Annals of Operations Research 95 (2000) 205–215

Generating beta random rate variables from probabilistic estimates of fireline production times

J. Keith Gilless [a] and Jeremy S. Fried [b]

[a] *Department of Environmental Science, Policy & Management, University of California, 207 Giannini Hall MC #3310, Berkeley, CA 94720-3310, USA*
E-mail: gilless@nature.berkeley.edu

[b] *Department of Forestry, Michigan State University, 110 Natural Resources Bldg., East Lansing, MI 48824-1222, USA*
E-mail: jeremy@pilot.msu.edu

An extension of probabilistic PERT/CPM is proposed as a framework for soliciting expert opinion to characterize random variables for stochastic treatment in simulation models. By eliciting minimum, modal, ninetieth percentile, and maximum estimates, the distribution of variables with probability density functions of beta form can be explicitly characterized without relying on the traditional, but empirically unverified, assumption of a standard deviation equal to one-sixth of the range. This practical and inexpensive technique is illustrated by application to a wildfire protection planning problem – estimating the time required to produce a given length of fireline by different firefighting resources under diverse conditions. The estimated production times are an essential input to a planning model of initial attack on wildland fires used by the California Department of Forestry and Fire Protection, and provide that agency with useful "rules-of-thumb" for use in firefighter training.

Keywords: expert opinion, stochastic simulation, fire control

1. Introduction

Simulation models can often be enhanced by stochastic treatment of the time required to perform some activity. However, practical and budgetary constraints on direct measurement of production processes often preclude such enhancements. Expert knowledge is an attractive alternative source of quantitative information on processes for which stochastic treatment is essential, but on which direct measurement is infeasible or impractical. Fireline production, a critical element of the simulation of initial attack on wildland fires, is an example of such a phenomenon.

There have been numerous attempts to improve planning for wildland fire control through the use of simulation modeling [6]. The California Department of Forestry and Fire Protection (CDF) initiated such an effort in response to a request from the State Board of Forestry to conduct an efficiency analysis of the agency's wildland fire control activities to justify its expenditures on such activities. At the outset of this effort, the CDF decided to focus attention on initial attack. This decision reflected the large share of the agency's budget going to initial attack, a policy of aggressive action to prevent the loss of property and resources (i.e., to keep fires small), and the availability of a deterministic initial attack simulator [11]. This model simulates the production of fireline around wildfires spreading at different rates at each of a limited number of representative fire locations over the course of an average year. Simulated annual fire losses resulting from alternative deployments and positioning of firefighting resources provide an objective basis for making changes in actual practice. Experimentation with this model has shown that fireline productivity critically affects simulation results, and by extension, all decisions based on such information. Experience with the simulator as a decision-making tool brought to light a number of serious shortcomings, particularly those stemming from ignoring the pervasive “real world” variability in the effectiveness of initial attack efforts. For a system in which preventing extreme events is acknowledged to be of paramount importance, this variability must be captured if simulation results are to have either credibility or reliability [1]. This situation motivated the development of a new stochastic initial attack simulation model that explicitly accounts for variability in, among other things, fireline production rates.

The fire literature contains dozens of fireline production rate studies for different firefighting resources [4,5,9,10]. The rates reported in these studies depend upon crew size, fuel type, percent slope, and whether line construction moves up or downhill. The rates vary widely among studies: differences of 500 percent are not uncommon for what appear to be the same type of firefighting resources operating under the same conditions. Even allowing for inconsistencies in experimental design, this degree of variability clearly indicates that fireline production is most appropriately simulated as a stochastic variable. Unfortunately, these studies usually report only tables of rates by fuel and resource type, typically with many cells empty. The rates in these tables are usually interpreted as averages, despite the fact that they are often based upon a single observation per cell. Not surprisingly, basic distributional parameters (e.g., variance) are seldom reported, and can rarely be inferred from comparisons across studies.

In addition to their shortcomings as a basis for stochastic simulation, most prior production rate studies were conducted using personnel from other fire control agencies, and in vegetation not characteristic of CDF protected areas. Several costly attempts to directly measure production rates using CDF produced only a limited number of observations for conditions that represented a small fraction of the spectrum of firefighting situations encountered by the CDF.

