Optimizing Seasonal Firefighting Resource Allocations with Random Consequences: A Stochastic Integer and Chance-Constrained Programming Estimation Approach

Michael Bevers
USDA Forest Service
Rocky Mountain Research Station
2150A Centre Avenue, Suite 361
Fort Collins, Colorado USA 80526
Phone (970) 295-5911
Fax (970) 295-5959
Email mbevers@fs.fed.us

January 18, 2006

In review: Canadian Journal of Forest Research

Optimizing Seasonal Firefighting Resource Allocations with Random Consequences: A Stochastic Integer and Chance-Constrained Programming Estimation Approach

Abstract

Recent studies for optimally deciding seasonal allocations of fire fighting resources suggest that stochastic integer programming and chance-constrained integer programming approaches should be explored. These problems can be very difficult to solve. This study develops estimation models that might be useful and more tractable than exact formulations, particularly considering the quality of information often available and the often high cost of obtaining better information. The estimation models are tested with a heuristic algorithm developed for this study that borrows from more rigorous statistical procedures used in ranking and selecting of alternatives based on simulation optimization. Results from two experiments indicate that these estimation models are potentially useful and warrant further investigation.

Introduction

Deciding the best numbers and station placements of wildfire suppression resources for a fire season is an important and complex problem. Both the relatively deterministic costs of acquiring and stationing seasonal resources, and the highly random costs of actual firefighting and resulting losses, can be expensive. Optimization modeling approaches to this problem date back at least to the work by Parks (1964), but available methods tend to have difficulty either addressing the scale of planning presented by large jurisdictions with numerous resources to manage (Boychuck and Martell 1988) or coping with the stochastic nature of the problem.

Boychuck and Martell (ibid.) were able to roughly evaluate the seasonal requirements for the province of Ontario using a Markov chain model that accounted for some random effects. For a related problem, Mees and Strauss (1992) used integer programming models of expected utility based on constant probability coefficients to allocate a fixed set of resources to various fireline segments. Both of these methods at least partially address random effects, but large and often intractable models still result, especially when random outcomes are interdependent. The expected utility approach of Mees and Strauss also presents complications if fire managers behave risk-aversely (see Hardaker et al. 1997), as Mees and Strauss suggest might happen when the set of available resources is not fixed a priori.

Optimization models that address large-scale problems have more often been deterministic. For example, Donovan and Rideout (2003) present an integer programming model used to identify the most efficient organization for a given fire, which Kirsch and Rideout (2005) have adapted and extended to address the larger-scale problem of planning the initial attack organization for multiple fires (e.g., a fire season). Although both formulations are structured much like typical stochastic programming problems with simple recourse, where one set of variables models possible decisions and a second set of variables models the resulting random consequences, both studies treat the problems deterministically.

The models presented by Donovan and Rideout (ibid.) and by Kirsch and Rideout (ibid.) suggest that formulating large-scale seasonal fire organization decision problems as stochastic integer programming problems (to optimize expected value objectives) or as chance-constrained integer programming problems (to optimize more risk-averse objectives) should be explored. Solving these types of problems, however, can be quite challenging (Birge and Louveaux 1997). In this paper, I begin examining this approach by defining a simple test problem small enough to be solved by enumeration. I then reformulate the full random variable problem into separate instantiated integer programming problems and investigate the performance of a heuristic algorithm that combines simulation with optimization of the separate problems to estimate solutions for the full problem. The focus here is on how to simplify solving these kinds of problems rather than on how to formulate the fire planning problem per se. Because more realistic problems might often prove difficult to solve with algorithms that converge to exact

solutions, testing estimation methods and heuristic solution algorithms seems a useful step toward optimizing stochastic integer fire planning problems.

