

ADAPTIVE STOCHASTIC MANPOWER SCHEDULING

Elmira Popova
David Morton

Graduate Program in Operations Research
and Industrial Engineering
The University of Texas at Austin
Austin, TX 78712, U.S.A.

ABSTRACT

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. Stochastic programs, unlike deterministic optimization models, 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 scheduling problem with uncertain up-times of manufacturing equipment and uncertain production rates.

1 INTRODUCTION

Many of the real-world problems the operations research community addresses require decision making with respect to a system that stochastically evolves over time. Having the right information at the right time increases our ability to make good decisions. However, obtaining the right data is just one of the problems that needs to be solved. For a given *decision making process* (DMP) we are interested in a number of additional issues:

- When is it optimal to gather information?
- Once we have it, how should the information be analyzed for the DMP (this could involve any type of data analysis or even visual inspection)?
- How should the analyzed information be used within the DMP?

- And, what information should we create? (The result of a DMP is typically a decision that should be acted on and this can be regarded as a piece of information.)

This paper attempts to address, at some level, the last three of these questions in the context of a stochastic manpower scheduling problem. Our main goal is to create a methodology that allows information, gathered dynamically over time, to be incorporated in the DMP so that better informed decisions can be made.

We address processes that are nonstationary, i.e., the stochastic evolution of the system changes over time. At each decision point in time, our methodology incorporates the observed history of the process via Bayesian methods and as a result it dynamically “adapts” the associated probability models. Given these adapted models a new decision is made. We repeat this procedure over and over again in a rolling-horizon manner. The proposed methodology combines techniques from Bayesian statistics, stochastic programming and simulation.

The class of problems that we intend to attack has a high-dimensional decision space and the uncertainties (usually) depend on the decision taken. Our approach partially captures decision-dependent randomness (by mixing Bayesian analysis and stochastic optimization) and provides more realistic sampling procedures for the stochastic optimization model (by mixing simulation and Bayesian analysis).

The use of simulation in Bayesian estimation has recently received a great deal of attention. Markov chain Monte Carlo methods allow the computation of Bayesian estimates that otherwise are intractable. Our approach is related, but instead of embedding simulation within a Bayesian estimation procedure we instead embed Bayesian updates within simulation. Specifically, we simulate sample paths from the “predictive distributions” obtained via Bayesian analysis. We illustrate the methodology by

applying it to a stochastic employee scheduling problem in a manufacturing system with uncertain equipment up-times and uncertain production rates.

2 BACKGROUND

Since the beginning of use of simulation models, researchers and practitioners have wanted not only to analyze system performance but also to improve system performance. There is an extensive literature in stochastic optimization and stochastic programming; simulation and optimization; gradient estimation; and Bayesian analysis. We only discuss the work most relevant to ours.

Consider the following stochastic optimization problem

$$z^* = \min_{x \in X} Ef(x, \xi_x), \quad (1)$$

where x is the decision vector, $f(x, \xi_x)$ is the random cost function, ξ_x is a random vector (whose distribution might depend on x), and X is a deterministic set of feasible decisions.

There are two main classes of simulation-based methods for solving (1). The one we employ is based on “external sampling” (also called the “nonrecursive” method, Pflug (1996), “sample-path optimization”, Robinson (1996), and the “stochastic counterpart” method, Shapiro (1991), and it constructs an approximation of (1) by generating, for example, an i.i.d. sequence of random vectors $\xi^1, \xi^2, \dots, \xi^n$ that have the same distribution as ξ (which in this case does not depend on x) to obtain the approximating problem

$$z_n = \min_{x \in X} \frac{1}{n} \sum_{i=1}^n f(x, \xi^i). \quad (2)$$

Results concerning asymptotic properties of (2) allow us to use it as an approximation of (1) when the sample size n is sufficiently large. Specifically, under appropriate assumptions, $\lim_{n \rightarrow \infty} z_n = z^*$, and the limit points of a sequence of optimizers to (2) solve (1), w.p.1; see Dupačová (1991), Dupačová and Wets (1988), King and Rockafellar (1993), King and Wets (1991), Robinson (1996), Shapiro (1991) for these and other related results. Methods for determining when n is large enough have been proposed via estimation of the optimality gap, Dantzig and Infanger (1995), Higle and Sen (1996a), Mak et al. (1997) as well as methods based on statistical verification of the (generalized) Karush-Kuhn-Tucker optimality conditions, see Higle and Sen (1996b) and Shapiro and Homem-de-Mello (1996).

