Forest Science, Vol. 34, No. 4, pp. 948–959.

Stochastic Representation of Fire Occurrence in a Wildland Fire Protection Planning Model for California

J. S. FRIED
J. K. GILLESS

ABSTRACT. A fire occurrence module was developed for CFES-IAM Version 2, a stochastic simulation model of initial attack on wildfires in California. The module is designed to generate annual sequences of fire start times that are consistent with local fire history. A three-stage approach was employed: (1) For each day of a simulated fire season, a random draw from a Bernoulli distribution is used to determine if any fires occur; (2) if any do occur, a random draw from a geometric multiplicity distribution determines their number; (3) ignition times for each fire are then randomly drawn from a time of day (beta or Poisson) distribution. This approach and specific distributional forms were selected after analysis of historical fire records from California's Sierra foothills and Central Valley. Fire sequences generated with the module appear to capture historical patterns with respect to diurnal distribution, interfire times, and total number of fires per year. FOR SCI. 34(4):948–959.

ADDITIONAL KEY WORDS. Simulation, initial attack, CFES-IAM, fire load, fire ignitions.

DESPITE THE STOCHASTIC CHARACTER OF FIRE OCCURRENCE, most wildland fire protection planning models rely on a deterministic treatment of this phenomenon. Developers of FOCUS (Fire Operational Characteristics Using Simulation) envisioned a model that generated fires stochastically, but ultimately compromised on simulating a list of historical fires (Bratten et al. 1981). The developers of the IAA (Initial Action Assessment) model generalized this approach by simulating initial attack on several fires spreading at different rates at each of a limited number of representative fire locations. Simulation results for a management unit are extrapolated to an annual basis using the historical average number of fires per year (USDA Forest Service 1985). CFES-IAM Version 1 (California Fire Economics Simulator-Initial Attack Module) is based on the same underlying assumptions as IAA (Fried and Gilless 1988). FEES (Fire Economics Evaluation System) comes closer to a truly stochastic treatment of fire occurrence. It considers many different kinds of fires breaking out at purely hypothetical locations and uses historical fire records to determine relative probabilities for each of these fires. Probabilistic financial and physical impacts are then estimated at the management unit level on the basis of a probability distribution for the total number of fires per year (Mills and Bratten 1982).

All of these models made significant contributions to wildland fire protection planning, but each has serious limitations as a tool for evaluating the capability of a fire protection organization to address the problems of fighting several fires simultaneously, time-of-day constraints on the effectiveness of some fire-fighting resources (e.g., air tankers), and determining

---

The authors are Associate Specialist and Assistant Professor, Department of Forestry and Resource Management, University of California, Berkeley, respectively. Manuscript received March 7, 1988.

Reprinted from the *Forest Science*, Vol. 34, No. 4, December 1988

appropriate seasonal staffing. Further, IAA and CFES simulate an average fire year, ignoring the variability likely to be encountered between fire years. FEES recognizes this variability, but its informational requirements and operational characteristics make it difficult to use. These limitations in existing wildland fire planning models prompted the California Department of Forestry and Fire Protection (CDF) to commission the development of CFES-IAM Version 2. The research reported here is an outgrowth of the development of this model.

Multiple fire starts are a major problem for the CDF (and other wildland fire protection organizations). To permit realistic simulation of multiple fire starts, CFES-IAM Version 2 will necessarily be a next-event, clock-driven simulator, i.e., after processing an event (e.g., fire ignition, arrival of a fire-fighting resource, or fire containment), the simulator's "clock" will advance to the next event, explicitly recognizing the importance of chronological and spatial proximity of events. For example, the fact that fire engines committed to one fire at 12:05 P.M. are not available for dispatch to another fire at 12:10 P.M. can only be represented using a clock-driven formulation.

The degree to which CFES-IAM Version 2 (or any clock-driven simulator) captures the dynamics of initial attack on forest fires will be strongly influenced by the validity of its occurrence module, i.e., upon its ability to generate fire starts in patterns that are consistent with an area's fire history. This paper describes the structure of the fire-occurrence module developed for CFES-IAM Version 2 and presents evidence as to its validity.

OBJECTIVES

Four evaluation criteria guided the development of the fire occurrence module for CFES-IAM Version 2. Specifically, close correspondence was sought between generated sequences of fires and historical fire records with respect to

1. Distribution of number of fires per year.
2. Distribution of fires by time of day.
3. Frequency and severity of multiple fire days (days with more than one fire).
4. Distribution of fires by season (or time of year).