Methods

Expected Value Problems

We consider a fire planning unit with nine subunits where the problem is to allocate one of three possible fire organizations to each subunit so as to minimize either the expected cost of fires for the fire season, or the seasonal cost associated with a pre-specified, more risk-averse probability level (i.e., a "probable" or "chance-constrained" cost). The expected value programming problem (Problem EVP) is formulated and revised below; formulations for the chance-constrained programming problem (beginning with Problem CCP) follow later:

<u>Problem EVP</u>

[1] $$\text{Minimize} \quad Y$$

subject to:

[2] $$Y - \sum_i \sum_j \phi_{ij} X_{ij} \geq 0$$

[3] $$\sum_j X_{ij} = 1 \qquad \forall\, i$$

[4] $$X_{ij} \in \{0, 1\} \qquad \forall\, i, j$$

[5] $$\sum_i \sum_j c_{ij} X_{ij} \leq b$$

where:

i indexes the $I = 9$ subunits of the fire planning unit,

j indexes the $J = 3$ fire organizations being considered for each subunit,

ϕ_{ij} is the mean variable cost for the fire season in subunit i given fire organization j,

c_{ij} is the fixed seasonal cost of using fire organization j in subunit i,

b is the budget for fixed seasonal cost available to the planning unit,

X_{ij} is a binary decision variable set to one when fire organization j is selected for subunit i (and set to zero otherwise), and

Y is a simple recourse variable representing the total expected (or mean) variable cost for the fire season.

Equation [1] minimizes the expected seasonal total of variable fire costs (referred to hereafter as "fire costs") for the planning unit, as summed across subunits and fire organization choices in Eq. [2]. Equations [3] and [4] require that each subunit be allocated exactly one of the three fire organizations being considered for that subunit. Equation [5] requires that the total cost of employing the nine selected fire organizations (referred to hereafter as "budgeted costs") not exceed the available fixed cost budget. The budgeted costs constrained in Eq. [5] would often be minimized in Eq. [1] along with expected fire costs in stochastic programming problems; Eq. [5] is used here instead so we can examine results for two different budget levels. The optimal fire organization decision $\boldsymbol{X}_{\mathbf{EVP}}$* resulting from Problem EVP is based on knowledge of the expected fire

costs for the season (although actual fire costs in any one season might differ considerably).

Problem EVP is straight-forward if the mean (ϕ_{ij}) of the fire costs for the season can be estimated exogenously and entered directly into the problem. Where a complex system of constraints affects expected values, as in Kirsch and Rideout (2005), and might preclude knowing them a priori, one alternative is to estimate the expected values endogenously, as in Problem SEVP below:

<u>Problem SEVP</u>

[6] $$\text{Minimize} \quad \sum_{n=1}^{N} \frac{1}{N} Y_n$$

subject to:

[7] $$Y_n - \sum_i \sum_j f_{ijn} X_{ij} \geq 0 \qquad \forall\, n$$

[8] $$\sum_j X_{ij} = 1 \qquad \forall\, i$$

[9] $$X_{ij} \in \{0, 1\} \qquad \forall\, i, j$$

[10] $$\sum_i \sum_j c_{ij} X_{ij} \leq b$$

where n indexes samples f_{ijn} of the random variables (F_{ij}) representing fire costs for the season in each subunit under each organization. The optimal fire organization decision ($\boldsymbol{X}_{\mathbf{SEVP}}$*) is now based on knowledge of the sample outcomes (hence, imperfect

knowledge of the expected values); i.e., the decision is only optimal with respect to the samples drawn. Accuracy in finding the true optimal solution of interest ($\boldsymbol{X}_{\mathbf{EVP}}$*) is sacrificed to an unknown extent, and an additional computing cost is incurred due to the expansion from Eq. [2] with dimension 1 to Eq. [7] with dimension *N*.