The second class of sampling-based approaches to solving (1) uses “internal sampling.” The sampling is internal because new observations of ξ are generated within the optimization algorithm and only when they are required. Stochastic approximation, e.g., Kiefer

and Wolfowitz (1952), Robbins and Monro (1951), and stochastic quasigradient methods, e.g., Ermoliev (1988) are sampling-based adaptations of deterministic gradient search algorithms. Other related procedures include stochastic adaptations of cutting-plane algorithms, see Dantzig and Glynn (1990), Higle and Sen (1996a) and Infanger (1993).

Gradient estimates of $Ef(x, \xi)$ play a key role in these methods. In stochastic programming, gradient (or subgradient) estimates of $Ef(x, \xi)$ are typically available via duality. For more general stochastic systems, there are two widely used methodologies to compute estimates of $\nabla Ef(x, \xi)$ — infinitesimal perturbation analysis (IPA), Glasserman (1991), Ho and Cao (1991) and the score function (SF) method, Rubinstein and Shapiro (1993), which is also called the likelihood ratio method, Glynn (1990). IPA requires certain structures of the process whereas the SF method is more widely applicable (but often leads to estimators with higher variance). When gradient estimates are unavailable one can resort to methods that use finite differences, Kiefer and Wolfowitz (1952). A survey of Monte Carlo methods applied to solving stochastic optimization problems is provided in Morton and Popova (1998).

Most the literature assumes that the source of randomness, ξ , does not depend on the decision taken, x . 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 decision-dependent probability distributions that may arise.

3 MANPOWER SCHEDULING

We consider a manpower scheduling problem in a production system in which a nominal weekday work schedule must be 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 schedule as closely as possible. However, overtime work for the weekends may be adaptively scheduled during the month. We utilize a Bayesian stochastic programming approach to this problem. Nominal weekday schedules are found by solving a two-stage stochastic program with recourse that is described in Section 3.2. The random parameters in this optimization model are the production rates and machine availabilities. We construct predictive distributions for these elements using Bayesian models described in Section 3.1. Based on observations made during the month, the Bayesian distributions are then updated at the end of the month prior to re-optimizing to find the schedule for the coming month.

There are many authors who analyze scheduling under uncertainty. Optimal strategies for single machine

scheduling are obtained in the papers by Smith (1956), Rothkopf (1966), Sevcik (1974) and Pinedo (1983). Gittins and Glazebrook (1977) and Hamada and Glazebrook (1993) solve Bayesian stochastic job scheduling problems on a single machine. Birge and Dempster (1996) provide a hierarchical framework of scheduling problems from a stochastic programming perspective, emphasizing the distinction between strategic, tactical, and operational models.

Kao and Queyranne (1985) use a two-stage stochastic program for a nurse scheduling problem with uncertain demand. First-stage variables schedule regular-time staffing levels while second-stage adaptive decisions schedule overtime work and the hiring of temporary employees. The model we develop in this paper schedules tasks for employees to perform during three daily shifts over the course of one month while Kao and Queyranne use monthly time periods and a year-long budget cycle. And, the stochastic parameters in our model are machine availabilities and production rates while demand is uncertain in Kao and Queyranne's work. Nevertheless, the notion of recourse we use is similar in spirit to theirs in that we schedule regular-time work with first-stage variables and schedule overtime work with adaptive second-stage decisions.

The system we consider is a production line in a major manufacturer of automobile parts. The production line consists of several cells, each of which produces, or assembles, different parts of the final product. The production equipment consists of a set of lathe machines located in the initial cell of the production line. These machines produce different types of shafts that are necessary components for all the items produced by the production line. Due to the key role that these machines play in the production process, they operate up to 24 hours a day, on 3 shifts, each 8 hours long.

The production line has its own manager. At the end of each month the manager receives the demand forecast for the next month which specifies the required number of shafts per type per week. Because the plant does not have a large warehouse where production can be stored, the line operates almost as a just-in-time system, i.e., finished products are shipped to customers as soon 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 would like 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) as soon as production is completed. If the delay is too long, then the manufacturer must also pay a contractual penalty cost.