Correspondence between the historical and generated distributions of number of fires per year is important for capturing the scale of the fire management problem. Agreement between historical and generated distributions for time of day is critical for simulating initial attack on multiple fire days, and to reflect the fact that some firefighting resources can only be used during daylight hours. Correspondence between historical and generated sequences of fires with respect to the frequency and severity of multiple fire days is necessary to evaluate the ability of the CDF to deal with severe fire seasons (i.e., those that can be characterized by longer response times and decreased response capability). Finally, a match between historical and generated fire sequences with respect to the distribution of fires by season (or time of year) is important given seasonal differences in fire organizations' staffing and response capabilities.

DATA

The CDF is building microcomputer databases for each of its 29 ranger units and contract counties. These databases will contain the date and time of occurrence, location, size at arrival and upon control, and estimated fire behavior parameters for each wildland fire since 1980 or 1981. The data used for the research reported here comes from recently completed databases for the Almador-El Dorado (AEU), Fresno-Kings (FKU), and Nevada-Yuba-Placer (NEU) ranger units, all located in the western foothills of the Sierra Nevada and on the eastern margin of the San Joaquin Valley, in central and northern California. These databases contain data describing 1906, 538, and 1680 wildland fires, respectively.

METHODS

Most next-event, clock-driven simulators are based on a single distribution describing the time between events. The initial structure contemplated for the fire occurrence module assumed that sequential fire ignitions over the course of a year or season could be generated from an estimated distribution for the time between fire events (TNext), as proposed by Johnson and Van Wagner (1985). From some first ignition, the time of each subsequent ignition would be determined by incrementing the simulation clock by a randomly drawn value from this distribution. Sequences of fires thus generated would include both periods of intense, possibly overlapping, fire activity as well as periods with relatively few fires.

The TNext distribution needed for a fire occurrence module based directly on fire frequency were best described for the AEU, FKU, NEU ranger units by an exponential form. This result is consistent with interarrival time distributions used in most simulation models (Law and Kelton 1982). For all three ranger units, the estimated distributions could be used to generate a sequence of TNexts corresponding to historical patterns. Unfortunately, the resulting distribution of fire occurrence by time of day did not demonstrate the diurnal pattern characteristic of real fires. In essence, fires had an equal probability of occurring at night or during the day. If the simulation of initial attack is sophisticated enough to reflect the influence of time of day on dispatch policies, firefighting tactics, and effectiveness, simulation results based on a ''TNext'' occurrence module might contain serious bias. Therefore, despite the appeal of its simplicity and the accuracy with which TNext distributions could be estimated, this approach was judged to be inappropriate given the stated objectives of this research. However, the concept of a TNext distribution proved useful in validating the structure ultimately chosen for the fire-occurrence module.

At this point, the short duration of most fires fought by the CDF and the wildfire occurrence prediction work of Haines et al. (1983) and Cunningham and Martell (1976) suggested an alternative structure, in which fire ignitions for any day are generated independently of those for preceding or subsequent days. This structure requires the estimation of not one, but several distributions, which together could be used to generate a sequence of fire ignitions over the course of a day. Although more complex, this structure seemed capable of producing a pattern of fires with a more acceptable distribution by time of day (ToD).

Specifically, this alternative structure uses three distributions to generate a sequence of fire ignitions. For each day in a year or season, Fireday, a randomly drawn value from a Bernoulli (0, 1) distribution would determine whether any fires occur on that day. Given that one or more fires occur, a randomly drawn Multiplicity value from a second, discrete distribution would determine their number. The ignition time for each of these fires would then be determined by randomly and independently drawn ToD values from a third distribution.

In recognition of California's annual pattern of fire occurrence, dates were identified for each ranger unit that divided the year into three seasonal classes of relatively homogeneous fire frequency. These classes are referred to below as the Low, Transition, and High fire seasons. The distributional forms that best described the probability of occurrence (Fireday), the number of fires per day (Multiplicity) and the time of day (ToD) of the fires for each ranger unit, by season, were then identified.