When total expected cost is a function of numerous constraints that create a large problem, sample size *N* might have to be kept small. In practice, optimal solutions to this estimated problem could easily miss the mark of identifying an optimal solution to Problem EVP. A further step then is to solve random replications of Problem SEVP, referred to below as Problem SEVP-*k*:

<u>Problem SEVP-*k*</u>

[11] $$\text{Minimize} \quad \sum_{n=1}^{N} \frac{1}{N} Y_{nk}$$

subject to:

[12] $$Y_{nk} - \sum_{i} \sum_{j} f_{ijnk} X_{ijk} \geq 0 \qquad \forall\, n$$

[13] $$\sum_{j} X_{ijk} = 1 \qquad \forall\, i$$

[14] $$X_{ijk} \in \{0, 1\} \qquad \forall\, i, j$$

[15] $$\sum_{i} \sum_{j} c_{ij} X_{ijk} \leq b$$

Problem SEVP-*k* is identical to Problem SEVP except that index *k* has been added to specify random replications of the Problem. This highlights that each replication *k* produces one estimated solution. Simulations, statistical procedures, and/or heuristic methods might then used to select an estimated global optimum from the *K* replications examined. Problem SEVP-*k* is the building block of the expected value optimization approach tested in this study.

Chance Constraint Problems

We noted above that after selecting and implementing the fire organizations that minimize expected fire costs, actual fire costs for a given season (such as the upcoming fire season) might differ considerably. For example, in the tests reported here based on lognormal distributions of fire costs in each subunit, observed seasonal fire costs exceeded average seasonal fire costs in about 40 percent of the cases simulated. Had fire costs been normally instead of lognormally distributed, average fire costs would have been exceeded in about 50 percent of the cases. Risk-averse fire managers might prefer to minimize and plan for cost levels that have a smaller probability of exceedance. Chance-constrained programming is one method for optimizing such "probable" rather than "expected" outcomes (Bevers and Kent, in review). Problem CCP, formulated below, states our fire organization planning problem as a chance-constrained programming problem:

<u>Problem CCP</u>

[16] $$\text{Minimize } B$$

subject to:

[17] $$\Pr(Y > B) < p$$

[18] $$Y = \mathrm{f}(\boldsymbol{X}, \boldsymbol{F})$$

[19] $$\sum_j X_{ij} = 1 \qquad \forall\, i$$

[20] $$X_{ij} \in \{0, 1\} \qquad \forall\, i, j$$

[21] $$\sum_i \sum_j c_{ij} X_{ij} \leq b$$

The chance constraint (Eq. 17) sets an upper bound B on total fire cost Y and requires that the probability (Pr) of exceeding that cost be less than parameter p, an accepted level of risk (e.g., 0.05). The resulting probable fire cost B is minimized in Eq. [16]. Because Eqs. [17] and [18] can make solving Problem CCP quite difficult, we are again interested in using endogenous estimation to simplify the problem. To that end, order statistics will be developed as a substitute for chance constraints in the problem that follows (see Fuessle et al. 1987 for another use of order statistics in chance-constrained programming).

We observe that in each replicate of Problem SEVP-k, the objective function (Eq. 11) is not just an estimate of the minimum expected fire cost that can be achieved, it is also (trivially) an estimate of the minimum expected value of the first order statistic of fire cost from an order statistic sample of size one. In Problem SCCP-k (described and

explained below), this idea is extended to allow estimation of the minimum expected value of the *M*th order statistic. Our motivation is that as *M* increases, the resulting observed fire cost and the cumulative probability associated with that fire cost both tend to increase; we will be minimizing a fire cost in the right-hand tail of the distribution where the probability of exceeding that fire cost becomes smaller. Although we do not know beforehand the probability of exceeding the fire cost associated with the *M*th order statistic, that probability can be estimated with post-optimization simulations of the resulting decision vector. Experiment 2 in the Results section will help demonstrate this concept. Problem SCCP-*k* is formulated below:

<u>Problem SCCP-*k*</u>

[22] $$\text{Minimize} \quad \sum_{n=1}^{N} \frac{1}{N} Y_{nk}$$

subject to:

[23] $$Y_{nk} - \sum_{i} \sum_{j} f_{ijmnk} X_{ijk} \geq 0 \qquad \forall\, m, n$$

[24] $$\sum_{j} X_{ijk} = 1 \qquad \forall\, i$$

[25] $$X_{ijk} \in \{0, 1\} \qquad \forall\, i, j$$

[26] $$\sum_{i} \sum_{j} c_{ij} X_{ijk} \leq b$$

where *m* now indexes the set of observations used for each of the *N* independent observations of the *M*th order statistic. The dimension of Eq. [12] from Problem SEVP-*k*

is multiplied by M in Eq. [23], and the expected value estimated in Eq. [22] is now the expected value of the Mth order statistic of fire cost for the season.

