Statistical Model for Forecasting Monthly Large Wildfire Events in Western United States

HAIGANOUSH K. PREISLER
Pacific Southwest Research Station, U.S. Department of Agriculture Forest Service, Albany, California

ANTHONY L. WESTERLING
School of Engineering, University of California, Merced, Merced, California

(Manuscript received 30 March 2006, in final form 7 November 2006)

ABSTRACT

The ability to forecast the number and location of large wildfire events (with specified confidence bounds) is important to fire managers attempting to allocate and distribute suppression efforts during severe fire seasons. This paper describes the development of a statistical model for assessing the forecasting skills of fire-danger predictors and producing 1-month-ahead wildfire-danger probabilities in the western United States. The method is based on logistic regression techniques with spline functions to accommodate nonlinear relationships between fire-danger predictors and probability of large fire events. Estimates were based on 25 yr of historic fire occurrence data (1980–2004). The model using the predictors monthly average temperature, and lagged Palmer drought severity index demonstrated significant improvement in forecasting skill over historic frequencies (persistence forecasts) of large fire events. The statistical models were particularly amenable to model evaluation and production of probability-based fire-danger maps with prespecified precisions. For example, during the 25 yr of the study for the month of July, an area greater than 400 ha burned in 3% of locations where the model forecast was low; 11% of locations where the forecast was moderate; and 76% of locations where the forecast was extreme. The statistical techniques may be used to assess the skill of forecast fire-danger indices developed at other temporal or spatial scales.

1. Introduction

Wildland managers have long desired to know the risks of severe fire events in advance of their occurrence. A number of actions are available to improve the efficiency of wildfire suppression efforts during severe fire seasons, including reallocating resources from other activities to fire suppression, reallocating suppression resources from low- to high-risk regions, public education campaigns to reduce ignitions, curtailment of fire use for vegetation management, and changing the number and timing of temporary hires. Because the annual expenditure on wildfire suppression is so high—averaging more than $1.3 billion per year for federal land management agencies since 2000 after adjusting for inflation [Fire suppression costs were obtained from the National Interagency Fire Center (online at http://www.nifc.gov) and adjusted for inflation using the western urban consumer price index from the Bureau of Labor Statistics.]—that even marginal improvements in suppression efficiency have the potential to save significant sums of money through either reduced suppression costs or reduced losses resulting from wildfire. However, contingency actions have associated costs and timeliness issues. For example, temporary hires for fire suppression require training before they can be put to work suppressing wildland fires, but may be retained for only a limited time period (6 months). Hiring them too early or too late incurs the risk that they will not be available when they are most needed. Another need for fire-danger forecasting is for planning prescribed fires. Making effective decisions in light of these issues requires information about the likelihood of large fire events.

Managers rely on a variety of fire-danger indicators to make their forecasts of the likelihood of large fire events at various lead times. For example, the National Fire Danger Rating system (Bradshaw et al. 1984) produces fire-danger maps for 1-day-ahead forecasts (in-

Corresponding author address: Haiganoush Preisler, USDA Forest Service, Pacific Southwest Research Station, 800 Buchanan St., West Annex, Albany, CA 94710.
E-mail: hpreisler@fs.fed.us

DOI: 10.1175/JAM2513.1

formation was available online at http://www.wfas.us/content/view/17/32/). Seasonal (3–6 months ahead) fire risk forecast maps are produced by the Mapped–Atmosphere–Plant–Soil System team (information was available online at http://www.fs.fed.us/pnw/about/programs/mdr/mapped.shtml). Fire-danger indicators are useful for describing burning conditions at a particular location and time. However, their utility may be enhanced if their relationship with actual fire occurrence and spread is quantified. Specifically, a useful tool to managers will be a system whereby it is possible to forecast probability of a large fire event given a particular value of a fire-danger index or a list of values of various fire-danger predictors.

In this work we present a statistical model for estimating the probability of a large fire event given a list of fire-danger predictors. We define a large fire event as an area greater than 400 ha burning in a 1° grid cell in a month. Other definitions of a large event may also be used. We will demonstrate how these models may be used to produce fire-danger maps that are based on the forecast probability of large fire events. We use historic fire occurrence and fire weather predictors to demonstrate the utility of this statistical tool for estimating 1-month-ahead forecasts. Similar statistical models may be used to estimate and study the utility of probability forecasts with longer leads. The latter will require availability of long-range forecasts of weather predictors.

Several climatic factors with apparent demonstrated effects on wildfire occurrence and size were used as predictors in this study. Numerous authors have reported relationships between wildfire and moisture anomalies concurrent with and antecedent to the fire season (see, e.g., Balling et al. 1992; Swetnam and Betancourt 1998; Kipfmueller and Swetnam 2000; Veblen et al. 2000; Donnegan et al. 2001; Heyerdahl et al. 2001; Westerling et al. 2003b). Swetnam and Betancourt (1990, 1998) and Westerling and Swetnam (2003) discuss the use of indices of Pacific Ocean sea surface temperatures [El Nino–Southern Oscillation and Pacific decadal oscillation (PDO)] as predictors for wildfire activity. Westerling et al. (2002, 2003a,b) demonstrate the utility of the Palmer drought severity index (PDSI) in forecasting wildfire area burned for wildfires in a variety of vegetation types, and Westerling et al. (2006) and Westerling (2007) have examined the relationship among temperature, seasonal temperature forecasts, and forest wildfire.

For this research, we demonstrate the use of the statistical model for studying the utility of 1-month-ahead forecast temperature values, PDSI values for the preceding 12 months, El Niño–Southern Oscillation, and the PDO in forecasting 1-month-ahead probabilities of large fire events. One of the strong points of the statistical model is the ability to produce forecasts with specified precision and, as a consequence, the ability to study the accuracy of the outputs when compared with historic data.

2. Data

This work relied on fire history datasets compiled from federal land management agency fire reports. Westerling et al. (2003a) compiled a gridded 1° latitude $\times$ 1° longitude dataset of monthly fire starts and the size of area burned from approximately 350 000 fire reports reported by the U.S. Department of Agriculture (USDA) Forest Service and the U.S. Department of the Interior Bureaus of Land Management and Indian Affairs and the National Park Service for 1980–2004.