To address these deficiencies in the available data, we conducted a statewide expert opinion survey, asking firefighters to estimate local production rates for the types of firefighting resources that they ordinarily command (fire engines, bulldozers, or handcrews). The survey methodology was inspired by probabilistic PERT/CPM methods, and produced time estimates suitable for characterizing fireline production rates as beta random variables in a stochastic simulator. (More accurate “point” estimates for use in the CDF’s existing deterministic initial attack simulator were also obtained.) Beta distributions are usually unimodal, have a finite range, and can be symmetric or highly skewed. This flexibility makes beta distributions well suited to describing nondeterministic production processes for which the "true" distribution is unknown [8].

The mechanics of the survey method and the analysis of the responses are outlined below in the context of describing the rates at which firefighting resources can produce fireline around a wildfire's perimeter.

2. Probabilistic PERT/CPM

Probabilistic PERT/CPM is a project management technique for describing the various activities that comprise a project, their precedence relationships, costs, and time requirements. The technique is "probabilistic" in the sense that it regards the time required to complete any activity, and by extension the time to complete the project, as stochastic variables [8]. This probabilistic treatment is based on expert estimates of the minimum, modal and maximum time required to complete each activity in a project – T_{min}, T_{mode}, T_{max}. When scaled over the interval $(0,1)$:

$$\tau_{\text{min}} = 0, \qquad \tau_{\text{max}} = 1,$$

and

$$\tau_{\text{mode}} = (T_{\text{mode}} - T_{\text{min}})/(T_{\text{max}} - T_{\text{min}}). \tag{1}$$

These transformed estimates are assumed to be the endpoints and modal value of a beta distribution with probability density function

$$f(\tau) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right]\left[\tau^{(\alpha-1)}(1-\tau)^{(\beta-1)}\right]. \tag{2}$$

Implicitly, this says that the time required to perform some task has a unimodal distribution with finite minimum and maximum bounds. This distribution may be symmetric or asymmetric. Integrating $\tau f(\tau)$ with respect to τ, the expected value and variance of τ are

$$E(\tau) = \frac{\alpha}{\alpha+\beta} \tag{3}$$

and

$$V(\tau) = \frac{\alpha\beta}{[(\alpha+\beta)^2(\alpha+\beta+1)]}. \tag{4}$$

Setting $\mathrm{d}[f(\tau)]/\mathrm{d}\tau = 0$ and solving for τ, the modal value of τ is

$$\tau_{\text{mode}} = \frac{\alpha-1}{\alpha+\beta+2}. \tag{5}$$

In probabilistic PERT/CPM practice, the objective is to put confidence limits on the time needed to complete an entire project. As a result, it is not necessary to explicitly estimate the α and β parameters of the beta distributions describing the time needed to complete any particular activity (e.g., building fireline along a flank of a fire). The three time estimates are used only to calculate an *approximate* mean:

$$\widehat{T} = \frac{[T_{\min} + [4T_{\mathrm{mode}}] + T_{\max}]}{6}. \tag{6}$$

3. Extension

An approximate mean, however, is an insufficient basis for stochastic simulation of fireline production. Even with an approximate variance, it would be difficult to devise a sampling process to compensate for skewness of the true distribution. However, by eliciting estimates of a fourth point on the distribution of T, the parameters of the underlying beta distribution can be calculated directly, allowing for explicit consideration of skewness, without relying on equation (6) or its implicit assumption of a standard deviation of production times equal to one-sixth of the range:

$$V(T) = \left[\frac{(T_{\max} - T_{\min})}{\rho}\right]^2, \tag{7}$$

where the parametric constant ρ is set to 6, or in terms of the beta distribution:

$$V(\tau) = \left(\frac{1}{\rho}\right)^2 \tag{8}$$

again with $\rho = 6$. A "fourth point" amenable to estimation by experts is the ninetieth percentile value (T_{90}), i.e., the one that would be exceeded ten percent of the time [7].