Note that Problem SCCP-k adds a third level to our hierarchical sampling effort: 1) we take independent samples indexed by m up to some number M to obtain each observation of the Mth order statistic of fire cost; 2) we make multiple observations of the Mth order statistic to endogenously estimate the expected value of that order statistic using the N overarching samples indexed by n; 3) we make one independent observation of the minimum possible expected value of the Mth order statistic of fire cost for the season with the solution of each replicate, indexed by k, of Problem SCCP-k. The instantiated estimation Problems SEVP-k and SCCP-k are easy to build and to solve, as would be many more realistic problems. How the number K of these problems to solve is determined in our test case, and how a single solution is chosen from the set of K resulting solutions, are addressed in the section that follows.

Before proceeding, we also note that the expected value estimation Problem SEVP-k is simply the special case of the expected order statistic value estimation Problem SCCP-k where $M = 1$. Hereafter, we use only Problem SCCP-k, relying on the setting for parameter M to indicate whether we are searching for solutions to a stochastic integer (expected value) programming problem ($M = 1$) or for solutions to a chance-constrained integer programming problem ($M > 1$).

A Heuristic Solution Algorithm

The heuristic solution procedure used in this study is one of many possible procedures that might be used (e.g., see Reeves 1993). As a heuristic method, this algorithm makes no attempt to run until it converges on a true optimal solution (i.e., a solution to either Problem EVP or to Problem CCP, whichever is intended by the magnitude of parameter *M* in our estimation Problem SCCP-*k*). Instead, the algorithm continues searching by constructing and solving Problem SCCP-*k* replicates until the estimated probability (τ) of finding a better solution with the next replicate decreases to a pre-specified stopping point (τ_0). The steps of the algorithm are as follows:

<u>Algorithm SCCP-*k*</u>

Step 0: Initialize the best objective function value found (μ^) to an arbitrarily large number. Initialize the stack of best solution vectors found and the stack of discarded solution vectors to zeroes; initialize to zeroes counters for each decision variable, a counter *k*^ for the number of replications tested since finding the best solution examined, and the counter *k* for the number of the replication currently being examined.

Set: min~*k* = the minimum number of Problem SCCP-*k* replications to examine,

max~*k* = the maximum number of replications to examine,

τ_0 = the probability point for terminating the search, where

τ is the probability of finding a better solution with the next replication,

α = an acceptable probability of rejecting hypothesis H_0: $\tau = \tau_0$ when H_0 is true (the probability of committing a Type I error),

ß = an acceptable probability of accepting H_0 when it is false (the probability of committing a Type II error),

$\tau_0 + \varepsilon$ = the probability of finding a better solution assumed true for computing 1 – ß, the probability of accepting H_1: $\tau > \tau_0$ when $\tau = \tau_0 + \varepsilon$ (with $\varepsilon > 0$),

δ = the width of an interval of indifference, or tolerance, used as a confidence interval to distinguish estimated objective function values (μ_k),

l = the confidence level required for interval estimation of μ_k such that $\Pr(\mu_k^* - \delta/2 \leqslant \mu_k \leqslant \mu_k^* + \delta/2) \geqslant l$ where μ_k is the true objective function value for replicate k and μ_k^* is the observed objective function value,

M and N.

Step 1: Set $k = k + 1$. Build and solve replicate Problem SCCP-k. Increment counters by one for the selected decision variables.

Step 2: Check stacks for decision vector $\boldsymbol{X}_k^*$. If found, go to Step 5.

Step 3: Simulate the selected decision vector $\boldsymbol{X}_k^*$ until confidence level l is

achieved for the interval specified around objective function value μ_k.