Average monthly temperature and PDSI from U.S. climate divisions (NCDC 1994) were projected onto a 1° grid to provide a monthly climate record for each grid cell. PDSI is an index of combined precipitation, evapotranspiration, and soil moisture that represents cumulative precipitation and temperature anomalies (Alley 1985; Guttman 1991). The index is far from being a perfect proxy for soil moisture (Alley 1984; Karl and Knight 1985), but it appears to be well correlated with wildfire dangers in the western United States, in particular in numerous studies (see, e.g., Balling et al. 1992; Larsen and MacDonald 1995; Swetnam and Betancourt 1998, Westerling et al. 2002, 2003a; Westerling and Swetnam 2003). PDSI is convenient for these purposes because of both its easy availability and the fact that it is a normalized index, the values of which provide a moisture index comparable across a diverse landscape.

We also employ two indices describing patterns of sea surface temperatures in the Pacific Ocean known to be associated with multiyear- to decadal-scale variability in western climate: El Niño–Southern Oscillation (the Niño-3.4 index, hereinafter Niño), and the PDO (Gershunov and Barnett 1998; Dettinger et al. 1998). Historical values of these indices were obtained from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (information was available online at http:// www.cpc.ncep.noaa.gov/data/indices/).

3. Statistical methods

a. *Estimating wildfire-danger probabilities*

We were interested in estimating wildfire danger at a given location and time. As a measure for fire danger we used the probability that an area greater than some specified value (e.g., 400 ha) will burn in a given 1° grid cell during a given month (henceforth referred to as the probability of a large fire event). We used a two-stage evaluation procedure to estimate the probability of a large fire event. In stage one we estimated the probability of at least one fire occurring in a given grid cell in a given month, that is, the probability of "ignition." In stage two we estimated the conditional probability, given at least one ignition, that an area greater than a specified value burns, that is, the probability of "spread" or "escape." Last, the probability of a large fire event was evaluated by multiplying the probability of at least one ignition with the conditional probability of escape given ignition.

We used logistic regression techniques with piecewise polynomials to estimate the probabilities of interest as functions of explanatory variables (predictors; see appendices A and B for further details). The explanatory variables used were monthly average temperature, (forecast from previous months), PDSI value in the previous month, maximum PDSI in the last 12 months, and values of Niño and PDO in addition to location (latitude, longitude) and month. The use of piecewise polynomials, rather than logistic regression with linear terms, was necessitated by the fact that many of the relationships between the probabilities and the explanatory variables appear to be nonlinear. For example, the relationship between the probability of ignition and month is nonlinear, with more fires (higher probabilities) occurring in the middle of the fire season (summer months) than at the beginning or end of the season. Similarly, the relationships between fire occurrence or size and location is likely to be complex and nonlinear. For example, it does not appear realistic to assume that the probability of fire occurrence changes linearly as one travels from south to north (with latitude) or from east to west (with longitude). It is more realistic to expect different regions to have different probabilities depending on vegetation, topography, and other variables specific to a given location. Spatially explicit variables account for most of the correlations between nearby grid cells. For example, neighboring grid points are more likely to have similar vegetation and fuel type, and hence a similar probability of large fire events. The procedure proposed here allows the inclusion of a location effect (a smooth function of latitude and longitude) that acts as a surrogate for locally specific predictors (e.g., elevation, fuel type) that are not included in the model. Probabilities estimated using this model have a unique value for each grid cell and each date (month and year).

Historic averages (persistence forecasts) may be estimated using the frequency of times (during the 25 yr) when a large fire event was observed at a given location and date. An alternative, more efficient procedure is one that takes advantage of the fact that nearby locations (and dates) are similar (correlated). A logistic model with only location and time as predictors is such a procedure. Forecasts produced from a model with only location and time as predictors have a unique value for each grid cell and each month; however, the forecasts do not change from year to year.

Comparing forecast probabilities for a given time (using a model with time-varying predictors) with forecast probabilities from the persistence model makes it possible to quantify departures from "normal conditions" and to study the skill of the model with predictors relative to a model based on historic averages (see section c below).

b. *Forecasting monthly temperatures*

One of the predictors in our model was 1-month-ahead forecast temperatures. These temperature forecasts were obtained by using 114 yr (1890–2004) of data on monthly temperature values from the NOAA Climate Prediction Center. We used an autoregressive model (arima module in the statistical package R, which was available online at http://lib.stat.cmu.edu/R/CRAN/) with month as a predictor to forecast 1-month-ahead temperature values for each climate division in the United States. The autoregressive model produced reasonable forecasts for 1-month-ahead temperatures when the procedure was used to forecast historic values (Fig. 1a). Although similar forecast values may have been obtained for the PDSI, the skill of the forecast was adequate only when the change in PDSI was small (Fig. 1b). As a consequence, in this study we used the previous month's PDSI values as a predictor. This limited our estimations of fire danger to only 1-month-ahead forecasts. We hope to be able to produce longer-range probability forecasts using the same statistical procedures as longer-range forecasts of predictors become available. For example, the methods may be used to study the skill of North Pacific sea surface temperatures (Alfaro et al. 2004, 2006) or the skill of forecast national fire-danger indices from weather models (Roads et al. 2005).

c. *Fire-danger maps*

We produced two types of fire-danger maps. The first was based on the 1-month-ahead forecast probabilities of large events using the following rule: fire danger in a given grid cell at a given time was defined as low if the probability of an area >400 ha burning is less than 10%,