There are scheduled down-times for production equipment for reasons such as preventive maintenance or because the crew is training, attending a meeting, or on a break. In addition, lathe machines fail at random times and then require corrective maintenance. Another source of uncertainty involves crew production rates. While the engineering department produces specifications of how many shafts can be produced per hour for each shaft type on each lathe machine, there is considerable variability in the actual production rates, both within a crew over time as well as between crews.

The production-line manager is allocated a budget by upper management for the coming month. 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 shift, the crews fill out a time card describing what was done during the 8 hour shift. A typical example is: 6 hours 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.

An important task that the line manager must perform is to construct a manpower schedule for shaft production for the coming month so that the requested demand for shafts will be met in a timely fashion. The decision maker (manager) is motivated to construct a good schedule. The main goal is to deliver the requested production on time. In practice, the manager decides on an initial schedule which is then adjusted over the course of the month. For instance, if, due to low production rates and machine availabilities, insufficient shifts were scheduled to meet demand, then overtime production shifts would be scheduled for the weekends. The workers' overtime wages are higher than the regular-time wages.

If the allocated production budget is exceeded by more than 2% in one 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 percentage wage increases for production managers and workers as well as to decide how bonuses are distributed when profits are high.

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.

Decision-dependent randomness arises in this setting for the following reason: As described in Section 3.1, at

the end of each 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 (see Section 3.2), in a rolling-horizon fashion. It would require a more sophisticated model in order to explicitly capture the potential advantage of making decisions now that could "gather information" for the future. In work along these lines: 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.

3.1 Bayesian Model for Production Rates and Up-Time of the Equipment

In this section we introduce Bayesian 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 unsatisfactory results. We have empirical evidence that the stochastic behavior of the system is nonstationary and we propose a Bayesian time-dynamic model to capture this nonstationarity.

A total of four months (January–April) of data was collected. For each eight-hour shift we have the number of machine up-hours and the type of shaft and number produced. We model the up-time rather than the downtime. The primary reason for this concerns the nature of the data that was collected. As described above, lathe operators account for each of the eight hours of their shift on a time card. Because of management's emphasis on production accounting, a premium is placed on workers accurately documenting the time spent producing shafts, i.e., the up-time data is very reliable. On the other hand, 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 downtime. During the week, the production equipment is virtually never idle for lack of something to produce. (In fact the high utilization is what lead to the desire for better scheduling.)

The stochastic optimization model defined in Section 3.2 is executed at the end of each month to obtain a manpower 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.

First we describe the Bayesian model for up-hours. From the data cards, it was clear that the up-time values entered by the operators are discrete rather than continuous. This is not surprising as it is natural for the operators to estimate such times to within an hour. As a result, we divide the eight-hour shift into eight one-hour subintervals. Denote by \mathbf{AH}_{cmt} the up-time of the equipment (available hours) during the eight-hour shift of crew c on machine m on day t . We assume a hierarchical Bayesian structure, and model the (random) probability mass function (pmf) of \mathbf{AH}_{cmt} . Let \mathbf{Z} be a random vector with realizations of the form (r_1, \dots, r_8) , where $r_i = 0, 1, 2, \dots$ and $\sum_{i=1}^8 r_i = n$. Here n represents the total number of observations over a month and r_i is the number of instances when the observed up-hours fell into the interval $(i-1, i]$, $i = 1, \dots, 8$. We assume that \mathbf{Z} is distributed as a multinomial random vector with parameters n and $\mathbf{W} = (\mathbf{W}_1, \dots, \mathbf{W}_8)$, where $0 \leq \mathbf{W}_i \leq 1$ and the values of the components of \mathbf{W} are unknown. The pmf for \mathbf{AH}_{cmt} is obtained by dividing each of the components of \mathbf{Z} by n . While the pmf for \mathbf{AH}_{cmt} depends on the shift, machine, and time, we suppress the dependency of \mathbf{Z} and \mathbf{W} on c , m , and t for clarity of the presentation.

