilities and production rates. For the problem under consideration, monthly demand forecasts are relatively certain compared to production rates and machine up-times. However, we note that the model we develop, as well as the proposed solution methodologies, can also incorporate random demands.

The stochastic model is time-dynamic with respect to production scheduling decisions but is more limited in time-dynamics with respect to its handling of uncertainty. The days of the month are denoted by T, which is partitioned into $T = \mathcal{R} \cup \mathcal{O}$, where $\mathcal{R}$ and $\mathcal{O}$ represent weekdays (regular-time) and weekend-days (overtime). The random up-times and production rates are denoted $\boldsymbol{AH}_{cmt}$ and $\boldsymbol{PD}_{scmt}$. The corresponding sample space is Ω, and a realization of $(\boldsymbol{AH}, \boldsymbol{PD})$ is denoted $(AH^{\omega}, PD^{\omega})$.

A key issue in extending the optimization model to the stochastic domain concerns the timing of decisions and observations of the random parameters as well as the type of adaptive, or recourse, production scheduling decisions that are permitted. We utilize a two-stage stochastic program with recourse (see Beale, 1955; Dantzig, 1955; Wets, 1974) that has the following structure.

At the beginning of the month a nominal weekday production schedule is specified for each week, $[x^{\max}_{scmt}]_{t\in\mathcal{R}}$. This decision constitutes the "first-stage" decision because it is made with only distributional knowledge of the machine up-times and crew production rates. This nominal schedule is subject to the following variant of constraint (5):

$$\sum_s x^{\max}_{scmt} \le AH^{\max}_{cmt} \quad \forall\, c, m, t \in \mathcal{R}. \tag{8}$$

Here, $AH^{\max}_{cmt}$ is an optimistic bound on machine availability, i.e., 8 hours per crew shift less scheduled down-time for reasons such as preventive maintenance. After making this decision, an observation of machine availabilities and production rates, $(AH^{\omega}_{cmt}, PD^{\omega}_{scmt})$, is revealed for all $t \in T$ (i.e., for both the weekdays and the weekends for the entire month). Knowing this sample point ω, a set of second-stage recourse decisions is made that consists of four parts: (i) *actual* weekday (regular-time) production schedules, $[x^{\omega}_{scmt}]_{t\in\mathcal{R}}$; (ii) weekend (overtime) production schedules, $[y^{\omega}_{scmt}]_{t\in\mathcal{O}}$; (iii) unmet demand variables for shipments, $[u^{\omega}_{st}]_{t\in T}$; and (iv) production costs v^{ω}.

The actual weekday production schedule, $[x^{\omega}_{scmt}]_{t\in\mathcal{R}}$, is a scaled version of the nominal schedule, which ensures that weekday machine availability constraints, i.e., constraint (5) with stochastic availability $AH_{cmt} = AH^{\omega}_{cmt}$, are obeyed. This is effected via:

$$x^{\omega}_{scmt} = \min\left\{ x^{\max}_{scmt}, \left[\frac{AH^{\omega}_{cmt}}{AH^{\max}_{cmt}}\right] x^{\max}_{scmt} \right\} \quad \forall\, s, c, m, t \in \mathcal{R}. \tag{9}$$

This method for specifying the actual weekday schedule is related to an idea that Powell and Frantzeskakis (1994) call "restricted recourse" from their work in dynamic stochastic network optimization. Note that the distribution of $\boldsymbol{AH}_{cmt}$ is for a "typical" day. Occasionally, $AH^{\max}_{cmt}$ has a small value for irregular but scheduled down-times and Equation (9) ensures that the actual schedule reflects these unavailabilities. The weekend schedules, $[y^{\omega}_{scmt}]_{t\in\mathcal{O}}$, are recourse decisions in the usual sense while the shortage variables, $[u_{st}]_{t\in T}$, and production cost variables v^{ω} are "accounting" variables from which appropriate late-delivery and budget-deviation penalties are assigned.

The two-stage stochastic optimization model can be stated as:

$$\min_{x^{\max}} \quad Ef(x^{\max}, \boldsymbol{AH}, \boldsymbol{PD})$$

subject to

$$\begin{aligned} \sum_s x^{\max}_{scmt} &\le AH^{\max}_{cmt} \quad \forall\, c, m, t \in \mathcal{R}, \\ x^{\max}_{scmt} &\ge 0 \quad \forall\, s, c, m, t \in \mathcal{R}, \end{aligned} \tag{10}$$

where

$$f(x^{\max}, AH, PD) = \min_{x,y,u,v} \quad \sum_{s,t} g_{st}(u_{st}) + \lambda h(v),$$

subject to

$$\begin{aligned} & x_{scmt} \le \left[\frac{AH_{cmt}}{AH^{\max}_{cmt}}\right] x^{\max}_{scmt} \quad \forall\, s, c, m, t \in \mathcal{R}, \\ & \sum_{c,m,t'\le t} PD_{scmt'}\, x_{scmt'} \\ & \quad + \sum_{c,m,t'\le t} PD_{scmt'}\, y_{scmt'} + u_{st} \ge D_{st} \quad \forall\, s, t, \\ & \sum_s y_{scmt} \le AH_{cmt} \quad \forall\, c, m, t \in \mathcal{O}, \\ & \sum_{s,c,m,t} WR_c\, x_{scmt} + \sum_{s,c,m,t} WO_c\, y_{scmt} - v = 0, \\ & x_{scmt} \le x^{\max}_{scmt} \quad \forall\, s, c, m, t \in \mathcal{R}, \\ & x_{scmt}, y_{scmt}, u_{st}, v \ge 0 \quad \forall s, c, m, t. \end{aligned} \tag{11}$$