RESULTS

The left half of Figure 1 shows histograms of the number of fires per week for each ranger unit during the periods covered by the data. Inspection of these histograms and Tukey multiple range tests of TNext by week suggested the classification of weeks into fire seasons by ranger unit (Figure 2). The Transition season typically included 3 to 5 weeks in the spring and another few weeks in the fall when fire incidence was higher than in the Low season but less than in the High season. For all three ranger units, mean

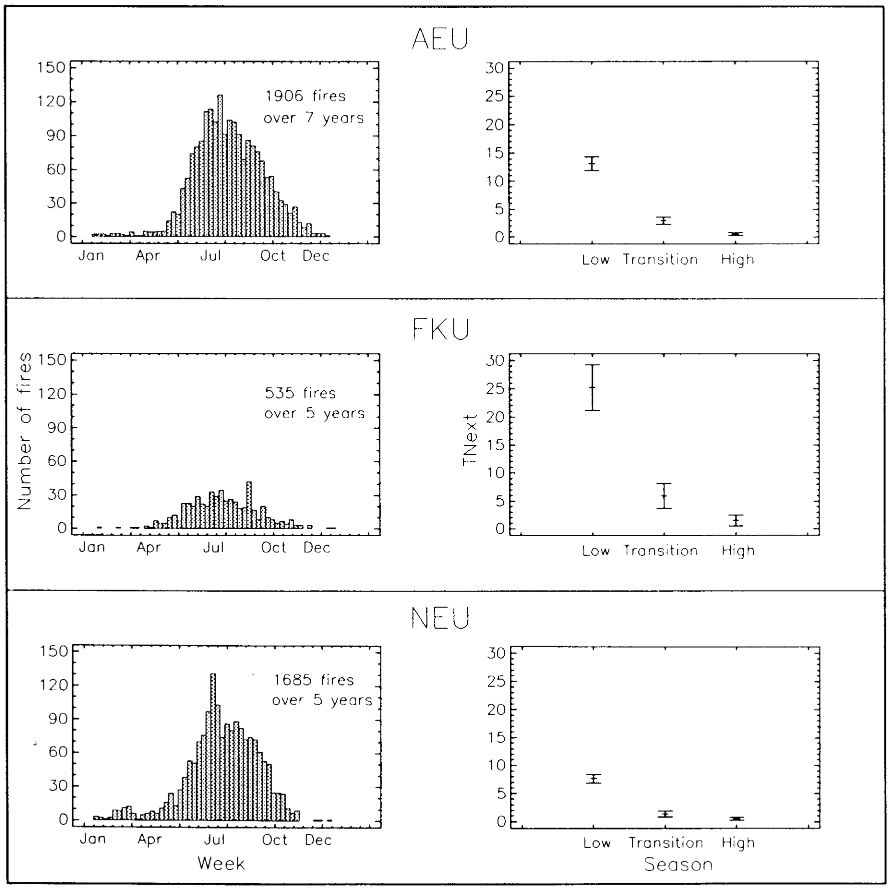

FIGURE 1. Left half: Number of fires per week over the periods represented in the AEU, FKU, and NEU databases. Right half: Mean and 95% confidence interval plots of TNext for the Low, Transition, and High fire seasons as defined in Figure 2 for the AEU, FKU, and NEU ranger units.

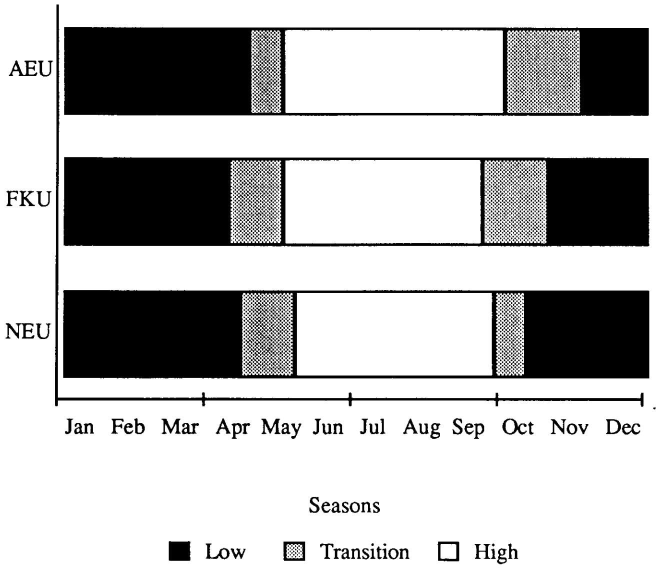

FIGURE 2. Low, transition, and high fire seasons for the AEU, FKU, and NEU ranger units.

TNext was significantly different for each season as shown by the Means/ 95% confidence interval plots of TNext in the right half of Figure 1.