Suppose that \mathbf{W} is a Dirichlet random variable with parametric vector $\alpha \equiv 0$ (which is the improper prior density, see 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 , $i = 1, \dots, 8$ observations in each interval. The Dirichlet distribution is a conjugate prior (see for instance DeGroot (1970), p.174) for the multinomial distribution, and as a result, the posterior distribution of \mathbf{W} given r_i , $i = 1, \dots, 8$ is a Dirichlet distribution with parametric vector $(\alpha_1 + r_1, \dots, \alpha_8 + r_8)$. The predictive distribution for \mathbf{Z} is a multinomial-Dirichlet distribution with parameters $(\alpha_1 + r_1, \dots, \alpha_8 + r_8)$ (for details see Bernardo and Smith (1994), p.441. Thus, generating an observation of \mathbf{AH}_{cmt} in a "fully Bayesian" manner involves two steps. First, we generate an observation of the pmf of \mathbf{AH}_{cmt} , i.e., $\frac{1}{n}\mathbf{Z}$, given r_1, \dots, r_8 from the predictive multinomial-Dirichlet distribution. Then, given this instance of the pmf, we generate an observation of \mathbf{AH}_{cmt} . We use a simpler approach that still captures time-dynamic updating. In particular, using the conditional mean of $\frac{1}{n}\mathbf{Z}$ (see Bernardo and Smith (1994), given the previous months observations yields the following

distributional forecast for \mathbf{AH}_{cmt} :

$$P[\mathbf{AH}_{cmt} = i] = \frac{\alpha_i + r_i}{\sum_{j=1}^8 (\alpha_j + r_j)}, i = 1, \dots, 8. \quad (3)$$

In generating observations of the up-hours we sample from (3).

Now we turn to the Bayesian model for production rates. Denote by \mathbf{PD}_{scmt} the hourly production rate for crew shift c of a shaft of type s on machine m on day t . We follow the model given in DeGroot (1970), p.169 and assume that \mathbf{PD}_{scmt} has a normal distribution with unknown mean \mathbf{M} and precision \mathbf{R} . (The precision is the reciprocal of the variance.) Furthermore, let \mathbf{M} and \mathbf{R} have an improper joint prior distribution $\pi(\mathbf{M}, \mathbf{R}) = 1/R, R > 0$. Then, (see DeGroot (1970), p.170) given the observed production rates for one month $pd_{scmt} = (pd_{scmt}^1, pd_{scmt}^2, \dots, pd_{scmt}^m)$, the posterior distribution for \mathbf{M} is 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 \mathbf{R} is Gamma with parameters $(n-1)/2$ and $ns^2/2$. The predictive distribution for \mathbf{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 constructing the Monte Carlo approximation, we actually draw observations from a truncated version of the predictive Student- t to ensure nonnegative rates, i.e.,

$$\max \left\{ t \left(\bar{x}, [(n-1)(n+1)s^2]^{-1}, n-1 \right), 0 \right\}. \quad (4)$$

To assure that we are sampling up-times from a representative distribution we performed 9 Kolmogorov–Smirnov tests (one for each of the 3 shifts in February, March, and April). The null hypothesis states that the difference between observed and predicted up-times does not exceed the differences that would be expected to occur by chance. In all 9 of the tests we could not reject the null hypothesis at 0.05 level of significance. Therefore, we regard the mathematical model (3) from which we sample up-times to feed into the scheduling model as an appropriate one. Similar tests can be performed for the production rates. However, the four-month data set has 35 different types of shafts and three crews; as a result, there are many possible combinations and for some crew-shaft pairs there are a small number of observations.

3.2 Stochastic Programming Formulation

We propose a stochastic programming model for scheduling shaft production — at the level of daily crew shifts — for one month to minimize the expected value of a weighted sum of penalties for late and nondelivered shipments plus a penalty for exceeding the target budget. We adopt the

following assumptions and notation. The days of the month are denoted by the set T . This set is partitioned into $T = \mathcal{R} \cup \mathcal{O}$, where the disjoint sets \mathcal{R} and \mathcal{O} represent weekdays (regular-time) and weekend-days (overtime). As before, the random up-times and production rates are denoted \mathbf{AH}_{cmt} and \mathbf{PD}_{scmt} . The corresponding sample space is denoted Ω , and a realization of the random vector, $(\mathbf{AH}, \mathbf{PD})$, is denoted $(\mathbf{AH}^\omega, \mathbf{PD}^\omega)$.