Some additional remarks on the above notion of recourse are in order. In practice, a nominal weekday schedule is published at the beginning of the month and management makes a concerted effort to follow this schedule. This is captured in the model because the nominal schedule must be stated prior to observing the up-times and production rates for the month and the model's actual weekday production schedule is determined by Equation (9). In the model, overtime production on Saturdays and Sundays is adaptively scheduled after the machine availabilities and production rates are known for the entire month. A potential concern is that overtime work under a specific scenario will be scheduled, say, at the end of the first week because it is already known that low production rates and/or machine availabilities will occur in week 2. In practice, this is not a significant issue because of the restrictions (dictated by storage constraints) on which shafts are eligible for overtime production.

4. Bayesian estimation and optimization

4.1. *Bayesian point forecasts and deterministic optimization*

Our simplest optimization model incorporates Bayesian point estimates in the deterministic model of models (3)–(7). At the end of a month, the manager has observed up-times and production rates. These are imported into the Bayesian estimation model of Section 3.1 which, in turn, provides point estimates for the number of up-hours and production per shift for each shaft type, crew shift, and lathe machine combination. These point estimates are then used to populate the deterministic optimization model of Section 3.2. A schedule for the following month is then obtained and this procedure is repeated monthly.

4.2. *Bayesian distributional forecasts and stochastic optimization*

Here we consider a solution methodology that is very similar, in principle, to that of the previous section except that the deterministic optimization model used to generate the production schedules is replaced by the stochastic programming model (10), of Section 3.3 and the Bayesian point estimates that feed the optimization model are replaced by the Bayesian distributional forecasts (predictive distributions).

We compactly state model (10) as:

$$z^* = \min_{x \in X} Ef(x, \boldsymbol{\xi}) \tag{12}$$

where x is the first-stage weekday production schedule, $[x_{scmt}^{\max}]_{t \in \mathcal{R}}$; $x \in X$ denotes the constraints of model (10); $\boldsymbol{\xi} = (\boldsymbol{AH}, \boldsymbol{PD})$ is the vector of random parameters; and $f(x, \xi^{\omega})$ is the second-stage cost of operating the production scheduling system, in terms of late-delivery and over-budget penalties, for a fixed first-stage schedule x and for the specific realization ξ^{ω} as defined in the recourse function (11).

While we can solve the deterministic optimization model (3)–(7) exactly (within numerical tolerances) using standard commercial software, due to the number of random parameters and the distributional assumptions of Section 3.1, it is impossible to solve model (12) exactly and so we resort to approximations.

The distributions of machine up-time and crew production rates, given the data observed up to that point in time, are defined by Equations (1) and (2), respectively. We replace these with an approximating distribution that is derived by sampling independent and identically distributed (i.i.d.) variates $\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^n$ from the distribution of $\boldsymbol{\xi}$ and forming the problem:

$$z_n = \min_{x \in X} \frac{1}{n} \sum_{i=1}^{n} f(x, \boldsymbol{\xi}^i). \tag{13}$$

Let x_n solve model (13). Large sample size properties of x_n and z_n can be obtained from the theory of epi-convergence. From Attouch and Wets (1990, Theorem 4.3), we know that the objective function of model (13) epi-converges to that of model (12), w.p.1, provided $f(x, \boldsymbol{\xi})$ is bounded below by a constant, w.p.1. This condition holds for our non-negative recourse function (11). Coupling the epi-convergence result with the fact that the constraint set X is compact implies consistency of z_n and x_n, i.e., z_n converges to z^*, w.p.1, and all limit points of $\{x_n\}_{n=1}^{\infty}$ solve model (12), w.p.1. (The existence of at least one limit point of $\{x_n\}_{n=1}^{\infty}$ in X is ensured as X is compact.)

In Section 3.1, we assumed a conjugate structure and hence eliminated any computational difficulties (such as multi-dimensional integration) in generating these samples. However, if one were to assume general distributions for machine up-times, production rates and their parameters, our methodology could still be applied. One approach would be to use Gibbs sampling (Gilks *et al.*, 1996) in order to generate the scenarios for the stochastic optimization problem.

When an instance of model (12) is approximated via model (13) it forms a large-scale linear program. If the number of scenarios is modest then we can use commercial implementations of the simplex method or an interior point method to solve the problem. Larger models require algorithms that exploit special structure, such as variants of the L-shaped cutting-plane algorithm; see Van Slyke and Wets (1969), Ruszczyński (1986), Birge and Louveaux (1988), and Ruszczyński and Świetanowski (1997).