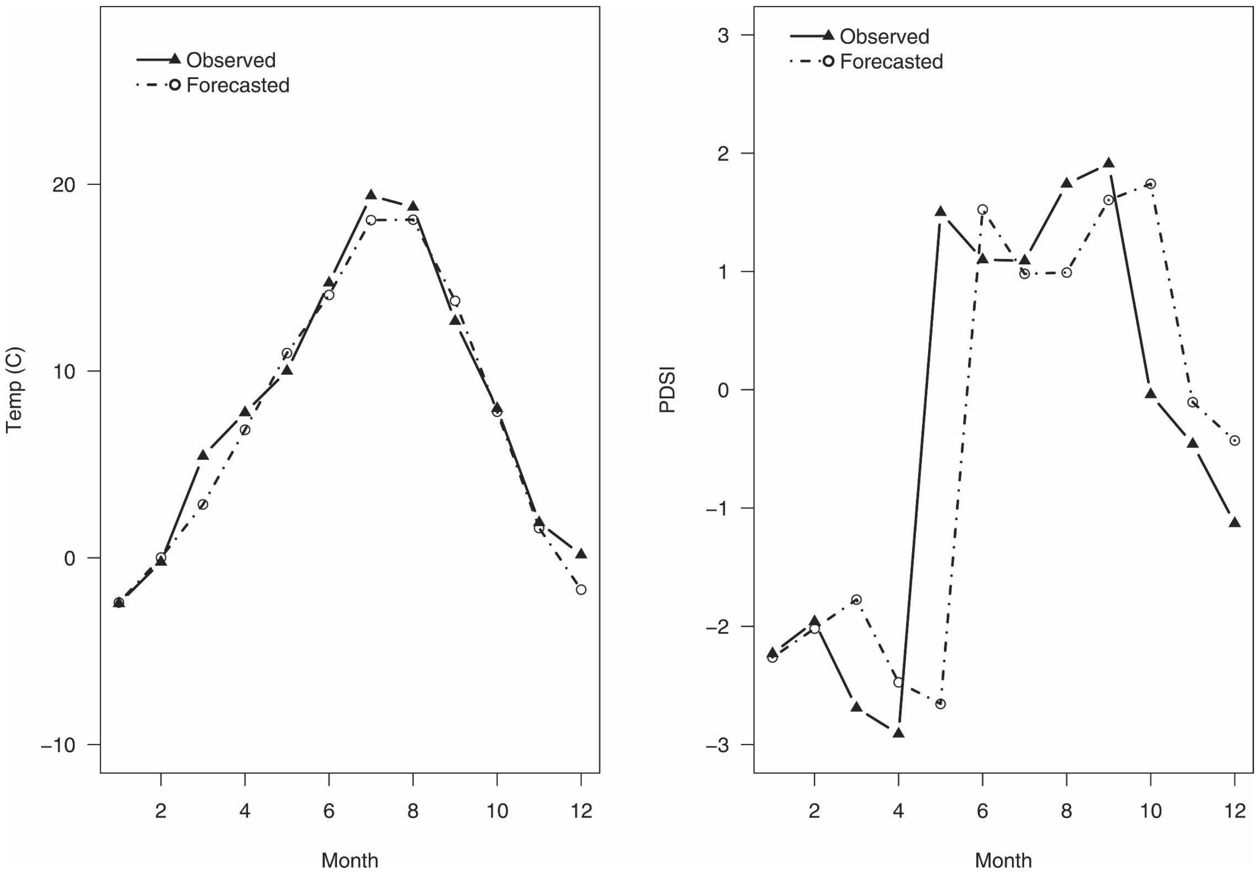

FIG. 1. Observed and 1-month-ahead forecasts of (left) temperature and (right) PDSI for 1 climate division and 1 yr.

as moderate if the probability of an area >400 ha burning is between 10% and 30%, as high if the probability of an area >400 ha burning is between 30% and 50%, and as extreme, if the probability of an area >400 ha burning is greater than 50%.

The formulas used to define the four danger levels are given in appendix C. The size of areas burned (400 ha) and the cutoff probabilities used above are for demonstration purposes. Managers may decide on other cutoff points for what may be considered a large fire event or acceptable levels of risk.

The second set of danger maps was produced to demonstrate whether the forecast probabilities are above or below the norm, where the norm is defined as the average probability for that location and date as estimated by the persistence model. In particular, these maps were based on the ratio of the odds of a large fire event relative to the historic monthly odds using the persistence forecast. The odds of an event are defined as the ratio $\pi/(1 - \pi)$, where π is the probability of the event occurring. Maps of fire danger were produced by designating each grid cell as being either low, normal, or high, depending on whether the odds of a large fire event was less, equal, or higher than the historic odds. Formulas for producing these maps are given in appendix D. Once more, managers may decide to use other definitions for what may constitute above- or below-average fire danger based on probabilities of large fire events.

d. *Model appraisal*

Model appraisal was done by comparing the observed frequencies of events with forecast probabilities using 25 yr of historic data. For binary data, as is the case here, the observations (0 or 1) need to be grouped, based on some criteria, and then a fraction of responses in each group is compared with the average forecast value for that group (Hosmer and Lemeshow 1989). Graphs of observed versus forecast probabilities are sometimes referred to as "reliability diagrams" (Wilks 1995).

There are many ways to group binary response data. We used three different groupings. Each grouping enabled us to study the performance of the forecasts at a different scale. The first grouping was done according to month and forecast probability level. All grid cells in a given month with an observed large fire event were grouped by the forecast probability level (low, medium, high, extreme). This grouping produced a table of the fraction of times an actual large fire event was observed when the 1-month-ahead forecast was low or high, and so on, thus allowing one to demonstrate the skill of the predictors.

The second grouping was done according to spatial location and month. For each month and each 1° grid cell, the number of years (out of 25) in which a large fire event was observed was compared with the forecast number of large fire events. Next, maps were generated for each month that highlight grid cells where the observed number of events was outside the 95% confidence bounds of the forecast numbers. Estimated confidence bounds included natural variation (binomial) and variation resulting from the error in the estimated model parameters. Maps generated in this fashion demonstrated the skill of the model for forecasting events at a given location. For example, if the maps show particular regions with many grid cells where the observed values are outside (below or above) the forecast confidence bounds, then one would conclude that the model is under- or overpredicting the outcomes in that region.

Last, we produced reliability diagrams by grouping all of the data into cells with similar forecast probabilities (similar within ~5% of each other). These diagrams demonstrate the overall performance of the model and, in particular, the skill of the model relative to the skill of the persistence forecasts.

In all cases the forecast probabilities for each grid cell were estimated using cross validation. Specifically, predictions for a given year were done by estimating the model parameters from all other years except the year being predicted.