Specifically, beta distribution parameters α and β can be derived from equations (4), (5), (8) and estimates of $T_{\min}$, T_{mode}, T_{90}, and $T_{\max}$. Setting the right hand side of equation (8) equal to the right hand side of equation (4) and substituting in a value for ρ (e.g., 6), τ_{mode} can be substituted directly into equation (5) to obtain a system of two equations in two unknowns:

$$\alpha = \frac{[\tau_{\mathrm{mode}}(\beta + 2) + 1]}{1 - \tau_{\mathrm{mode}}}, \tag{9}$$

$$\begin{aligned} 0 = {} & \beta^3 + \beta^2 \big[\tau_{\mathrm{mode}}\big(-\rho^2\tau_{\mathrm{mode}}^2 + 2\rho^2\tau_{\mathrm{mode}} - \rho^2 - 7\big) + 4\big] \\ & + \beta\big[16\tau_{\mathrm{mode}}^2 - 18\tau_{\mathrm{mode}} + 5 - \rho^2\big(-2\tau_{\mathrm{mode}}^3 + 5\tau_{\mathrm{mode}}^2 - 4\tau_{\mathrm{mode}} + 1\big)\big] \\ & + \big[-12\tau_{\mathrm{mode}}^3 + 20\tau_{\mathrm{mode}}^2 - 11\tau_{\mathrm{mode}} + 2\big]. \end{aligned} \tag{10}$$

Solving the cubic equation for β yields either 1 or 3 real roots. In cases where 3 roots are found, β can be assigned the value of the largest root.

The parameter ρ reflects the degree of dispersion of a beta distribution. As ρ increases, $V(\tau)$ decreases. In the absence of more specific information, a value of 6 is a reasonable choice for a unimodal distribution [8]. The value assigned ρ can be critical in simulation applications, however, since it determines the relative frequency of "extreme" values in the tails of a beta distribution.

Estimates of T_{90} allow the value of ρ to be assessed in a straightforward manner. By scaling T_{90} over the interval $(0,1)$:

$$\tau_{90} = \frac{(T_{90} - T_{\min})}{(T_{\max} - T_{\min})}. \tag{11}$$

An estimate of ρ can be obtained using iterative numerical integration of the probability density function in equation (2). In the first iteration, equations (9) and (10) are solved α and β with ρ set to 6. The resulting beta distribution is then numerically integrated between 0 and τ_{90}. If this definite integral is not equal to 0.9, ρ must be incremented or decremented as appropriate, and the integration/area check process repeated. Eventually, either the area test will pass or it will be discovered that there is no value of ρ consistent with τ_{90} being the ninetieth percentile value of the distribution. A sample of estimates for "problem" observations (usually because τ_{90} is "too close" to τ_{mode}) can be analyzed with successively larger τ_{90} values until the probability density function can be integrated. The average of the ρ values obtained via this modified approach can be regarded as a lower bound estimate of ρ for "problem" observations.

4. Application: estimating fireline production rates

We surveyed more than two hundred California Department of Forestry and Fire Protection (CDF) firefighters to develop a stochastic representation of fireline production. These firefighters were asked to make best-case $T_{\min}$, most-likely case T_{mode}, ninetieth percentile T_{90}, and worst-case $T_{\max}$ estimates of the time required to build a length of fireline using the type of resource they ordinarily commanded (fire engines, bulldozers, or handcrews) for each of several different local control conditions in which differences in productivity were deemed significant. In each of the California Department of Forestry and Fire Protection Ranger Units in which the survey was conducted, three firefighters made estimates for each type of resource. Details of the survey administration are described in [3].

Values of ρ were calculated as described above for all sets of time estimates. An ANOVA ($P = 0.01$) of these ρ values revealed no significant effects from control condition or resource type, but some effect from administrative unit (CDF Ranger Unit or Contract County). The mean, mode, and median values of this distribution were 7.22, 6.99, and 6.69, respectively. The lower bound for ρ for "problem" observations was 12.8. The similarity of distributions of ρ across firefighting resources (figure 1) and administrative units led us to derive production time distributions throughout the state assuming $\rho = 7.22$, allowing estimation of T_{90} to be dropped from some later surveys. The "true" value of ρ, it can be assumed, lies somewhere between 7.22 and 12.8.

To obtain a single distribution representing production times for each firefighting resource/control condition combination, the three sets of estimates were aggregated using two methods: (1) α and β calculated from means of the three estimates of $T_{\min}$, T_{mode}, and $T_{\max}$, and (2) α and β calculated as means of the three α and three β