5. Computational results

This section provides a computational comparison of four strategies for solving the employee scheduling problem that we have described. The results illustrate the benefits of using Bayesian forecasting and stochastic programming over classical statistical methods and deterministic optimization.

Our data covers January through May. In order to compute the first Bayesian update we begin the numerical comparisons in February. The optimization model has daily time periods with three 8-hour crew-shifts per day. There are demands for a total of 35 shaft types and the system has one lathe machine. The overtime salary rate is 1.75 times the regular-time rate. The marginal penalties increase by one order of magnitude for wage costs falling in each of the three intervals (0, *B*), (*B*, 1.02*B*), and (1.02*B*, ∞), where *B* is the budget.

We use February as an example to give the approximate size of the linear programs being solved. The deterministic (point forecast) model has 209 structural constraints, 1380 decision variables and 5088 nonzero coefficients. The stochastic programming model, under an approximating probability distribution with 100 observations of the random parameters, has 14 861 rows, 48 504 decision variables and 383 131 nonzero coefficients. The model has 695 random parameters (but only a fraction of these are active depending on the production schedule decision). Solving this model takes just under 40 seconds using GAMS 2.25 (Brooke *et al.*, 1998) calling the CPLEX 7.0 default parallel interior point method solver. All computational tests reported here were performed on a dual processor 1.7 GHz Pentium IV machine with 2 GB of memory. We are primarily concerned with demonstrating the potential value of our approach (as opposed to computational efficiency). No attempts were made to construct stochastic programs that better lend themselves to solution via interior point methods (see, for example, Lustig *et al.*, 1991 for a discussion of split-variable formulations) or to use cutting-plane methods.

It is common practice to demonstrate the value of a stochastic programming solution (Birge, 1982) by finding a solution based on a point forecast, e.g., the population mean, and then calculating the objective function value of that deterministic solution in the stochastic model. We use this approach to compare solutions obtained using Bayesian predictive distributions and point forecasts. To illustrate the value of using a dynamic Bayesian updating of a distribution over a classical stationary distribution we use as a benchmark the empirical distribution from all 5 months of data for production rates and machine availabilities. Like the Bayesian distributions of Section 3.1, there is a separate marginal empirical machine availability distribution for each crew shift. Similarly, the empirical production rate distributions depend on the shaft type and crew. All 5 months of data are used to construct the empirical distribution as an approximation of what would be done under the assumption of stationarity for the stochastic behavior of the system.

In Section 5.1 we present a computational comparison of the quality of the solutions obtained using four different approaches. Section 5.2 then presents evidence that our solutions to the sampled approximating problems are of high-quality relative to their "true" distribution counterparts. Finally, Section 5.3 investigates sensitivity of our approach with respect to the variability of $\boldsymbol{\xi}$ and the value of the budget B.

5.1. *Comparing solution approaches*

This section compares the quality of the nominal regular-time employee production schedule obtained via: (i) an empirical point forecast; (ii) a Bayesian point forecast; (iii) an empirical distribution; and (iv) a Bayesian distributional forecast. We use x_{ep}, x_{bp}, x_{ed}, and x_{bd} to denote the respective nominal schedules produced using these four forecasts. We perform an all pairwise comparison for $Ef(x_{\text{ep}}, \boldsymbol{\xi})$, $Ef(x_{\text{ed}}, \boldsymbol{\xi})$, $Ef(x_{\text{bp}}, \boldsymbol{\xi})$, and $Ef(x_{\text{bd}}, \boldsymbol{\xi})$. Here, the expectation is taken with respect to the Bayesian predictive distributions since these represent our best forecasts of how the system will behave. We form joint confidence intervals for the six differences between the objective function values (Law and Kelton, 2000, Section 10.3.2) using a paired Student-t approach with common random numbers. In order for the six confidence intervals to hold jointly at a 0.95 level we have, via the Boole-Bonferroni inequality, that each is made at level $1 - 0.05/6 \approx 0.9917$.

The differences estimated in the first row under each month in Table 2 indicate the expected cost savings obtained by using the x_{bd} schedule over the other three approaches. The second and third rows show similar results for x_{ed} and x_{bp}. For example, $Ef(x_{\text{ed}}, \boldsymbol{\xi}) - Ef(x_{\text{bd}}, \boldsymbol{\xi})$ indicates the savings from solving the (sample approximation of the) stochastic program using the Bayesian distribution over the empirical distribution. The point estimate of this difference in February is 510.68, about 8% of the point estimate of $Ef(x_{\text{bd}}, \boldsymbol{\xi})$ (see subsequent Table 4), and this value ranges from 4 to 20% over the 4 months. Analogous savings for x_{bd} over x_{bp} and x_{ep} range from 8 to 42%. The results in Table 2 suggest the quality of the schedules (best to worst) for the four forecasting strategies is: Bayesian predictive distribution, empirical distribution, and the two point forecasts. There is not a clear ordering between the Bayesian and empirical point forecasts. The computational results do seem to indicate that using an inaccurate distributional forecast (in our case, the empirical one) may be preferable to using point forecasts, even if the latter are accurate in mean (the Bayesian point forecast).