For each day during the period 1981–1985 for the NEU and FKU ranger units, and 1980–1986 for the AEU ranger unit, Fireday was defined as a Bernoulli variable equal to 1 if any fires occurred on that day and 0 otherwise. A Bernoulli distribution of the form

$$p(x) = \begin{array}{ll} 1 - \pi & \text{if } x = 0 \\ \pi & \text{if } x = 1 \\ 0 & \text{otherwise} \end{array}$$

where x = Fireday, was fit for each fire season and ranger unit. The parameter π can be interpreted as the probability of one or more fires occurring on any one day. Estimated values for π are shown in Table 1.

For each ranger unit, histograms showing the relative frequency of Multiplicity (number of fires per day) for days in the High season on which fire(s) occurred are shown in the left half of Figure 3. For each season, the transform (Multiplicity-1) was best described by a geometric distribution with probability mass function

$$p(x) = \phi(1-\phi)x \text{ if } x \in \{0,1,\ldots\}$$

$$0 \text{ otherwise,}$$

TABLE 1. Probabilities of one or more fires occurring on one day by season and ranger unit.

Season	AEU	FKU	NEU
Low	0.043	0.027	0.102
Transition	0.261	0.190	0.411
High	0.660	0.444	0.822

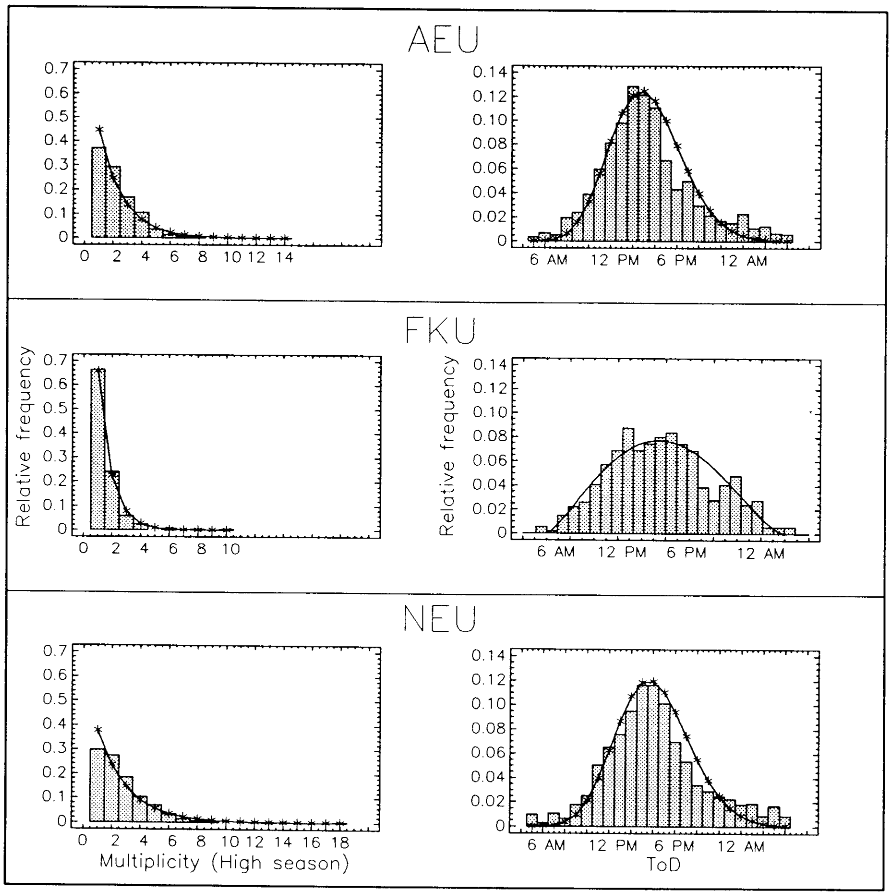

FIGURE 3. Left half: Relative histograms of historical high fire season days having one or more fires and the corresponding fitted Multiplicity distributions for the AEU, FKU, and NEU ranger units. Right half: Histograms of the relative frequency of historical fires by time of day and the corresponding fitted ToD distributions for the AEU, FKU, and NEU ranger units.

where x = the number of fires on one day -1. Estimated geometric distributions are shown superimposed on the Multiplicity histograms. Estimates of ϕ are given in Table 2 along with χ^2 goodness-of-fit statistics for each fire season. Although the degree of Multiplicity represented in the Low and

TABLE 2. Estimated ϕ parameters and χ^2 goodness-of-fit significance levels for geometric distributions fitted to (Multiplicity-1) by season and ranger unit.