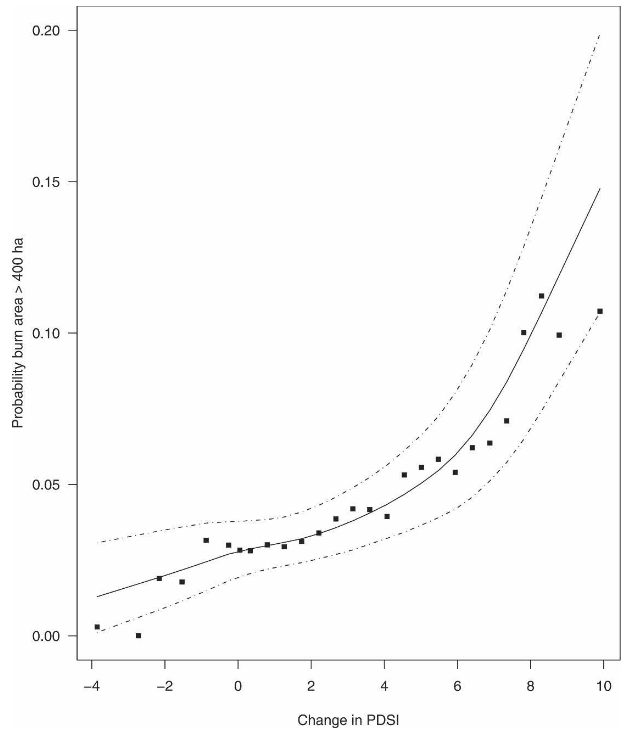

FIG. 2. Observed (dots) and predicted (solid curve) fraction of cases with large fire events plotted against the difference between maximum PDSI in the past 12 months and PDSI at present. Dashed lines are the approximate pointwise 95% bounds.

4. Results

a. Estimated effects of predictors

The explanatory variables (predictors)—forecast temperature, PDSI in previous month, and maximum PDSI in last 12 month, in addition to spatial location and month—had significant effects (P value $\ll 0.01$) on both probabilities of fire occurrence and of fire spread given ignition. On average, fire danger (probability of ignition and spread) appeared to increase with decreasing (i.e., drier) PDSI. That is, dry conditions that would foster fuel flammability appear to be important. Also, fire danger appeared to be increasing as the maximum PDSI of last 12 month increases. This indicates an apparent increase in fire danger when there are large shifts from high to low PDSI values (Fig. 2). The latter is consistent with previous research, indicating that antecedent moisture anomalies are important for fire risks in ecosystems where the availability and continuity of fine fuels (e.g., grasses) can be a limiting factor for the ignition and spread of fires (Swetnam and Betancourt 1998; Westerling et al. 2003a)

The variables Niño and PDO also had significant, albeit small, effects on the probabilities of large fire events. The pattern of these effects depended on the location. For example, in the southern region (latitude <42°) the probability of fire occurrence seemed to increase with increasing values of PDO and the probability of fire spread seemed to increase with decreasing values of Niño. No such affects were apparent in the northern region of the study area (latitude >42°). PDO is a decadal influence with a multidecadal cycle, and is significantly correlated with Niño over the model estimation period. As a consequence, the 25 yr of fire history available to estimate our statistical model is not likely to be sufficient to resolve a PDO signal in the fire data. Furthermore, El Niño–Southern Oscillation described by Niño is most likely to be associated with fire danger via its association with interannual variability in

TABLE 1. Relative frequency of locations with observed large fire events (>400 ha) between 1980 and 2004 in the western United States for each month and in each forecast category.

Forecast probability							
Level	Range	May	Jun	Jul	Aug	Sep	Oct
Low	($P < 0.1$)	0.01 (7094)*	0.02 (4900)	0.03 (2635)	0.02 (2943)	0.03 (5581)	0.02 (7512)
Medium	($0.1 \leq P < 0.3$)	0.07 (784)	0.11 (2616)	0.11 (3899)	0.14 (3715)	0.08 (2067)	0.12 (409)
High	($0.3 \leq P < 0.5$)	0.26 (47)	0.33 (409)	0.36 (1370)	0.38 (1253)	0.20 (277)	0.50 (4)
Extreme	($P \geq 0.5$)	— (0)	— (0)	0.76 (21)	0.64 (14)	— (0)	— (0)

* Total number of 1° square grid cells over 25 yr.

winter precipitation, the effects of which are already incorporated into the PDSI variables included in the model. Therefore, model coefficients for these variables may not be directly interpretable.

b. Model skill

The skill of the model in producing reliable fire-danger ratings is demonstrated in Table 1 where the observed number of large fire events in each of the forecast levels was produced for the period 1980–2004. According to values in Table 1, the model never produced an extreme forecast, or false alarm, during the 25 yr of the study for the months of May, June, September, or October. For the month of July the model forecast extreme danger levels 21 times. In 16 out of the 21 cases (76%) an area greater than 400 ha did burn, and in 10 out of the 21 cases (48%) an area greater than 4000 ha burned (latter not shown in table). At the other end of the table, during July, the model forecast low danger in 2635 cases. Out of those we observed 79 cases (3%) where an area greater than 400 ha burned (missed events) and 0 cases (0%) where an area greater than 4000 ha burned (latter not shown in table).

The skill of the model in forecasting probabilities of an area greater than 400 ha burning at a particular location is demonstrated in the maps in Fig. 3. Maps of significant differences between observed and forecast probabilities at the 1° grid scale showed no apparent spatial patterns. For example, for the months of August and October (Fig. 3) the spatial pattern of significant departure from forecast values appeared to be random. Considering all 12 months (not shown), in 3.1% of the grid cells the observed frequency of large events was outside the upper 95% confidence bound. The above is an indication that there is no apparent spatial bias in the predicted probabilities. There are no apparent regions where the number of misclassified cases is higher or lower than expected.