Table 2. Paired Student-t tests for an all pairwise comparison using common random numbers at a joint 95% level with a sample size of 1000. The expectations are with respect to the Bayesian predictive distribution. The table contains confidence intervals for the column entry less the row entry; e.g., the first entry for the month of February of 510.68 ± 39.04 is a confidence interval for $Ef(x_{\text{ed}}, \boldsymbol{\xi}) - Ef(x_{\text{bd}}, \boldsymbol{\xi})$

	$Ef(x_{ed}, \boldsymbol{\xi})$	$Ef(x_{bp}, \boldsymbol{\xi})$	$Ef(x_{ep}, \boldsymbol{\xi})$
February			
$Ef(x_{\text{bd}}, \boldsymbol{\xi})$	510.68 ± 39.04	661.06 ± 40.64	713.58 ± 42.74
$Ef(x_{\text{ed}}, \boldsymbol{\xi})$		150.38 ± 41.31	202.901 ± 38.51
$Ef(x_{\text{bp}}, \boldsymbol{\xi})$			52.52 ± 43.62
March			
$Ef(x_{\text{bd}}, \boldsymbol{\xi})$	236.05 ± 29.45	883.93 ± 40.46	682.02 ± 43.73
$Ef(x_{\text{ed}}, \boldsymbol{\xi})$		647.88 ± 42.62	445.970 ± 41.29
$Ef(x_{\text{bp}}, \boldsymbol{\xi})$			-201.91 ± 39.55
April			
$Ef(x_{\text{bd}}, \boldsymbol{\xi})$	641.93 ± 48.21	1258.68 ± 52.19	1349.42 ± 53.67
$Ef(x_{\text{ed}}, \boldsymbol{\xi})$		616.75 ± 37.11	707.49 ± 38.22
$Ef(x_{\text{bp}}, \boldsymbol{\xi})$			90.74 ± 36.39
May			
$Ef(x_{\text{bd}}, \boldsymbol{\xi})$	90.35 ± 23.76	177.33 ± 20.43	158.53 ± 19.58
$Ef(x_{\text{ed}}, \boldsymbol{\xi})$		86.97 ± 30.65	68.17 ± 21.68
$Ef(x_{\text{bp}}, \boldsymbol{\xi})$			-18.80 ± 23.53

5.2. *Establishing quality of solutions to the approximating problems*

We cannot solve the stochastic programming model (12) exactly under the empirical distribution or the Bayesian distribution given the large number of scenarios under the former and the continuous random parameters used in the latter. As described in Section 4.2, we instead solve model (13) and obtain $x_{bd} = x_{bd}(n)$ and $x_{ed} = x_{ed}(n)$. Here, we parameterize these solutions by n to emphasize that they depend on the sample size used in solving model (13). We also know from Section 4.2 that as n grows large, $x_{bd}(n)$ and $x_{ed}(n)$ solve, w.p.1, their "true-distribution" versions of model (12). However, for finite n this solution approach may be viewed as a heuristic. In our numerical tests in Section 5.1 we obtained $x_{bd}(n)$ and $x_{ed}(n)$ using $n = 100$. Our goal here is to quantify the quality of these solutions.

We measure quality via the optimality gap, $E_P f(\hat{x}, \boldsymbol{\xi}) - z_P^*$, where $\hat{x} = x_{bd}(n)$ or $\hat{x} = x_{ed}(n)$, P is the Bayesian or empirical distribution, and z_P^* is the optimal solution value to model (12) under the corresponding distribution. We use the procedure of Mak *et al.* (1999) and ascertain the quality of $\hat{x}$ by constructing a one-sided confidence interval on $Ef(\hat{x}, \boldsymbol{\xi}) - z^*$.

Tables 3 and 4 contain results for the empirical and Bayesian predictive distributions, respectively. For example, Table 3 shows that for the month of February, under the empirical distribution, we solved an instance of model (13) with a sample size of 100 in order to generate the candidate solution $\hat{x} = x_{ed}(100)$ (i.e., the nominal schedule). To test the quality of $\hat{x}$, we generated 30 point estimates of the optimality gap with each point estimate based on a batch size of $n = 100$ (for details of the procedure see Mak *et al.* 1999). This yielded a point estimate of 58.23 and an approximate 95% confidence interval of [0, 62.31] for $Ef(\hat{x}, \boldsymbol{\xi}) - z^*$. Restated, we found the nominal schedule, $\hat{x} = x_{ed}(100)$, by solving an approximating problem and $\hat{x}$ is almost certainly not optimal under the empirical distribution. However, we are confident at a 95% level that the objective function value of $\hat{x}$, i.e., $Ef(\hat{x}, \boldsymbol{\xi})$, is within 62.31 units of the optimal objective function value. Given the upper and lower bound point estimates on z^* in Table 3 we regard this as an acceptable solution.