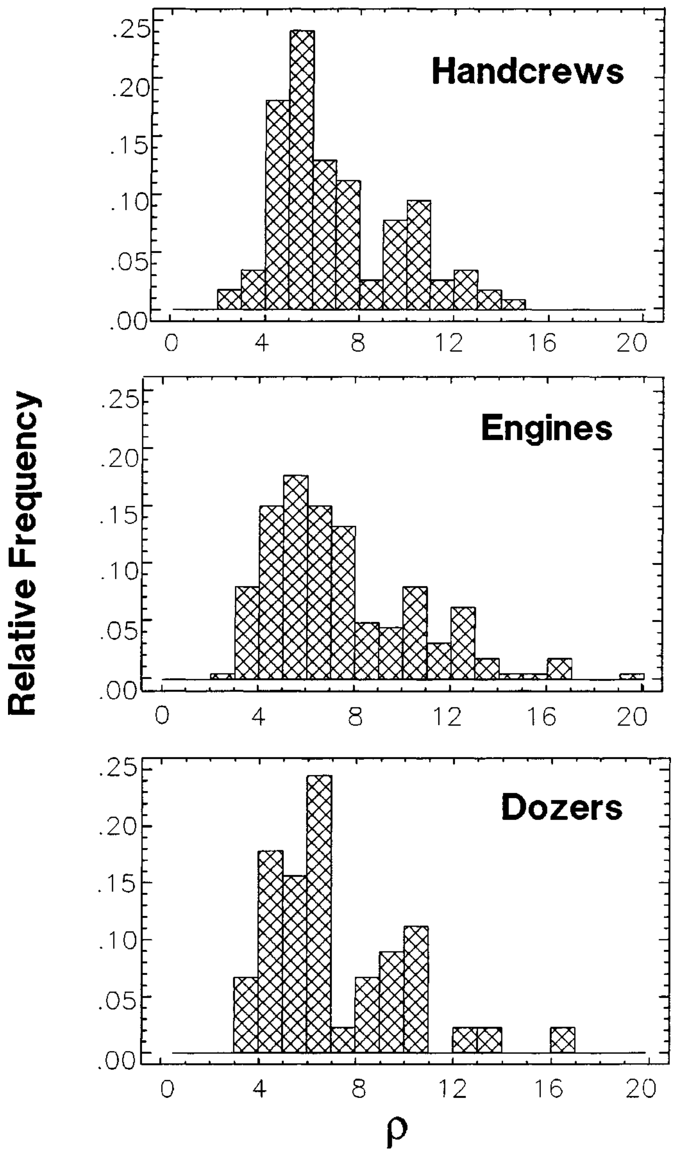

Figure 1. Relative frequency of ρ (a dispersion parameter) for beta distributions of the time required to build a length of fireline.

derived from the individual sets of estimates. Because the beta distribution parameters differed little between methods, and because of its stronger intuitive appeal, the first method appears preferable. Plots of aggregate beta distributions obtained via the preferred method well described the central tendency of the distributions estimated by individuals (figure 2).

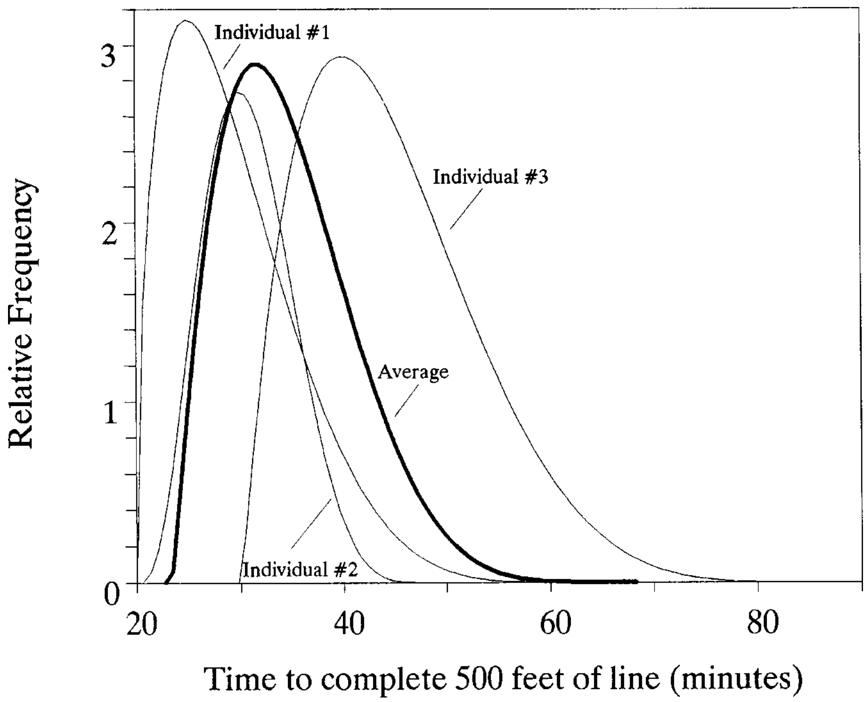

Figure 2. Beta distributions of the time required to build 500 feet of fireline by hoselay in brush fuels (Amador-El Dorado Ranger Unit).