Figure 4 shows the reliability diagrams for the persistence (historic average) forecasts and the model forecasts. Model forecasts appear to be an improvement on the persistence forecasts in two ways: first, fewer observed points fall outside the model-forecast confidence

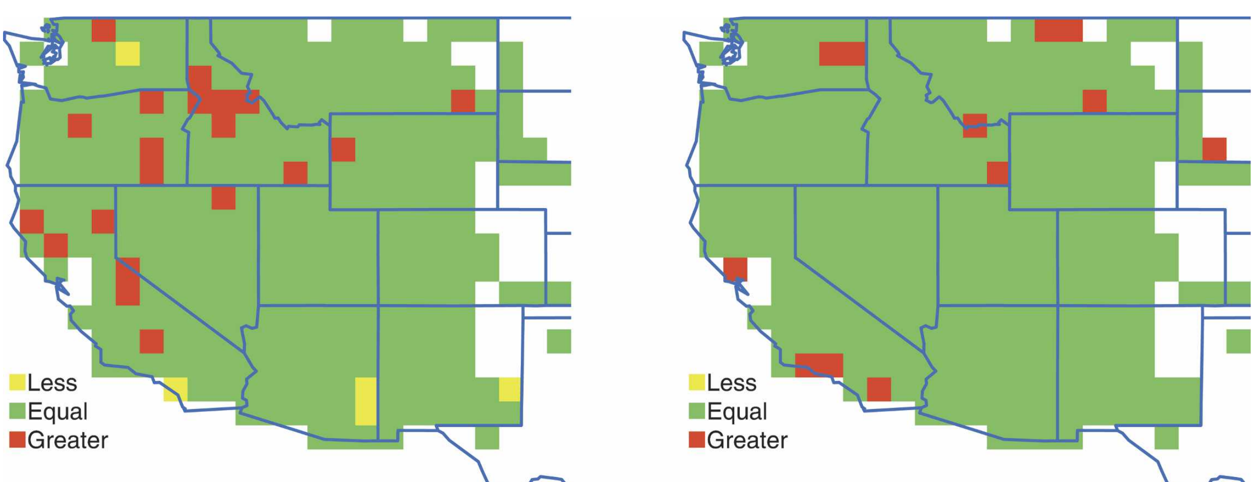

FIG. 3. Comparing observed and forecast number of year (out of 25) with large fire events for (left) August and (right) October. Cells marked as less (yellow) are those where an observed number of years with large events was below the 95% forecast bounds. Cells marked as greater (red) are those with an observed number of events larger than the 95% forecast bound.

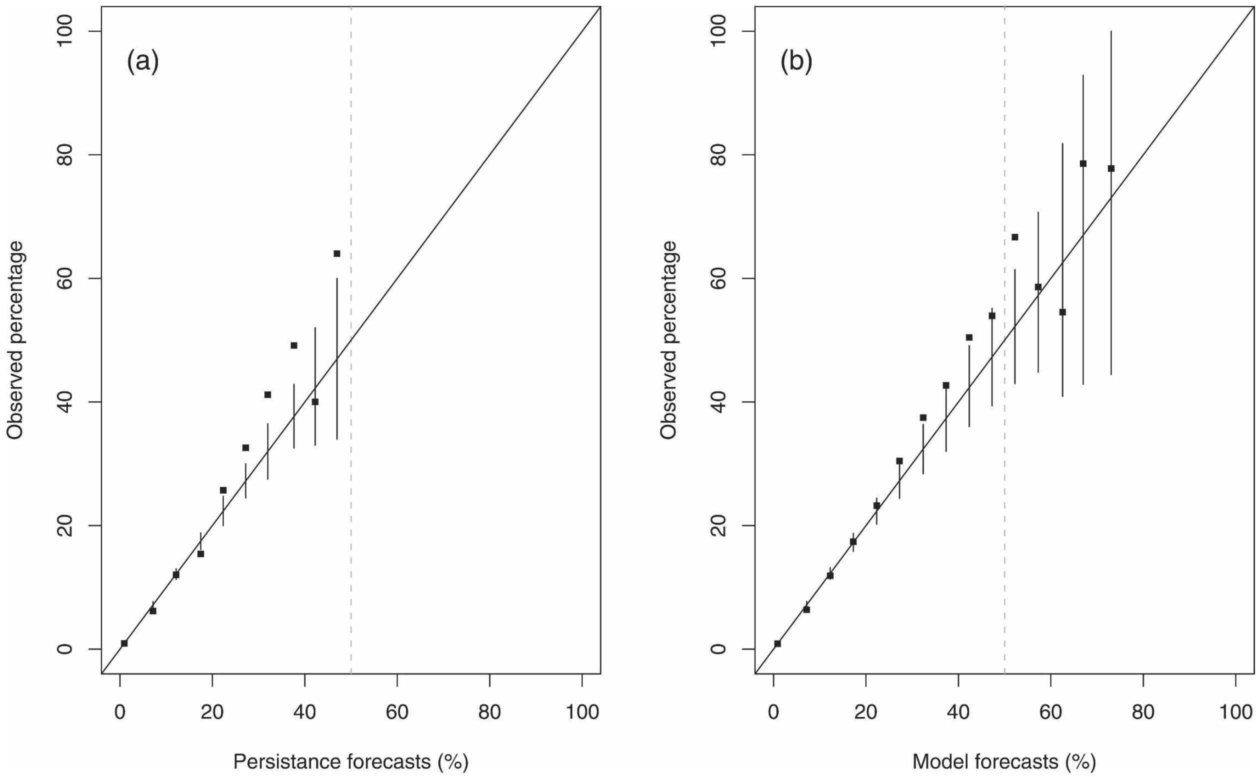

FIG. 4. Observed vs forecast probabilities of large fire events grouped according to similar values of (a) persistence probabilities and (b) model-forecast probabilities. Vertical lines are estimated pointwise 95% confidence bounds.

bounds than the persistence-forecast bounds; second, forecast model probabilities spanned a wider range of values (0%–76%) than those of the persistence forecasts (0%–47%). For a model with almost perfect forecasting skill, forecast probabilities will range from 0 to 1, with most values concentrated near 0 and 1. For a model with no skill there will be no range in forecast values. The forecast will be the same regardless of either location or date. For example, such a forecast for western United States will be 3.7%, that is, the total fraction of times a large fire event was observed in a 1° grid in the past 25 yr. The larger confidence bounds (Fig. 4) for larger forecast values are due to smaller numbers of observations in these groupings.

c. Probability maps

Maps of estimated persistence probabilities of large fire events (area burned >400 ha), that is, maps of historic averages, display the areas in the west of the United States with the highest probabilities of fire (Fig. 5). Over the last 25 yr (1980–2004), areas with the highest probabilities in August appear to be in southern California and around southern Idaho. In October, only one grid cell in southern California and one in southern Idaho had estimated probabilities greater than 10%. These maps demonstrate an overall pattern of large fire events over the last 25 yr.