Table 3. Test results for February–May using the empirical distribution

	Month			
	February	March	April	May
Sample size used to generate $\hat{x}$	100	100	100	100
No. of batches	30	30	30	30
Batch size	100	100	100	100
Upper bound estimate	5847.01	7837.78	4643.42	2366.73
Lower bound estimate	5788.78	7773.68	4580.19	2359.31
Gap estimate	58.23	64.10	63.23	7.42
Error estimate ($\alpha = 0.95$)	4.08	4.10	2.89	0.73
Confidence interval (95%)	[0,62.31]	[0,68.20]	[0,66.12]	[0,8.15]
Interval width/lower bound estimate	0.011	0.0088	0.014	0.0035

Table 4. Test results for February–May using the Bayesian predictive distributions

	Month			
	February	March	April	May
Sample size used to generate $\hat{x}$	100	100	100	100
No. of batches	30	30	30	30
Batch size	100	100	100	100
Upper bound	6431.07	5420.69	3208.91	2013.06
Lower bound	6209.17	5271.34	3011.48	1962.20
Gap estimate	221.90	149.35	197.43	50.86
Error estimate ($\alpha = 0.95$)	10.96	7.55	7.57	3.92
Confidence interval (95%)	[0,232.86]	[0,156.90]	[0,205.00]	[0,54.78]
Interval width/lower bound estimate	0.038	0.030	0.068	0.028

5.3. *Model sensitivity*

In this section we investigate the sensitivity of the Bayesian stochastic programming approach in two ways. First, we study the value of the Bayesian stochastic solution with respect to the variability of the Bayesian distributional forecasts. Then, we investigate the value of the solution under our Bayesian distributions to solutions under empirical distributions and Bayesian point forecasts as the budget deviates from its nominal value.

We consider the role of variability in the value of the stochastic programming solution for model (12) under the Bayesian distributions for each month, February through May. To do so, we parameterize the random elements by a scalar $\beta \in [0, 1]$ via $\boldsymbol{\xi}(\beta) = E\boldsymbol{\xi} + \beta(\boldsymbol{\xi} - E[\boldsymbol{\xi}])$. When $\beta = 1$, $\boldsymbol{\xi}(1) = \boldsymbol{\xi}$ yields the stochastic programs analyzed in Sections 5.1 and 5.2. When $\beta = 0$, $\boldsymbol{\xi}(0) = E\boldsymbol{\xi}$ has the degenerate distribution associated with the Bayesian point forecast. Values of $\beta \in (0, 1)$ allow us to systematically investigate the effect of variability on the value of the Bayesian stochastic solution, i.e.:

$$VBS(x_{bp}, \beta) = Ef(x_{bp}, \boldsymbol{\xi}(\beta)) - Ef(x_{bd}(\beta), \boldsymbol{\xi}(\beta)). \quad (14)$$

When $\beta = 1$, $VBS(x_{bp}, 1)$ is simply the $Ef(x_{bp}, \boldsymbol{\xi}) - Ef(x_{bd}, \boldsymbol{\xi})$ values in Table 2, and $VBS(x_{bp}, 0) = 0$. Table 5 displays point estimates of $VBS(x_{bp}, \beta)$ and 95%

Table 5. The table shows estimates of the value of the stochastic solution $VBS(x_{bp}, \beta)$, defined in Equation (14), as β shrinks from unity to zero for the months February through May. The values are expressed as a percentage of the point estimate of $Ef(x_{bd}(\beta), \boldsymbol{\xi}(\beta))$ and the values in parentheses are 95% confidence interval half-widths, expressed as the same percentage. When $\beta = 1$ the distribution of $\boldsymbol{\xi}(1)$ is that of the Bayesian predictive distributions given by our observed data. Smaller values of β yield less variable distributions until the Bayesian point forecast is achieved at $\beta = 0$

	$VBS(x_{bp}, \beta)$ as a percentage of $Ef(x_{bd}(\beta), \boldsymbol{\xi}(\beta))$			
β	*February*	*March*	*April*	*May*
1.00	9.60 (0.57)	16.42 (0.70)	37.95 (0.99)	8.57 (0.71)
0.90	9.63 (0.58)	15.00 (0.63)	33.49 (0.65)	7.11 (0.93)
0.80	9.32 (0.51)	14.36 (0.46)	28.42 (1.76)	5.69 (0.71)
0.70	7.88 (0.39)	13.56 (0.52)	24.75 (1.60)	4.46 (0.49)
0.60	6.86 (0.29)	11.58 (0.50)	20.11 (0.98)	3.76 (0.46)
0.50	5.59 (0.38)	10.22 (0.38)	16.50 (0.77)	3.64 (0.23)
0.40	3.85 (0.21)	9.40 (0.31)	11.13 (0.64)	3.39 (0.42)
0.30	2.20 (0.22)	7.19 (0.29)	8.00 (0.20)	2.66 (0.25)
0.20	1.09 (0.11)	4.63 (0.13)	4.40 (0.33)	2.23 (0.19)
0.10	0.47 (0.07)	1.82 (0.07)	2.28 (0.10)	1.03 (0.09)
0.00	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