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 production schedule is specified for Monday–Friday of each week, $[x_{scmt}^{max}]_{t \in \mathcal{R}}$, which states the number of hours crew shift c should spend producing shaft s on machine m on day t . This scheduling decision constitutes the “first-stage” decision because it must be made with only distributional knowledge of the machine up-times and crew production rates. This schedule is subject to the following constraint:

$$\sum_s x_{scmt}^{max} \leq \mathbf{AH}_{cmt}^{max} \quad \forall c, m, t \in \mathcal{R}. \quad (5)$$

Here, \mathbf{AH}_{cmt}^{max} is an optimistic bound on machine availability, i.e., eight hours per crew shift less scheduled down-time for reasons such as preventive maintenance. After this scheduling decision has been made, an observation of machine availabilities and production rates $(\mathbf{AH}_{cmt}^\omega, \mathbf{PD}_{scmt}^\omega)$ 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, $[x_{scmt}^\omega]_{t \in \mathcal{R}}$, (ii) weekend (overtime) production schedules, $[y_{scmt}^\omega]_{t \in \mathcal{O}}$, (iii) unmet demand variables for shipments, $[z_{st}^\omega]_{t \in T}$, and (iv) the amount by which the target budget is exceeded, v^ω .

The actual weekday production schedule, $[x_{scmt}^\omega]_{t \in \mathcal{R}}$, is essentially a scaled version of the nominal schedule which ensures that weekday machine availability constraints, with stochastic availability $\mathbf{AH}_{cmt} = \mathbf{AH}_{cmt}^\omega$, are obeyed. This is effected via

$$x_{scmt}^\omega = \min \left\{ x_{scmt}^{max}, \left[\frac{\mathbf{AH}_{cmt}^\omega}{\mathbf{AH}_{cmt}^{max}} \right] x_{scmt}^{max} \right\} \quad \forall s, c, m, t \in \mathcal{R}. \quad (6)$$

Thus, the weekday schedule, $[x_{scmt}^\omega]_{t \in \mathcal{R}}$, is determined by (6), given the up-time realizations and nominal schedule. 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 \mathbf{AH}_{cmt} is for a “typical” day. Occasionally, \mathbf{AH}_{cmt}^{max} has a small value for irregular but scheduled down-times

and (6) ensures that the actual schedule reflects these unavailabilities. The weekend schedules, $[y_{scmt}^\omega]_{t \in \mathcal{O}}$, are recourse decisions in the usual sense while the shortage variables, $[z_{st}]_{t \in \mathcal{T}}$, and budget exceedance variable v^ω are “accounting” variables from which appropriate late-delivery and up-side budget-deviation penalties are assigned via piecewise-linear convex penalty functions g_{st} and h .

Additional data include regular-time and overtime hourly wages, WR_c and WO_c ; the demand schedule D_{st} ; the production budget B ; and a weighting factor λ for the budget-deviation penalty. The value of λ is selected so that minimizing late shipments is the primary objective and minimizing up-side deviations from the target budget is the secondary objective.

The two-stage stochastic program can be stated as:

$$\begin{aligned} & \min_{x^{max}} Ef(x^{max}, \mathbf{AH}, \mathbf{PD}) \\ & \text{s.t. } \sum_s x_{scmt}^{max} \leq AH_{cmt}^{max} \quad \forall c, m, t \in \mathcal{R} \\ & \quad x_{scmt}^{max} \geq 0 \quad \forall s, c, m, t \in \mathcal{R}, \end{aligned} \quad (7)$$

where

$$\begin{aligned} & f(x^{max}, \mathbf{AH}, \mathbf{PD}) = \\ & \min_{x, y, z, v} \sum_{s, t} g_{st}(z_{st}) + \lambda h(v) \\ & \text{s.t. } x_{scmt} \leq \left[\frac{AH_{cmt}}{AH_{cmt}^{max}} \right] x_{scmt}^{max} \quad \forall s, c, m, t \in \mathcal{R} \\ & \quad \sum_{c, m, t' \leq t} PD_{scmt'} x_{scmt'} + \\ & \quad \sum_{c, m, t' \leq t} PD_{scmt'} y_{scmt'} + z_{st} \geq D_{st} \quad \forall s, t \\ & \quad \sum_s y_{scmt} \leq AH_{cmt} \quad \forall c, m, t \in \mathcal{O} \\ & \quad \sum_{s, c, m, t} WR_c x_{scmt} + \sum_{s, c, m, t} WO_c y_{scmt} - v \leq B \\ & \quad x_{scmt} \leq x_{scmt}^{max} \quad \forall s, c, m, t \in \mathcal{R} \\ & \quad x_{scmt}, y_{scmt}, z_{st}, v \geq 0 \quad \forall s, c, m, t. \end{aligned} \quad (8)$$

An external Monte Carlo sampling-based approximation of (7) is solved to obtain a manpower schedule to be posted by the manager.