	AEU		FKU		NEU	
Season	ϕ	χ^2 sig. level	ϕ	χ^2 sig. level	ϕ	χ^2 sig. level
Low	0.818	*	0.917	*	0.769	*
Transition	0.640	0.66	0.839	*	0.599	0.58
High	0.445	0	0.657	0.42	0.377	0

Note: * denotes too few classes to calculate χ^2 goodness-of-fit statistic.

Transition seasons was insufficient to calculate χ^2 statistics, the fit of these geometric distributions over all seasons combined was reasonably good, far better than any logical alternatives such as the exponential distribution.

Unlike Fireday and Multiplicity, ToD (Time of day) exhibited no seasonal differences. Thus, a single ToD distribution was estimated for each ranger unit. Frequency distributions of ToD varied in appearance, but all had central tendencies when left-shifted 5 hours (so that 0 corresponded to 5 A.M. and 23 to 4 A.M. the next day) (Figure 3, right half). The FKU ranger unit's ToD distribution exhibited a broad peak between 11 A.M. and 7 P.M. and was best fit by a beta distribution normalized to a 0-1 scale

$$f(x) = \frac{x^{\alpha_1 - 1}(1 - x)^{\alpha_2 - 1}}{\dfrac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}} \text{ if } 0 \leq x \leq 1$$

$$0 \text{ otherwise.}$$

By contrast, the AEU and NEU ranger units' ToD distributions had high, narrow frequency peaks from 12 P.M. to 4 P.M., and were best fit by Poisson distributions

$$p(x) = \frac{(e^{-\lambda}\lambda^x)}{x!} \text{ if } x \in \{0, 1, \ldots\}$$

$$0 \text{ otherwise.}$$

The fitted ToD distributions are shown superimposed on the corresponding ToD histograms. χ^2 goodness-of-fit significance levels and estimated parameter(s) for each ToD distribution are reported in Table 3, along with the transformations appropriate for those model forms.

VALIDATION

A primitive version of the fire-occurrence module based on the distributions described above generated 60 years of fires for each ranger unit. To test the validity of the overall model structure, we compared subsets of the generated fires with their historical counterparts using the time between fires variable, TNext. We also compared the generated and historical distributions for the number of fires per year.

For all three ranger units, there was satisfactory correspondence between

TABLE 3. Distribution parameters and χ^2 goodness-of-fit significance levels for beta and Poisson distributions fitted to Time of day (ToD) by ranger unit.

	AEU	FKU	NEU
Best fitting distribution	Poisson	Beta	Poisson
Transformation	1. subtract 5 2. if result is < 0 then add 24	1. subtract 5 2. if result is < 0 then add 24 3. divide by 23	1. subtract 5 2. if result is < 0 then add 24
Estimated parameter(s)	$\lambda = 10.279$	$\alpha_1 = 2.419$ $\alpha_1 = 2.598$	$\lambda = 11.088$
χ^2 significance level	0	0.09	0

historical and generated TNext distributions, as demonstrated by the paired histograms in Figure 4. Correspondence in diurnal fluctuations in relative frequency were apparent when these histograms were plotted using 3-hour intervals over a 2-day period. Over 10 days, the exponential decay of both historical and generated TNexts appeared extremely close. In no case were the historical and generated distributions wildly disparate. The graphical and tabular results of a more formal statistical comparison are summarized in Figure 5 and Table 4. The paired box and whisker plots for each season and ranger unit (Figure 5) clearly show similar central tendencies and degrees of variability for the historical and generated fires. Means, and to a lesser extent, medians, corresponded well. No consistent bias was observed for the differences in means, medians, or variances. Except for the High season on the AEU ranger unit, t-tests indicated that the hypothesis that the historical and generated TNext distributions are part of the same distribution could not be rejected at the 0.05 significance level (Table 4).

The box and whisker plots of historical and generated number of fires per year for each ranger unit demonstrate more than acceptable similarity