confidence interval half-widths, expressed as a percentage of the point estimate of $Ef(x_{bd}(\beta), \boldsymbol{\xi}(\beta))$, for these and intermediate values of β. As in Sections 5.1 and 5.2, a sample size of $n = 100$ was used to find $x_{bd}(\beta)$ and a sample size of 1000 was used to estimate $VBS(x_{bp}, \beta)$. As Table 5 indicates, the value of the Bayesian stochastic solution can be significant even when the variability of $\boldsymbol{\xi}(\beta)$ is significantly smaller than that indicated by the data we observed.

Next, we consider the value of the Bayesian stochastic solution relative to those of the empirical stochastic solution and the Bayesian point-forecast solution as the value of the budget B changes from its nominal value, which we denote B_o. Specifically, in Table 6 we estimate, and construct confidence intervals for:

$$VBS(x, B) = Ef(x, \boldsymbol{\xi}) - Ef(x_{bd}(B), \boldsymbol{\xi}),$$

for $x = x_{ed}(B)$ and $x = x_{bp}(B)$, where these denote optimal solutions under the empirical distribution and the Bayesian point forecast, respectively, when the value of the budget is B. Table 6 contains these values for the months of February and April, which are representative of all 4 months. Several observations stand out. First, $Ef(x_{bd}(B), \boldsymbol{\xi})$ shrinks as B grows and then becomes constant when a larger budget no longer improves scheduling. This occurs with roughly a 50% increase over the nominal budget. Similarly, $VBS(x_{ed}(B), B)$ shrinks with growing budget until it becomes constant, although the scheduling improvement over the empirical-distribution schedule remains significant. As B shrinks from its nominal value, the expected cost $Ef(x_{bd}(B), \boldsymbol{\xi})$ continues to grow because of increased shortage penalties and increased overtime wage costs. Interestingly, $VBS(x_{ed}(B), B)$ becomes essentially constant for sufficiently small B, meaning that the relative difference between the empirical- and Bayesian-distribution schedules stops changing. This occurs at about 80–90% of the nominal budget. The values of $VBS(x_{bp}(B), B)$ roughly follow those of $VBS(x_{ed}(B), B)$ but are more erratic. This is due to the presence of multiple optimal solutions and the fact that small changes in the deterministic linear program's data can lead to significant changes in the recommended schedule. The stochastic programs, under either distribution, tend to produce schedules that change more smoothly with varying B. Again, a sample size of $n = 100$ was used to find $x_{bd}(B)$ and a sample size of 1000 was used to estimate $VBS(x_{ed}(B), B)$ and $VBS(x_{bp}(B), B)$. As in the case under modified variability of the distributional forecast, Table 6 indicates that the value of the Bayesian stochastic solution can be significant even when the value of B deviates from its nominal value.

Table 6. The table shows estimates of the value of the Bayesian stochastic solution $VBS(x, B)$ for $x = x_{ed}(B)$ and $x = x_{bp}(B)$ as the budget is varied from its nominal value of B_o. The values in parentheses are 95% confidence interval half-widths. Point estimates of $Ef(x_{bd}(B), \boldsymbol{\xi})$ are provided for reference.

B	$VBS(x_{ed}(B), B)$	$VBS(x_{bp}(B), B)$	$Ef(x_{bd}(B), \boldsymbol{\xi})$
February			
$1.7B_o$	427.19 (39.94)	610.83 (28.56)	5547.56
$1.6B_o$	427.19 (39.94)	610.83 (28.56)	5547.56
$1.5B_o$	426.57 (39.94)	457.32 (23.38)	5547.63
$1.4B_o$	428.05 (39.49)	455.25 (23.51)	5551.03
$1.3B_o$	444.68 (41.03)	602.38 (29.68)	5574.62
$1.2B_o$	493.90 (37.33)	488.24 (22.45)	5679.23
$1.1B_o$	510.67 (36.39)	516.29 (25.62)	5941.30
$1.0B_o$	510.68 (38.83)	537.49 (28.90)	6352.78
$0.9B_o$	508.81 (38.60)	672.69 (39.19)	6803.09
$0.8B_o$	508.83 (38.59)	658.03 (31.46)	7253.41
$0.7B_o$	508.83 (38.59)	658.79 (35.65)	7703.74
April			
$1.7B_o$	412.32 (37.52)	742.61 (37.05)	2832.36
$1.6B_o$	412.32 (37.52)	732.32 (37.46)	2832.36
$1.5B_o$	412.72 (37.49)	732.32 (37.46)	2832.36
$1.4B_o$	417.05 (38.45)	732.24 (37.43)	2832.76
$1.3B_o$	435.57 (38.48)	741.30 (38.61)	2833.45
$1.2B_o$	472.13 (40.29)	798.11 (41.23)	2864.11
$1.1B_o$	530.68 (42.90)	1004.60 (45.15)	2957.08
$1.0B_o$	641.93 (42.75)	1257.34 (40.63)	3169.32
$0.9B_o$	770.39 (44.44)	1418.30 (39.49)	3504.46
$0.8B_o$	800.30 (45.06)	1392.50 (44.67)	3969.09
$0.7B_o$	799.64 (45.43)	1395.52 (43.92)	4465.12

6. Summary

In this paper we have considered an employee scheduling problem in a production environment in which employees are assigned to production tasks over the course of a month.