3.3 Bayesian Predictive Distributions and Stochastic Optimization

Here we consider a solution methodology that generates the production schedules from the stochastic program using the Bayesian predictive distributions. We can compactly state the stochastic optimization model as

$$z^* = \min_{x \in X} Ef(x, \xi) \quad (9)$$

where x denotes the first-stage weekday production schedule, $[x_{scmt}^{max}]_{t \in \mathcal{R}}$; $x \in X$ denotes the constraints of (7); $\xi = (\mathbf{AH}, \mathbf{PD})$ represents 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 a specific realization of the random up-time and production rate vector, as defined in (8). It is not possible to solve (9) exactly and we must resort to approximations.

We use the external sampling approach described in Section 2. In our case, the expectation in (9) is with respect to the Bayesian distributions of machine up-time (3) and crew production rates (4), given the data observed up to that point in time. We sample i.i.d. variates ξ^1, \dots, ξ^n from the Bayesian distribution of ξ and form the approximating problem

$$z_n = \min_{x \in X} \frac{1}{n} \sum_{i=1}^n f(x, \xi^i). \quad (10)$$

We assumed a conjugate structure and hence eliminated any computational difficulties (such as multidimensional integration) in generating these samples. However, if one were to assume general distributions for machine up-times, production rates and their parameters, the above methodology may still be applied.

4 CONCLUSIONS AND FUTURE WORK

In this paper we have presented a methodology which dynamically incorporates information, as it becomes available, into the scheduling process. In a forthcoming paper we will compare the quality of the nominal regular-time production schedules obtained using four forecasting procedures: (i) an empirical point forecast, (ii) a Bayesian point forecast, (iii) an empirical distribution, and (iv) a Bayesian distributional forecast. Specifically, we will perform an all pairwise comparison for $z_{ep} = Ef(x_{ep}, \xi)$, $z_{ed} = Ef(x_{ed}, \xi)$, $z_{bp} = Ef(x_{bp}, \xi)$, and $z_{bd} = Ef(x_{bd}, \xi)$. Here, x_{ep} , x_{ed} , x_{bp} , and x_{bd} denote the respective nominal schedules based on the four forecasting techniques and the expectations are taken with respect to the Bayesian distributions since these represent our best forecasts of how the system will behave. Our goal is to investigate the value of using Bayesian forecasting and stochastic programming over classical statistical methods and deterministic optimization.

ACKNOWLEDGMENTS

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. David Morton’s research was partially supported by the National Science Foundation through

grant DMI-9702217. Elmira Popova's research was partially supported by grant #0003658-472 from the State of Texas Advanced Technology Program.