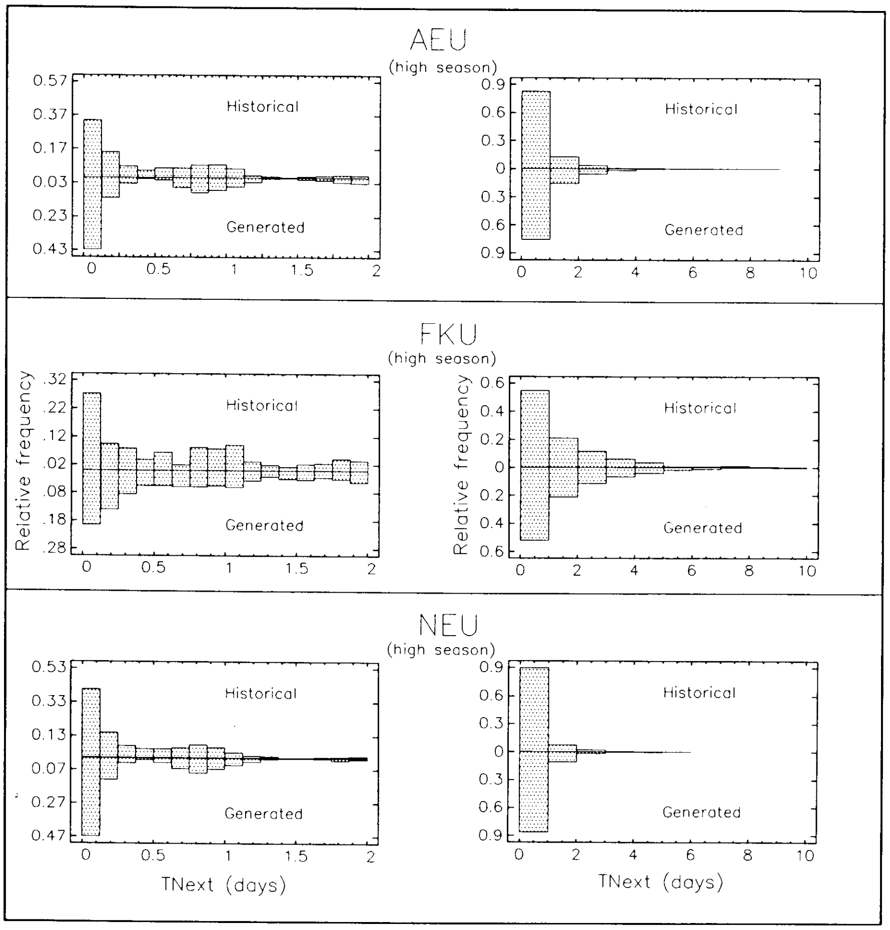

FIGURE 4. Left half: Paired, 3-hour interval, relative frequency histograms of historical and generated TNext distributions over a 2-day span. Right half: Paired, 1-day interval, relative frequency histograms of historical and generated TNext distributions over a 10-day span.