The optimization problem is stochastic because machine availabilities and production rates are uncertain. We utilized a Bayesian forecasting model for these random parameters because there is limited historical data and there is evidence that the up-time and production rate distributions are not stationary over time. We regard the Bayesian predictive distributions as our most accurate distributional estimate for these uncertain parameters. Our computational results suggest that significantly better schedules are obtained using distributional (Bayesian or empirical) models for these random parameters compared to schedules generated using point estimates, and that there is a significant advantage to using a dynamic Bayesian methodology for updating of distributions over classical stationary schemes.

Acknowledgements

The authors are grateful for the support of Eaton Corporation's Fluid Power Division and in particular thank Ed Carlyle, now Division Operations Manager, Actuator and Sensor Division. The authors also thank two anonymous referees for helping to improve the paper. David Morton's research was supported by the National Science Foundation through grants DMI-9702217 and DMI-0217927 and the State of Texas Advanced Research Program through grant 003658-0405. Elmira Popova's research was partially supported by grant 0003658-472 from the State of Texas Advanced Technology Program.

References

Artstein, Z. and Wets, R.J.-B. (1993) Sensors and information in optimization under stochastic uncertainty. *Mathematics of Operations Research*, **18**, 523–547.

Attouch, H. and Wets, R.J.-B. (1990) Epigraphical processes: laws of large numbers for random lsc functions, in *Proceedings of the Séminaire d'Analyse Convexe*, Montpellier, France, pp. 13.1–13.29.

Bailey, J., Alfares, H. and Lin, W.Y. (1995) Optimization and heuristic models to integrate project task and manpower scheduling. *Computers and Industrial Engineering*, **29**, 473–476.

Beale, E.M.L. (1955) On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society*, **17B**, 173–184.

Bernardo, J.M. and Smith, A.F.M. (1994) *Bayesian Theory*, J. Wiley, Chichester, UK.

Bertocchi, M., Dupačová, J. and Moriggia, V. (2000) Sensitivity of bond portfolio's behavior with respect to random movements in yield curve: a simulation study. *Annals of Operations Research*, **99**, 267–286.

Birge, J.R. (1982) The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, **24**, 314–325.

Birge, J.R. and Dempster, M.A.H. (1996) Stochastic programming approaches to stochastic scheduling. *Journal of Global Optimization*, **9**, 417–451.

Birge, J.R. and Louveaux, F.V. (1988) A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research*, **34**, 384–392.

Brooke, A., Kendrick, D., Meeraus, A. and Raman, R. (1998) *GAMS: A User's Guide*, GAMS Development Corporation, Washington DC.

Chen, T. and Popova, E. (2000) Bayesian maintenance policies during a warranty period. *Communications in Statistics: Stochastic Models*, **16**, 121–142.

Dantzig, G.B. (1955) Linear programming under uncertainty. *Management Science*, **1**, 197–206.

DeGroot, M.H. (1970) *Optimal Statistical Decisions*, McGraw-Hill, New York, NY.

Dupačová, J. and R.J.-B. Wets (1988) Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *The Annals of Statistics*, **16**, 1517–1549.

Easton, F.F. and Rossin, D.F. (1996) A stochastic goal program for employee scheduling. *Decision Sciences* **27**, 541–563.

Futschik, A. and Pflug, G.Ch. (1997) Optimal allocation of simulation experiments in discrete stochastic optimization and approximative algorithms. *European Journal of Operational Research*, **101**, 245–260.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (eds) (1996) *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

Gittins, J.C. and Glazebrook, K.D. (1977) On Bayesian models in stochastic scheduling. *Journal of Applied Probability*, **14**, 556–565.

Glover, F. and McMillan, C. (1986) The generalized employee scheduling problem: an integration of MS and AI. *Computers and Operations Research*, **13**, 563–573.

Hamada, T. and Glazebrook, K.D. (1993) A Bayesian sequential single machine scheduling problem to minimize the expected weighted sum of flowtimes of jobs with exponential processing times. *Operations Research*, **41**, 924–934.

Jonsbråten, T.W. (1997) Optimal selection and sequencing of oil wells under reservoir uncertainty. Working paper, Department of Business Administration, Stavanger College, Norway.

Jonsbråten, T.W., Wets, R.J.-B. and Woodruff, D.L. (1998) A class of stochastic programs with decision dependent random elements. *Annals of Operations Research*, **82**, 83–106.