Virtually all of the production time distributions were skewed left (towards shorter times), often significantly so (figure 3). For the sets of estimates obtained in this study (assuming $\rho = 7.22$), $\widehat{T}$ proved to be a robust estimator of $E(T)$. The mean absolute deviation of $\widehat{T}$ from $E(T)$, expressed as a percentage of $E(T)$, was only 2.1 percent. For a variety of firefighting resources and control conditions, the percentage difference between $\widehat{T}$ and $E(T)$ was greatest when the ratio α/β was small, i.e., when distributions were most skewed (figure 4). (Since β depends on the value ρ^2, the difference between 7.22 and 6 has a significant impact on a distribution's shape – see figure 5.) When a distribution function of this type is embedded in a complex stochastic simulator, the effects of an erroneous assumption about the degree of dispersion become difficult to predict. At best, it may lengthen the time required for the distributions of simulation outputs to stabilize within a specified tolerance. At worst, it may lead to spurious predictions of expected variability in the system. In any study of this kind, the most knowledgeable individuals should be polled for T_{90} estimates to calculate a value of ρ specific to the process being modelled. Another approach to ascertaining ρ was explored early in the study, but proved impractical to implement. Firefighters were asked for an increment of time such that when added to and subtracted from their most likely (modal) estimate it would define a range of values that contained the actual value 50% of the time. Skewness of the distributions in question made it difficult to

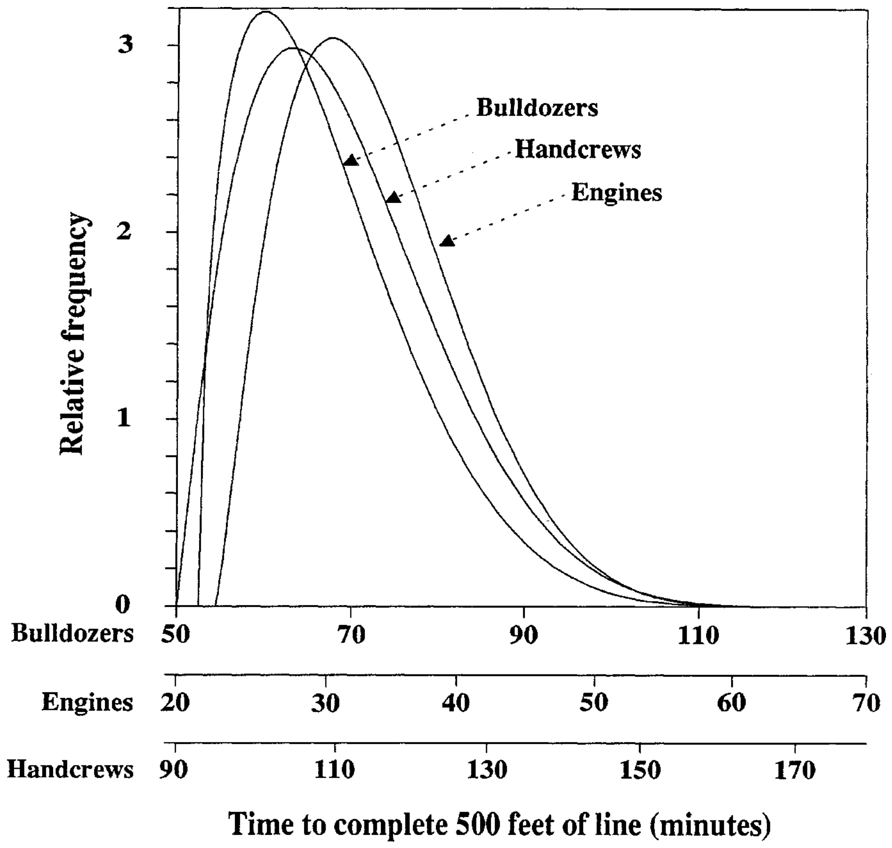

Figure 3. Beta distributions of the time required to build 500 feet of fireline in brush fuels (Amador-El Dorado Ranger Unit).

define such a symmetric range about the mode, and firefighters found this increment more difficult to estimate than T_{90}.