REFERENCES

- Artstein Z. and R.J.-B. Wets. 1993. Sensors and information in optimization under stochastic uncertainty. *Mathematics of Operations Research*. 18:523–547.
- 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 A.F.M. Smith. 1994. *Bayesian Theory*. John Wiley and Sons. Chichester.
- Birge J.R. and M.A.H. Dempster. 1996. Stochastic programming approaches to stochastic scheduling. *Journal of Global Optimization*. 9:417–451.
- Dantzig G.B. 1955. Linear programming under uncertainty. *Management Science*. 1:197–206.
- Dantzig G.B. and P.W. Glynn. 1990. Parallel processors for planning under uncertainty. *Annals of Operations Research*. 22:1–21.
- Dantzig G.B. and G. Infanger. 1995. A probabilistic lower bound for two-stage stochastic programs. Technical Report SOL 95-6. Department of Operations Research. Stanford University.
- DeGroot M.H. 1970. *Optimal Statistical Decisions*. McGraw-Hill. New York.
- Dupačová J. 1991. On non-normal asymptotic behavior of optimal solutions for stochastic programming problems and on related problems of mathematical statistics. *Kybernetika*. 27:38–52.
- 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.
- Ermoliev Y. 1988. Stochastic quasigradient methods. In Y. Ermoliev and R.J.-B. Wets editors. *Numerical Techniques for Stochastic Optimization*. pages 141–185. Springer-Verlag. Berlin.
- Futschik A. and G.Ch. Pflug. 1997. Optimal allocation of simulation experiments in discrete stochastic optimization and approximative algorithms. *European Journal of Operational Research*. 101:245–260.
- Gittins J.C. and K.D. Glazebrook. 1977. On Bayesian models in stochastic scheduling. *Journal of Applied Probability*. 14(5):556–565.
- Glasserman P. 1991. *Gradient Estimation via Perturbation Analysis*. Kluwer Academic Publishers. Boston.
- Glynn P.W. 1990. Likelihood ratio gradient estimation for stochastic systems. *Communications of the ACM*. 33(10):75–84.
- Hamada T. and K.D. Glazebrook. 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(5):924–934.
- Higle J.L. and S. Sen. 1996a. Duality and statistical tests of optimality for two stage stochastic programs. *Mathematical Programming*. 75:257–275.
- Higle J.L. and S. Sen. 1996b. *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*. Kluwer Academic Publishers. Dordrecht.
- Ho Y.C. and X.R. Cao. 1991. *Perturbation Analysis of Discrete Event Dynamic Systems*. Kluwer Academic Publishers. Boston.
- Infanger G. 1993. *Planning Under Uncertainty: Solving Large-Scale Stochastic Linear Programs*. The Scientific Press Series. Boyd & Fraser. Danvers, Mass.
- Jonsbråten T.W. 1997. Optimal selection and sequencing of oil wells under reservoir uncertainty. working paper.
- Jonsbråten T.W. and R.J.-B. Wets and D.L. Woodruff. 1998. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research*. to appear.
- Kao E.P.C. and M. Queyranne. 1985. Budgeting costs of nursing in a hospital. *Management Science*. 31(5):608–621.
- Kiefer J. and J. Wolfowitz. 1952. Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics*. 23:462–466.
- King A.J. and R.T. Rockafellar. 1993. Asymptotic theory for solutions in statistical estimation and stochastic programming. *Mathematics of Operations Research*. 18:148–162.
- King A.J. and R.J.-B. Wets. 1991. Epi-consistency of convex stochastic programs. *Stochastics*. 34:83–91.
- Mak W.K. and D.P. Morton and R.K. Wood. 1997. Monte Carlo bounding techniques for determining solution quality in stochastic programs. Technical report. Graduate Program in Operations Research. The University of Texas at Austin.
- Morton D. and E. Popova. Monte Carlo simulations for stochastic optimization. 1998. In C.A. Floudas and P.M. Pardalos editors. *Encyclopedia of Optimization*. Kluwer Academic Publishers. forthcoming.
- Pflug G.Ch. 1996. *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*. Kluwer Academic Publishers. Boston.
- Pinedo M. 1983. Stochastic scheduling with release dates and due dates. *Operations Research*. 31:559–572.
- Powell W.B. and L.F. Frantzeskakis. 1994. Restricted recourse strategies for dynamic networks with random arc capacities. *Transportation Science*. 28:3–23.

- Robbins H. and S. Monro. 1951. A stochastic approximation method. *Annals of Mathematical Statistics*. 22:400–407.
- 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.
- Rubinstein R.Y. and A. Shapiro. 1993. *Discrete Event Systems: Sensitivity and Stochastic Optimization by the Score Function Method*. John Wiley & Sons. Chichester.
- 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.
- Shapiro A. and T. Homem-de-Mello. 1998. A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming*. 81:301–325.
- Smith W.E. 1956. Various optimizers for single stage production. *Naval Research Logistics Quarterly*. 3:59–66.
- Wets R.J.-B. 1974. Stochastic programs with fixed recourse: the equivalent deterministic program. *SIAM Review*. 16:309–339.

AUTHOR BIOGRAPHIES

ELMIRA POPOVA is an Assistant Professor in the OR/IE Group at The University of Texas at Austin. She received an M.S. in Mathematics from University of Sofia, Bulgaria and a Ph.D. in Operations Research from Case Western Reserve University. Her research interests are in stochastic modeling, Bayesian statistical analysis, and application of applied probability to problems in manufacturing, reliability, and finance.

DAVID MORTON is an Assistant Professor in the OR/IE Group at The University of Texas at Austin. He received a Ph.D. in Operations Research from Stanford University. His research interests include computational stochastic optimization.