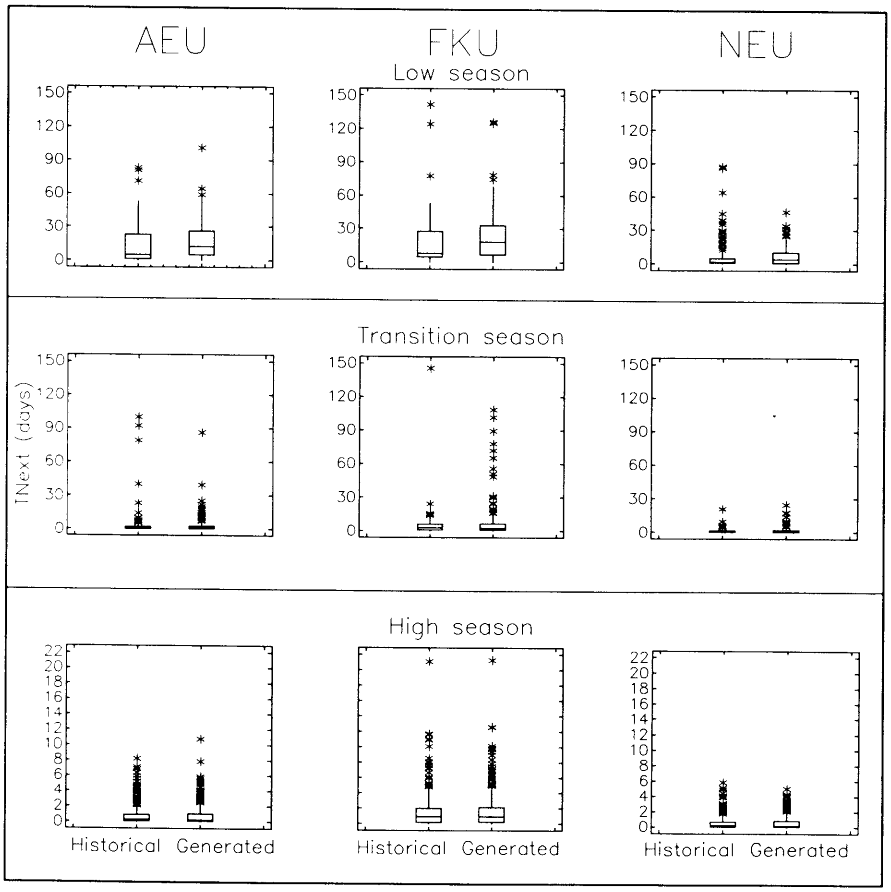

FIGURE 5. Box and whisker plots of historical and generated TNext distributions and corresponding vital statistics for the AEU, FKU, and NEU ranger units for the Low, Transition, and High fire seasons. The center horizontal line in each box represents the median of that distribution, while the upper and lower horizontal lines mark the upper and lower quartiles. The vertical lines (whiskers) extend out to the extremes (minimum and maximum) of the distribution. Unusual values that occur beyond 1.5 times the interquartile range (the length of the box) are plotted as separate points.

(Figure 6). For FKU and NEU, the historical and generated mean number of fires per year differed by less than 1% (Table 5).

DISCUSSION

Wildland fire protection agencies operate in an environment characterized by great variation in fire activity and its consequences. CFES-IAM Version 1 is a strictly deterministic simulator and was not designed to provide or utilize any information on such variation. While effective for evaluating fire protection alternatives for the average fire year, a deterministic model offers no insights for nonaverage years. This is a serious shortcoming in a state like California, where the consequences of being ill-prepared for an extreme fire year can be catastrophic. Concern over this issue has prompted the

TABLE 4. Descriptive statistics for historical and generated distributions of the time between fires (TNext) by season and ranger unit.

	AEU		FKU		NEU	
	Historical	Generated	Historical	Generated	Historical	Generated
Low						
Years	7	10	5	20	5	10
Fires	64	69	25	102	120	265
Median	4.97	11.90	8.06	18.17	1.80	4.09
Mean	13.12	19.24	25.21	24.06	7.64	6.62
Variance	355.9	403.5	1381.0	564.6	227.6	55.58
Significance level	0.07		0.85		0.37	
Transition						
Years	7	10	5	20	5	10
Fires	223	337	80	347	192	427
Median	0.913	1.002	3.009	2.613	0.736	0.760
Mean	2.980	2.602	5.935	6.274	1.342	1.460
Variance	117.9	36.50	269.1	162.6	4.362	6.145
Significance level	0.6		0.84		0.57	
High						
Years	7	10	5	20	5	10
Fires	1619	2006	430	1658	1373	2787
Median	0.303	0.249	0.919	0.940	0.203	0.153
Mean	0.593	0.707	1.491	1.553	0.461	0.457
Variance	0.605	0.845	4.246	3.553	0.383	0.339
Significance level	0		0.54		0.82	

Note: Significance levels are for the t test with H_O: Distribution means are identical, H_A: Distribution means are different.

development of a stochastic simulator, CFES-IAM Version 2.

A critical component of any stochastic simulation of initial attack on wildland fires is the treatment of fire occurrence. The fire occurrence module for Version 2 could have been based on stochasticly generated

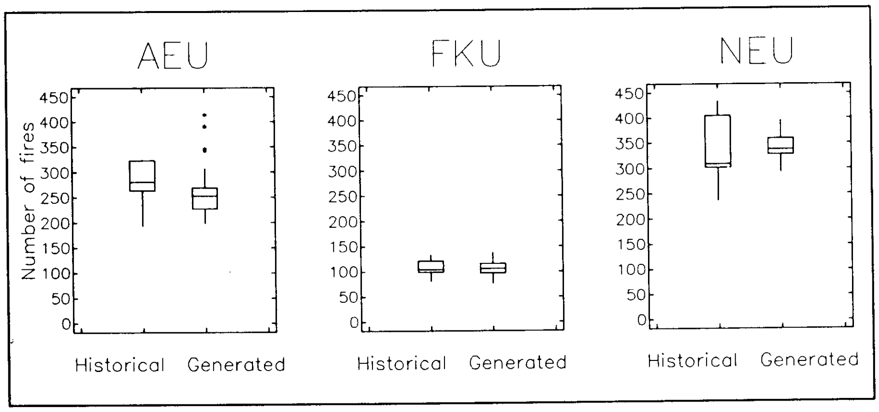

FIGURE 6. Box and whisker plots of historical and generated Number of Fires per Year distributions and corresponding vital statistics for the AEU, FKU, and NEU ranger units. The center horizontal line in each box represents the median of that distribution, while the upper and lower horizontal lines mark the upper and lower quartiles. The vertical lines (whiskers) extend out to the extremes (minimum and maximum) of the distribution. Unusual values that occur beyond 1.5 times the interquartile range (the length of the box) are plotted as separate points.