Step 4: Test μ_k:

If $(\mu_k^* + \delta/2) < (\mu^\wedge - \delta/2)$ then μ_k is declared a better solution than the previous best solution $\mu^\wedge$.

Set: $k^\wedge = 1$, $\mu^\wedge = \mu_k^*$.

Move decision vectors from the best solution stack to the discarded solution stack.

Place $\boldsymbol{X}_k^*$ on the best solution stack.

Else if $(\mu_k^* - \delta/2) \leqslant (\mu^\wedge + \delta/2)$ then μ_k is declared an alternate best solution.

Set: $k^\wedge = k^\wedge + 1$.

Place $\boldsymbol{X}_k^*$ on the best solution stack.

Else μ_k is declared a suboptimal solution.

Set: $k^\wedge = k^\wedge + 1$.

Place $\boldsymbol{X}_k^*$ on the discarded solution stack.

Step 5: Construct a new decision vector $\boldsymbol{X}_k$, prioritizing allocations based on the counts recorded for each decision variable from the preceding k solutions. Repeat Steps 2 – 4 for $\boldsymbol{X}_k$ in place of $\boldsymbol{X}_k^*$, but without returning to Step 5 and without incrementing $k^\wedge$ if $\boldsymbol{X}_k$ is not a better solution.

Step 6: If $k^\wedge \geqslant 100$, test hypothesis H_0 as described below.

If $(k < \min{\sim}k)$ or if $((k < \max{\sim}k)$ and (H_0 tests false or was untested)), go to Step 1.

Step 7: Accept $\mu^\wedge$ as a qualified estimate of the true global optimal solution (μ^*),

as described below.

Simulate the decision vectors from the best solution stack to estimate *p*, the probability that μ^ will be exceeded in any given fire season.

Report those decision vectors and probability estimates along with μ^.

Stop.

The decision variable counters incremented in Step 1 and used in Step 5 to construct a new decision vector are intended to account for the possibility in large problems that none of the replicates may actually produce a globally optimal decision vector. For some problems, the larger number of times some decision variables are selected relative to others might be indicative that those variables belong in the optimal decision.

Maintaining a stack of the discarded decision vectors, as well as the set of best decision vectors found, allows the current decision vector to be checked against all previously examined decision vectors in Step 2. This prevents unnecessarily repeating simulations to estimate the objective function, which might be quite time-consuming depending on the indifference interval and the confidence level specified for Step 3.

The test in Step 4 for whether a better, worse, or equivalent solution has been found is relatively strict. Ordinarily, a two-sample t-test might be used, but that test can be problematic if the variances differ for the two solutions. Here, each of the two solutions is simulated in batches of 100. The sample mean from each batch is used as a single

observation, and the distribution of sample means is assumed to be approximately normal. Batches of simulations continue until the probability that the true mean value is within a ± δ/2 interval of the overall sample mean is at least confidence level *l*. At least 200 batches are simulated and z-values of the standard normal distribution are used to approximate Student t-statistics for determining the confidence interval achieved with each new batch. The test in Step 4 then establishes that new objective function values declared to be either superior or inferior to the previous best value are smaller or larger, respectively, by more than δ with confidence l^2. Solutions that do not differ by more than δ (with confidence l^2) are treated as being equally good; i.e., if optimal, they are alternate optima.

The decision to terminate the search for better solutions is based on the likelihood of finding a better solution given another replication effort. Each time a solution better than the previous best solution is found, the probability of finding an even better solution decreases and the magnitude reached by counter *k*^ before finding that "even better solution" tends to increase. Following each replication, *k*^ is treated as the size of a sample from the binomial distribution for which the observations are one success (i.e., a better solution was found) and *k*^ - 1 failures (i.e., a better solution was not found). At *k*^ ⩾ 100, we use the approximation:

$$W = \frac{1-(k\hat{})(\tau_0)}{\sqrt{(k\hat{})(\tau_0)(1-\tau_0)}} \sim \mathrm{N}(0,1)$$

to test H_0.