Examples of fire-danger maps using forecast probabilities and forecast odds ratios for 2 yr demonstrate some of the outputs from our modeling (Fig. 6). Managers may be interested in studying such maps for a range of past years in order to further assess the utility of the probability model in their management practices. For example, the forecasts for August 2000 predicted higher-than-average odds for most areas in Utah and Colorado (orange regions Fig. 6a). This result can also be deduced if one compares the estimated persistence probabilities (Fig. 5) with the estimated probabilities for August 2000 (Fig. 6b). Note that higher-than-normal odds do not necessarily imply high probabilities of large fire events. Compare, as an example, the estimated odds and estimated probabilities for August 2000 for Colorado, Arizona, and New Mexico (Figs. 6a,b). This emphasizes the need to look at not only departures from normal conditions but at actual probabilities of large fire events. As an example, suppose the probability of an event at a particular location and time is 0.1% and the historic average for that location is 0.01%; then, the odds are 11 times the historic average,

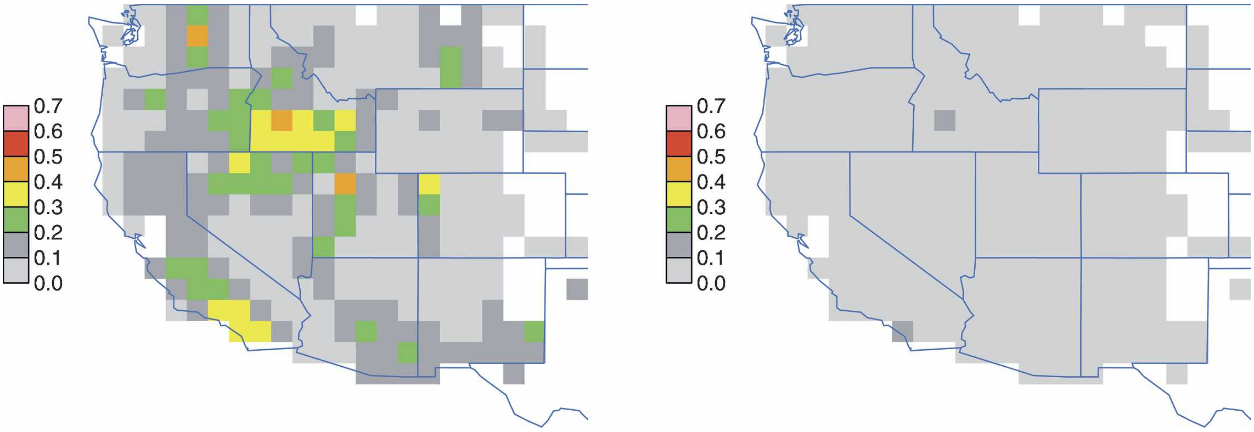

FIG. 5. Estimated persistence (historic average) probabilities of large fire event (>400 ha burned) in 1° cells in the western United States for (left) August and (right) October.

and yet the likelihood of a large fire event is still very small.

The second example is for August 1987 (Figs. 6c,d). During this period the odds for the regions of northern California and southwestern Oregon were mostly normal or higher than normal and the forecast probabilities (in particular, in the grid cells along the coast) were low (less than 20%), and yet many large fire events were observed in this region for this time. In other words, managers acting on these forecasts may have been unprepared for these events. August 1987 was associated with an anomalously large number of dry lighting strikes. Such instances are unavoidable with most any forecasting method. However, using the present probability forecasts one is able to quantify (and consequently attempt to minimize) the frequency of times that large events are missed. From Table 1, it is seen that with the present model large fire events were missed (forecast as low) 2% of the time during the month of August. If this failure rate is not acceptable, then it may be lowered by changing the criteria for a low forecast. However, that may be at the cost of forecasting many more cases with a moderate probability of a large fire event when in realty the probability is very small.

5. Discussion

In this paper we present methods for estimating, forecasting, and mapping fire danger. We found the methods useful for assessing the skill of predictor variables in forecasting 1-month-ahead large fire events. PDSI values in the previous 12 months together with forecast next-month temperatures were useful indicators for forecasting fire probabilities 1 month ahead. The methods are not limited to these predictors or to only 1-month-ahead forecasts. A similar study may be done with forecast fire-danger and fire weather indices (Burgan 1988; Roads et al. 2005) in order to study the skill of 3- or 6-month-ahead forecast indicators on future fire danger. In addition, recent advances in forecasting temperature and PDSI using observed North Pacific sea surface temperatures (Alfaro et al. 2004, 2006; Westerling 2005) could also be used in conjunction with the methods demonstrated here to produce high-resolution seasonal forecasts of wildfire danger with specified precisions.

One important feature of the model is its ability to develop a fire-danger rating system (based on the probabilities being low or high, and so on) where the manager can tell with confidence what the overall error rate will be; that is, how often a large fire event will be missed because the predicted forecast was low, or how often a large fire event will actually be observed when an extreme forecast is predicted. A second feature of the model is the facility with which the limitations of the model with a particular set of predictors may be studied. For example, with the present predictors, during the month of October in the southern California region observed large fire events appear to be higher than forecast for the 25 yr of the study (Fig. 3). One explanation is that Santa Ana winds (a variable not included in the model) are an important component of wildfire risk in southern California in the autumn and winter months (Keeley 2004; Westerling et al. 2004), and are especially salient prior to the start of winter precipitation. Consequently, they are likely a significant source of variability in fire risks in October in coastal southern California that is not captured by the variables used in our model (PDSI and temperature, ENSO, and PDO).