TNexts (times between fires), in accordance with standard practice in next-event based simulators. Historical distributions of TNext were well described by an exponential model; however, sequences of fire generated in this manner exhibited no diurnal variation. Given the clustering of real fires about midday, this potential source of bias was judged to be significant enough to warrant developing an alternative approach. Since fire events on wildlands protected by the CDF are relatively independent from one day to the next, an obvious alternative was to separate the problem into two components, determining: (1) how many fires occur on a given day, and (2) when they occur during that day.

No single distributional form proved capable of accurately representing the number of fires per day, largely due to the predominance of days with no fires. However, when Multiplicity was considered only for days on which at least one fire occurred, a geometric form provided a satisfactory representation of how many fires occurred on those days. This suggested a hybrid treatment of the first component—a simple Bernoulli "fire(s)/no fire(s)" distribution coupled with a geometric Multiplicity distribution. Treatment of the second component was complicated by the fact that distributions of the time of day at which fires occur differed significantly among the three rangers units. A Poisson distribution best described this phenomenon for some ranger units, while a beta distribution was superior for ranger units with a less pronounced clustering of fires about midday. Seasonal differences in the Multiplicity and fires/no fires component required that those distributions be estimated separately for Low, Transition, and High fire seasons.

Based on these results, a fire-occurrence module has been developed that operates as follows. For each day of the simulation, a Bernoulli distribution determines if any fires occur. If so, a geometric Multiplicity distribution determines their number. Ignition times for each fire are then independently selected from a Time of Day distribution (ToD). Experimentation with this module has shown that it generates sequences of fires with TNext patterns quite similar to those of historical fires. The central tendency and variability of the total number of fires in these generated sequences compare favorably to their historical counterparts.

The approach outlined in this paper for the stochastic representation of fire occurrence differs significantly from those previously employed in wildland fire planning models. It is, in some ways, more mechanical and less descriptive; e.g., no attempt is made to stratify fire occurrences by cause or weather conditions. No attempt was made to understand why certain distributional forms describe various components of fire occurrence better than

TABLE 5. Descriptive statistics for historical and generated distributions of the time between fires (TNext) by season and ranger unit.

	AEU		FKU		NEU	
	Historical	Generated	Historical	Generated	Historical	Generated
Years	7	60	5	60	5	60
Median	280.0	252.5	104.0	106	309.0	338.0
Mean	279.0	256.7	107.6	106.4	337.2	339.5
Variance	1987.0	1639.0	404.8	167.2	6462.0	551.7
Significance	0.17		0.84		0.86	

Note: Significance levels are for the t test with H_O: Distribution means are identical H_A: Distribution means are different.

others. However, it does satisfy the rigorous set of objectives guiding its development, it is simple enough to be incorporated into a microcomputer simulator, its data requirements are modest, and estimation of local distributional parameters does not require advanced statistical skills.

LITERATURE CITED

BRATTEN, F. W., et al. 1981. FOCUS: A fire management planning system—final report. USDA For. Serv. Gen. Tech. Rep. PSW-49. 34 p.

CUNNINGHAM, A. A., and D. L. MARTELL. 1976. The use of subjective probability assessments to predict forest fire occurrence. Can. J. For. Res. 6(3):348–356.

FRIED, J. S., and J. K. GILLESS. 1988. The California fire economics simulator—initial attack module (CFES-IAM): MS-DOS Version 1.11 User's Guide. Bull. 1925. Div. of Agric. and Nat. Resour., Univ. Calif., Berkeley. 84 p.

HAINES, D. A., et al. 1983. Fire-danger rating and wildfire occurrence in the Northeastern United States. For. Sci. 29(4):679–696.

JOHNSON, E. A., and C. E. VAN WAGNER. 1985. The theory and use of two fire history models. Can. J. For. Res. 15(1):214–220.

LAW, A. M., and W. D. KELTON. 1982. Simulation modeling and analysis. McGraw Hill, New York. 400 p.

MILLS, T. J., and F. W. BRATTEN. 1982. FEES: Design of a fire economics evaluation system. Gen. Tech. Rep. PSW-65. 26 p.

USDA For. Serv. Gen. Tech. Rep. PSW-65. 26 p.

USDA Forest Service. 1985. National fire management analysis system users' guide of the initial action assessment model (FPL-IAA2.2). USDA For. Serv., Wash., DC.