Type II errors pose the greater concern for this test because incorrectly accepting H_0 leads to stopping the search prematurely. Consequently, power of the test is calculated as:

$$1 - ß\hat{} = \int_{t_{\alpha, k\hat{}-1}}^{\infty} f_{T\Delta}(t)\, dt$$

where T_Δ is a noncentral T variable with $k\hat{}$ - 1 degrees of freedom and noncentrality parameter:

$$\gamma = (k\hat{})(\varepsilon)\sqrt{k\hat{}}\ .$$

When $k\hat{}$ is sufficiently large that $ß\hat{} \leqslant ß$, hypothesis H_0 is accepted as true and the search terminates in Step 6 unless the number of replications has not yet reached min~k, a pre-specified lower limit. Alternatively, the search also terminates in Step 6 if the number of replications has reached max~k, a pre-specified upper limit.

The idea here is that the best solutions reported in Step 7 are accepted, whether or not they are truly optimal, on the basis that more searches are too unlikely to find better solutions to be worth the effort. Whether the reported solutions are truly equivalent or are merely quite similar in performance is treated as a matter of indifference to the decision maker because results are similar enough to be considered equivalent. While this approach borrows ideas from the literature on simulation optimization methods (e.g., Boesel et al. 2003), it is important to note that Algorithm SCCP-k stops short of rigorously finding optimal solutions, or of even demonstrating that the solutions reported

are really the best of those examined (the purpose in Boesel et al.). There are at least two reasons why this more lax approach might be reasonable for wildland fire planning purposes. In contrast to many industrial engineering problems, such as choosing between alternative circuit designs, where basic information is often readily available and highly reliable (i.e., only the integrated performance is particularly uncertain) and where the product selected will be used many, many times, even basic information for fire planning can be uncertain and difficult to obtain and the selected decision might be implemented only once or a few times before dynamics have altered the system enough that a new decision is required.

Results

To introduce some symmetry to our test problem that will intentionally create cases that have alternate optimal solutions, we treat the subunits of the fire planning unit as square areas arranged in a 3x3 grid. Subunits are identified as 1 through 9 from upper left to lower right, so that subunit 3 is in the upper right-hand corner and subunit 7 is in the lower left-hand corner of the grid. Three alternative decisions are possible for each subunit, and decisions for the entire planning unit are described by vectors ranging from $\boldsymbol{X} = (1, 1, 1, 1, 1, 1, 1, 1, 1)$, where all subunits are assigned the least expensive alternative, to $\boldsymbol{X} = (3, 3, 3, 3, 3, 3, 3, 3, 3)$, where all subunits are assigned the most expensive alternative. We use this notation here for convenience; the actual decision vectors have 27 binary (0-1) elements. The budget cost on each subunit for alternative 1 is \$1 million, the cost for alternative 2 is \$1.5 million, and the cost for alternative 3 is \$2

million. The distribution of fire cost for the season on each subunit is a lognormally distributed function of the selected alternative (e.g., see Strauss et al. 1989). These distributions are described by the mean and variance of the underlying normal distributions. Under alternative 2 the underlying mean parameter for fire cost is 12.6 and the standard deviation is 0.8 for each subunit. Random fire costs by subunit are then generated for that alternative as:

$$f_{i,(j\ =\ 2),\,mnk} = e^{(12.6+0.8z)}$$

where z is a random deviate from the standard normal distribution. Under alternative 1, the means and standard deviations of the underlying normal distributions increase by 10 percent; under alternative 3 they decrease by 10 percent. Seasonal fire costs are spatially correlated in such a way that the correlations of z-values between subunits that share a common edge are 0.40, the correlations between subunits that share only a common vertex are 0.15, and no correlations exist between subunits that are not adjacent.