In a similar way, numerous large fires in northern

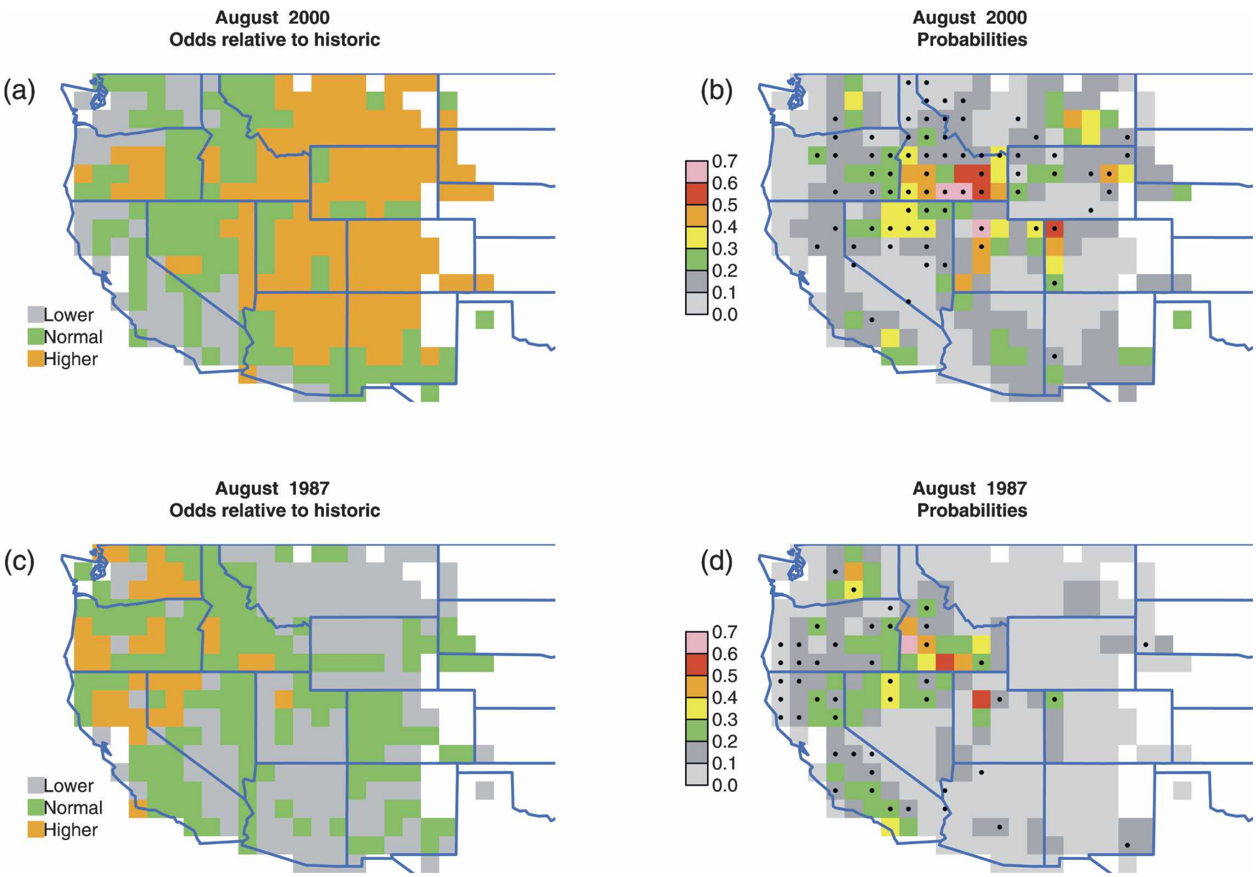

FIG. 6. Two examples of probability-based fire-danger maps. (a), (c) Maps of odds indicate forecast odds for a particular month and year relative to the historic odds of a large fire event (>400 ha). (b), (d) Maps are the forecast probabilities of large fire events. Dots in (b) and (d) indicate the cells with observed area burned >400 ha.

California and southern Oregon in August 1987 were associated with an anomalously large number of dry lighting strikes in inaccessible terrain during a moderate drought. A large number of ignitions in difficult terrain can overwhelm local suppression resources, and with appropriate climatic conditions, fires can subsequently grow to a size that is difficult to suppress by the time additional resources are mobilized from outside the area. While the variables used in our model can describe climatic conditions that make live and dead vegetation conducive to the ignition and spread of fires, they do not capture variability in the source of ignitions. Factors like dry lightning are currently beyond the scope of monthly or seasonal lead-time forecasts, and represent a source of error in forecasting wildfire that is not presently reducible.

There are many factors that affect fire occurrence and spread. Many of them are not systematically measured, and thus insufficient data are available for their inclusion in regional-scale analyses and models. Unmeasured sources of variation are handled by including a stochastic element in the models. Using a statistical model, such as the one described herein, is one way of quantifying not only the deterministic part of the process (relating to the predictors in the model) but also the remaining "unexplained" variability. The latter will enable managers to make decisions with specified precisions by producing estimates of the error rate associated with a given decision rule.

Acknowledgments. This work was supported by the NOAA Office of Global Programs via the Regional Integrated Science and Assessment program and by National Fire Plan grants through the USDA Forest Service's Southern and Rocky Mountain Research Stations.

APPENDIX A

Probability Models

Let Y_{ijk} be the (random) size of the area burned in grid i ($i = 1, \ldots, I$), month j ($j = 1, \ldots, 12$), and year

k (k = 1980, . . . , 2004). Let N_{ijk} be the (random) number of fires in grid i, month j, and year k. Let $\mathbf{X}_{ijk}$ be a matrix of explanatory variable values for grid i, month j, and year k. We define probability of fire danger as

$$\pi_{ijk} = \Pr(Y_{ijk} > C|\mathbf{X}_{ijk}, \boldsymbol{\theta}), \tag{A1}$$

where C is a critical size of interest (e.g., 400 ha) and $\boldsymbol{\theta}$ is a vector of parameters. The probability in Eq. (A1) may be written as a product of two probabilities,

$$\pi_{ijk} = \Pr(N_{ijk} > 0|\mathbf{X}_{ijk}, \boldsymbol{\theta}_1) \times \Pr(Y_{ijk} > C|\mathbf{X}_{ijk}, N_{ijk} > 0, \boldsymbol{\theta}_2). \tag{A2}$$