Kao, E.P.C. and Queyranne, M. (1985) Budgeting costs of nursing in a hospital. *Management Science*, **31**, 608–621.

Kenyon, A.S. and Morton, D.P. (2003) Stochastic vehicle routing with random travel times. *Transportation Science*, **37**, 69–82.

King, A.J. and Rockafellar, R.T. (1993) Asymptotic theory for solutions in statistical estimation and stochastic programming. *Mathematics of Operations Research*, **18**, 148–162.

King, A.J. and Wets, R.J.-B. (1991) Epi-consistency of convex stochastic programs. *Stochastics*, **34**, 83–91.

Koop, G.J. (1988) Multiple shift workforce lower bounds. *Management Science*, **34**, 1221–1230.

Lau, H.C. (1996) On the complexity of manpower shift scheduling. *Computers and Operations Research*, **23**, 93–102.

Law, A.M. and Kelton, W.D. (2000) *Simulation Modeling and Analysis*, 3rd edn., McGraw-Hill, Boston, MA.

Lustig, I.J., Mulvey, J.M. and Carpenter, T.J. (1991) Formulating two-stage stochastic programs for interior point methods. *Operations Research*, **39**, 757–770.

Mak, W.K., Morton, D.P. and Wood, R.K. (1999) Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, **24**, 47–56.

Morton, D.P. and Wood, R.K. (1998) On a stochastic knapsack problem and generalizations, in *Advances in Computational and Stochastic Optimization, Logic Programming, and Heuristic Search: Interfaces in Computer Science and Operations Research*, Woodruff, D. (ed.), Kluwer, Boston, MA, pp. 149–168.

Pinedo, M. (1983) Stochastic scheduling with release dates and due dates. *Operations Research*, **31**, 559–572.

Powell, W.B. and Frantzeskakis, L.F. (1994) Restricted recourse strategies for dynamic networks with random arc capacities. *Transportation Science*, **28**, 3–23.

Robinson, S.M. (1996) Analysis of sample-path optimization. *Mathematics of Operations Research*, **21**, 513–528.

Rothkopf, M.H. (1966) Scheduling with random service time. *Management Science*, **12**, 707–713.
Ruszczyński, A. (1986) A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical Programming*, **35**, 309–333.
Ruszczyński, A. and Świetanowski, A. (1997) Accelerating the regularized decomposition method for two stage stochastic linear problems. *European Journal of Operational Research*, **101**, 328–342.
Sevcik, K.C. (1974) Scheduling for minimum total loss using service time distributions. *Journal of the Association for Computing Machinery*, **21**, 66–75.
Shapiro, A. (1991) Asymptotic analysis of stochastic programs. *Annals of Operations Research*, **30**, 169–186.
Smith, W.E. (1956) Various optimizers for single stage production. *Naval Research Logistics Quarterly*, **3**, 59–66.
Süer, G.A. (1996) Optimal operator assignment and cell loading in labor-intensive manufacturing cells. *Computers and Industrial Engineering*, **31**, 155–158.
Van Slyke, R.M. and Wets, R.J.-B. (1969) L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, **17**, 638–663.
Verweij, B., Ahmed, S., Kleywegt, A., Nemhauser, G. and Shapiro, A. (2001) The sample average approximation method applied to stochastic vehicle routing problems: a computational study. Optimization Online, http://www.optimization-online.org. Accessed 10/23/03.
Watkins, Jr., D.W., Morton, D.P. and McKinney, D.C. (1998) Monte Carlo techniques for estimating solution quality in stochastic groundwater management models, in *Proceedings of the XII International Conference on Computational Methods in Water Resources*, Crete, Greece, pp. 67–74.
Wets, R.J.-B. (1974) Stochastic programs with fixed recourse: the equivalent deterministic program. *SIAM Review*, **16**, 309–339.
Wilson, J.G. and Popova, E. (1996) Bayesian approaches to maintenance intervention, in *Proceedings of the Section on Bayesian Science of the American Statistical Association*, Alexandria, VA, pp. 244–248.
Zakeri, G. (2000) MetaNEOS project: VERIFY optimization solver. Accessed at *http://www-unix.mcs.anl.gov/metaneos/* on 10/23/03.

Biographies

David P. Morton is an Associate Professor in the Graduate Program in Operations Research and Industrial Engineering at The University of Texas at Austin. He graduated from Stetson University with a B.S. in Mathematics and Physics and from Stanford University with an M.S. and Ph.D. in Operations Research. His research interests include developing applications, algorithms and theory for solving stochastic optimization models of complex systems that contain significant uncertainties.

Elmira Popova is an Assistant Professor in the Graduate Program in Operations Research and Industrial Engineering at The University of Texas at Austin. She has an M.S. in Mathematics from the University of Sofia, Bulgaria and a Ph.D. in Operations Research from Case Western Reserve University, Cleveland, OH. Her research interests are in the application of statistical, applied probability, management science and computer science techniques to problems in engineering, management, manufacturing and finance.

Contributed by the Engineering Statistics and Applied Statistics Department