Johnson and Kotz (1970) provide a formula that gives expected values of lognormal random variables based on the parameters of the underlying normal distribution. Using that formula, the expected value of fire cost for the entire planning unit was directly computed for each of the 3^9 = 19,683 possible decisions as the sum of the expected values computed for each subunit. These calculations revealed that the optimal budget level, i.e., the largest budget for which the incremental benefit exceeded the incremental cost, is $13.5 million. The sole optimal decision vector based on expected value is $\boldsymbol{X}$ =

(2, 2, 2, 2, 2, 2, 2, 2, 2) and the expected fire cost for the planning unit is about $3.7 million. These calculations also revealed that $13 million offers another interesting budget level to examine. The nine decision vectors ranging from $\boldsymbol{X}$ = (1, 2, 2, 2, 2, 2, 2, 2, 2) to $\boldsymbol{X}$ = (2, 2, 2, 2, 2, 2, 2, 2, 1) are all optimal based on expected value, as we might anticipate because spatial correlations have no impact on expected values. The expected fire cost under these alternatives is about $4.8 million.

No simple formula is available for the variance of the sum of lognormal random variables, which affects optimal chance-constrained solutions. Instead, 10 million simulations were used to estimate the overall standard deviation of fire cost for all alternatives with budget levels of $13 or $13.5 million. The simulations for a budget level of $13.5 million revealed the decision vector $\boldsymbol{X}$ = (2, 2, 2, 2, 2, 2, 2, 2, 2) has the smallest variance as well as the smallest expected value among those planning alternatives, indicating that vector will also be optimal for any chance-constrained problem we define. The simulations for a budget level of $13 million revealed that four of the nine decision vectors that are optimal based on expected value also have the smallest variance among those planning alternatives. Those four vectors, $\boldsymbol{X}$ = (1, 2, 2, 2, 2, 2, 2, 2, 2), $\boldsymbol{X}$ = (2, 2, 1, 2, 2, 2, 2, 2, 2), $\boldsymbol{X}$ = (2, 2, 2, 2, 2, 2, 1, 2, 2), and $\boldsymbol{X}$ = (2, 2, 2, 2, 2, 2, 2, 2, 1), will be optimal for any chance-constrained problem we define. Because spatial correlations affect overall variance, reduced staffing in the corner subunits is preferable to reduced staffing elsewhere when variance is considered.

Heuristic Search Results

463

464 Experiment 1

465

466 We begin by testing the effect of sample size N when $M = 1$ for a case with fairly
467 stringent settings for the heuristic algorithm parameters: min~k = 100, max~k =5000, α =
468 0.05, ß = 0.01, τ_0 = 0.0001, ε = 0.00001, $\delta = (0.001)(\mu_k^*)$, and l = 0.99. With the budget
469 parameter b = \$13.5 million and N = 1, the search concluded after 9 hours and 2 minutes
470 at K = 2244. The optimal decision was constructed from decision variable counts and
471 correctly identified (i.e., it was the last "better" solution found) at k = 13; the rest of the
472 search was required to satisfy the requirements ß = 0.01, $\tau_0 + \varepsilon = 0.00011$. The search
473 time was quite long because many different decision vectors were found and had to be
474 simulated at length to estimate with fairly high precision the resulting objective function
475 values. The objective function value and p (the probability of exceedance) were
476 estimated at \$3.7 million and 0.41, respectively.

477

478 With N = 10, the search concluded after 33 minutes at K =2232. The optimal decision
479 vector was found and correctly identified at k = 1. The search time was reduced
480 considerably because many fewer different decision vectors had to simulated; instead, a
481 few decision vectors were repeatedly found. Estimates of the objective function value
482 and p were the same as reported above.

483

With $N = 100$, the search concluded after 2 minutes at K=2232 with other results much the same as before. The optimal decision vector was found repeatedly and with few exceptions.

Experiment 2

For Experiment 2, varying M at a budget level of b = \$13 million, the heuristic parameter settings are relaxed somewhat: min~k = 100, max~k =5000, $\alpha = 0.05$, $ß = 0.01$, $\tau_0 = 0.001$, $\varepsilon = 0.0001$, $\delta = (0.005)(\mu_k^*)$, and $l = 0.975$. With N set to 100, M was tested across the values (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 40). All searches were completed either at $K = 362$ or 363 because one of the optimal solutions was found either at $k = 1$ or 2. The correct alternate optima (nine for $M = 1$, four for $M > 1$) were reported in all cases. The estimated objective function values, p values, and search times are reported in Table 1. For these runs, search times were lengthened during optimizations at large M by problem size. A few additional tests indicated that successful searches could be completed for $M = 40$ in a little over 7 minutes, a five-fold reduction in search time, with N set to 25 instead of 100.