5. Conclusions

The four-estimate, expert-opinion approach described in this paper proved to be a practical and cost-effective alternative to undertaking a “direct measurement” study of fireline production rates. The survey met both the CDF’s immediate requirement for valid “point” estimates of production rates (for a deterministic initial attack simulation model) and their longer-term need for a stochastic characterization of production rates. Aside from their value in simulation modeling of initial attack on wildland fires, the estimated production times provide the CDF with useful “rules-of-thumb” for use in firefighter training. Firefighters were not intimidated by the survey format, and expressed confidence in both the process and their estimates. The results of the survey have credibility with the CDF, and have been incorporated into ongoing decision-

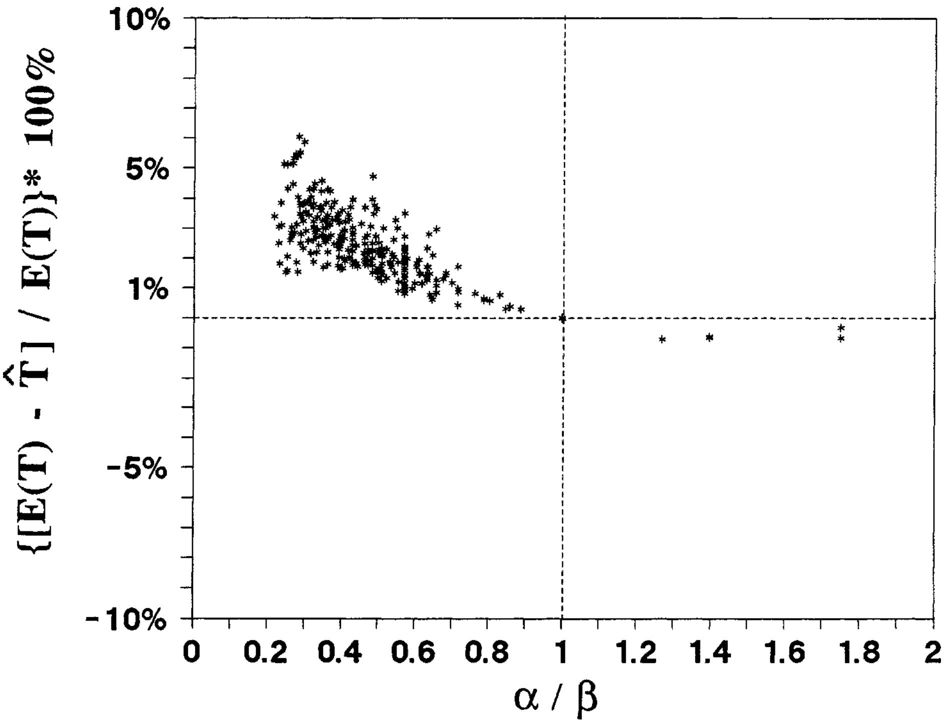

Figure 4. Percent error of $\widehat{T}$ from $E(T)$ calculated using α and β plotted against the ratio (α/β) $(n = 327)$.

making and planning efforts. Other researchers considering an expert-opinion approach to quantifying the stochastic properties of a production process may find this survey format to be of considerable value.

On a practical basis, the survey results confirmed a widely held suspicion that production rates reported in the literature are too high for CDF conditions. To revise dispatch strategies or make decisions on the basis of such overly optimistic rates could easily result in significant loss of property and resources. Simulations of initial attack on wildland fires with the more conservative (and presumably more realistic) production rates generated by the expert opinion survey is providing the CDF with a rigorous justification for its current policy of a quick and aggressive initial response, as well as helping identify incremental improvements to a system that already works well.