Conclusion

While test results reported here are very limited, they indicate that the use of endogenous estimation formulations can potentially help solve difficult stochastic integer and chance-constrained integer programming problems. Whether exact or heuristic solution

algorithms are the best approach to solving these estimation problems likely depends on the particular problem at hand. Many further investigations might be required to more fully characterize the usefulness of this approach and to refine preferred solution techniques. The formulation and algorithm tested here offer a start that at least appears promising for solving seasonal fire planning problems.

Acknowledgements

The author is grateful to Laurie Travis and to Eldon Gunn for many helpful discussions of related problems. Any mistakes, however, are my own.

References

Bevers, M., and Kent, B. In review. Managing forest fire risk with chance-constrained programming. *In* Wildfire and Fuels Management: Risk and Human Reaction. *Edited by* Wade E. Martin, Carol B. Raish, and Brian Kent. Resources for the Future, Washington, DC.

Birge, J.R., and Louveaux, F. 1997. Introduction to Stochastic Programming. Springer-Verlag, New York.

Boesel, J., Nelson, B.L., and Kim, S. 2003. Using ranking and simulation to “clean up” after simulation optimization. Operations Research 51:814-825.

Boychuck, D., and Martell, D.L. 1988. A Markov chain model for evaluating seasonal forest fire fighter requirements. Forest Science 34:647-661.

Donovan, G.H. and Rideout, D.B. 2003. An integer programming model to optimize resource allocation for wildfire containment. Forest Science 49:331-335.

Fuessle, R.W., Brill, E.D., Jr., and Liebman, J.C. 1987. Air quality planning: A general chance-constraint model. Journal of Environmental Engineering 113:106-123.

Hardaker, J.B., Huirne, R.B.M., and Anderson, J.R. 1997. Coping with Risk in Agriculture. CAB International, New York.

Johnson, N.L., and Kotz, S. 1970. Distributions in Statistics: Continuous Univariate Distributions – 1. John Wiley & Sons, New York.

Kirsch, A.G., and Rideout, D.B. 2005. Optimizing initial attack effectiveness by using performance measures. *In* Systems Analysis in Forest Resources: Proceedings of the 2003 Symposium, 7-9 October, Stevenson, WA. *Edited by* Michael Bevers and Tara M. Barrett. USDA Forest Service General Technical Report GTR-PNW-656, Portland, OR, pp.183-187.

Mees, R., and Strauss, D. 1992. Allocating resources to wildland fires: A model with stochastic production rates. Forest Science 38:842-853.

Parks, G.M. 1964. Development and application of a model for suppression of forest fires. Management Science 10:760-766.

Reeves, C.R. (Ed.). 1993. Modern Heuristic Techniques for Combinatorial Problems. Halstead Press, New York.

Strauss, D., Bednar, L., and Mees, R. 1989. Do one percent of forest fires cause ninety-nine percent of the damage? Forest Science 35:319-328.

Table 1. Estimated objective function values, probability of exceedance values, and search times for various magnitudes of *M* in Experiment 2 on the *M*th order statistic estimation Problem SCCP-*k*.

M	Objective Function Value (million $)	Exceedance Probability	Search Time (minutes)
1	4.8	0.39	1.0
2	6.0	0.22	1.5
3	6.8	0.16	2.0
4	7.4	0.12	2.5
5	7.9	0.10	3.0
6	8.3	0.08	3.6
7	8.6	0.07	4.1
8	8.9	0.06	4.7
9	9.1	0.055	5.3
10	9.4	0.049	6.2
15	10.3	0.028	8.9
20	11.1	0.024	13.7
40	12.9	0.012	38.5