When survey results were incorporated into the stochastic simulator CFES2 [2], they were not based on the traditional PERT/CPM assumption that production rates can be assumed to follow a beta distribution $\rho = 6$. Samples drawn from beta distributions based on the empirically estimated value for ρ (7.22) incorporated far fewer extremely slow production rates, and more that are close to the appropriate modes. An accurate frequency of extreme values is critical for stochastic simulation models of initial attack, since a major objective of fire planning is to minimize the number of fires that "escape"

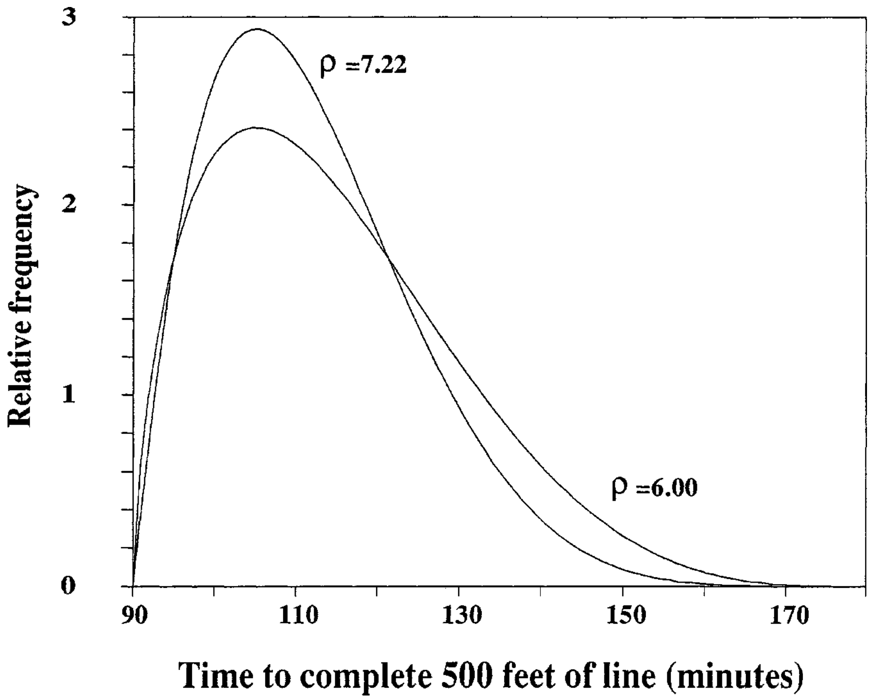

Figure 5. Beta distributions of fireline production times derived using alternative values for ρ.

initial attack (i.e., are not controlled within some size or time limit) due to perimeter growth exceeding aggregate fireline production.

References

[1] J.S. Fried and J.K. Gilless, Modification of an initial attack simulation model to include stochastic components, in: *Proceedings of the 1988 Symposium on Systems Analysis in Forest Resources*, Asilomar Conference Center, Pacific Grove, CA (March 19–April 1, 1988) pp. 235–240; General Technical Report RM-161, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, p. 278.

[2] J.S. Fried and J.K. Gilless, CFES2: The California Fire Economics Simulator Version 2, User's Guide, University of California Agricultural Experiment Station Publication 21580 (1999).

[3] J.S. Fried and J.K. Gilless, Expert opinion estimation of fireline production rates, Forest Science 35(3) (1989) 870–877.

[4] L. Haven, T.P. Hunter and T.G. Storey, Production rates for crews using hand tools on firelines, General Technical Report PSW-62, Pacific Southwest Forest and Range Experiment Station, Berkeley, CA (1982) p. 8.

[5] K.G. Hirsch and D.L. Martell, A review of initial attack fire crew productivity and effectiveness, International Journal of Wildland Fire 6(4) (1996) 199–215.

[6] D.L. Martell, A review of operational research studies in forest fire management, Canadian Journal of Forest Research 12 (1982) 119–140.

[7] J.J. Moder, C.R. Phillips and E.W. Davis, *Project Management with CPM, PERT and Precedence Diagramming* (Van Nostrand Reinhold, New York, 1983) p. 389.

[8] J.J. Moder and E.G. Rodgers, Judgement estimates of the moments of PERT type distributions, Management Science 15(2) (1968) B76–B83.

[9] C.B. Phillips, C.W. George and D.K. Nelson, Bulldozer fireline production rates – 1988 Update, Research Paper INT-392, U.S.D.A. Intermountain Experiment Station, Ogden, UT (1988) p. 13.

[10] C. Wilson, Determining production rates for fire engines and their crews, 1978-1979, Unpublished manuscript, California Department of Forestry, Sacramento, CA (1980) p. 16.

[11] U.S. Forest Service, National fire management analysis system users' guide of the initial action assessment model (FPL-IAA2.2), U.S.D.A. Forest Service, Washington, DC (1985).