APPENDIX B

Logistic Regression with Splines

We used logistic regression techniques with piecewise polynomials to estimate the probabilities in Eq. (A2). To be specific, we estimated the regression relationship

$$\mathrm{logit}(p_v) = \beta_o + g_1(\mathrm{lat}_v, \mathrm{lon}_v) + g_2(\mathrm{month}_v) + \sum_{m=1} g_{m+2}(X_{mv}), \tag{B1}$$

where the subscript v indicates the 1° × 1° × 1-month voxel, p is either one of the probabilities on the right side of Eq. (A2) , (lat, lon) are the latitude and longitude of the midpoint of the 1° grid cell, and X_m are explanatory variables. The terms $g()$ are semiparametric smooth functions (Hastie et al. 2001) such as piecewise polynomials, periodic splines (for estimating month-in-year effect), and thin-plate splines (for estimating the spatial surface as a function of lat, lon). The statistical package R (more information was available online at http:www.r-project.org) has a module bs() that determines the basis functions of a given vector. Once the basis functions are determined one may use any linear or logistic regression routine because the model is linear in these new expanded variables. For the spatial component g(lat, lon), we utilized the two-dimensional version of the basis function, that is, the thin-plate spline function. The required modules for fitting thin-plate splines within R were downloaded from the Internet (Geophysical Statistical Project; information available online at http://www.cgd.ucar.edu/stats/Software/Fields). Similar logistic models were used in Preisler et al. (2004), Preisler and Benoit (2004), and Preisler and Westerling (2005).

APPENDIX C

Probability-Based Fire-Danger Maps

Criteria for defining levels of fire danger are as follows:

$$\begin{array}{lll} \text{Low} & \text{if} & \hat{\pi}_1 + 2\hat{\sigma}_1 \le \alpha_1 \\ \text{Moderate} & \text{if} & \alpha_1 \le \hat{\pi}_1 + 2\hat{\sigma}_1 < \alpha_2 \\ \text{High} & \text{if} & \alpha_2 \le \hat{\pi}_1 + 2\hat{\sigma}_1 < \alpha_3 \\ \text{Extreme} & \text{if} & \hat{\pi}_1 + 2\hat{\sigma}_1 \ge \alpha_3 \end{array}, \tag{C1}$$

where $\hat{\pi}_1$ are the forecast probabilities [using Eqs. (A1), (A2), and (B1)] of area burned, in a particular grid cell and date, being greater than C_1; $\hat{\sigma}_1$ are jackknife standard error estimates (Efron 1982) of $\hat{\pi}_1$, and α_k (k = 1, 2, 3) are arbitrary probability cutoff points. For the purposes of this study, we used the cutoff values 0.1, 0.3, and 0.5 and C_1 = 400 ha (~1000 acres).

APPENDIX D

Maps of Departure from Normal Conditions

Let $\hat{\theta} = \log(\hat{\gamma})$ be the logarithm of the estimated odds relative to historic odds. To be specific,

$$\gamma = \frac{\pi}{1-\pi} \div \frac{\pi_H}{1-\pi_H},$$

with π being the model-forecast probability and π_H being the probability based on historic averages (persistence forecasts). Fire-danger maps were produced using the following rules:

$$\begin{array}{lll} \text{Lower than historic} & \text{if} & \hat{\theta}_1 + 2\hat{\sigma}_1 < 0 \\ \text{Normal} & \text{if} & -2\hat{\sigma}_1 \le \hat{\theta}_1 \le 2\hat{\sigma}_1. \\ \text{Higher than historic} & \text{if} & \hat{\theta}_1 - 2\hat{\sigma}_1 > 0 \end{array} \tag{D1}$$

REFERENCES

Alfaro, A., A. Gershunov, D. Cayan, A. Steinemann, D. Pierce, and T. Barnett, 2004: A method for prediction of California summer air surface temperature. *Eos, Trans. Amer. Geophys. Union,* **85,** 553–558.

——, ——, and ——, 2006: Prediction of summer maximum and minimum temperature over the central and western United States: The role of soil moisture and sea surface temperature. *J. Climate,* **19,** 1407–1421.

Alley, W. M., 1984: The Palmer drought severity index: Limitations and assumptions. *J. Climate Appl. Meteor.,* **23,** 1100–1109.

——, 1985: The Palmer drought severity index as a measure of hydrologic drought. *Water Resour. Bull.,* **21,** 105–114.

Balling, R. C., G. A. Meyer, and S. G. Wells, 1992: Relation of surface climate and burned area in Yellowstone National Park. *Agric. For. Meteor.,* **60,** 285–293.

Bradshaw, L. S., J. E. Deeming, R. E. Burgan, and J. D. Cohen, Eds., 1984: The 1978 National Fire-Danger Rating System: Technical documentation. U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station General Tech. Rep. INT-169, 44 pp.

Burgan, R. E., 1988: 1988 revisions to the 1978 National Fire-Danger Rating System. USDA Forest Service, Southeastern Forest Experiment Station Research Paper SE-273, 39 pp.

Dettinger, M. D., D. R. Cayan, H. F. Diaz, and D. M. Meko, 1998: North–south precipitation patterns in western North America on interannual-to-decadal timescales. *J. Climate,* **11,** 3095–3111.

Donnegan, J. A., T. T. Veblen, and S. S. Sibold, 2001: Climatic and human influences on fire history in Pike National Forest, central Colorado. *Can. J. For. Res.,* **31,** 1527–1539.

Efron, B., 1982: *The Jackknife, the Bootstrap and Other Resampling Plans.* Society for Industrial and Applied Mathematics, 92 pp.

Gershunov, A., and T. P. Barnett, 1998: Interdecadal modulation of ENSO teleconnections. *Bull. Amer. Meteor. Soc.,* **79,** 2715–2725.

Guttman, N. B., 1991: Sensitivity of the Palmer hydrologic drought index. *Water Res. Bull.,* **27,** 797–807.

Hastie, T. J., R. Tibshirani, and J. Friedman, 2001: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 533 pp.

Heyerdahl, E. K., L. B. Brubaker, and J. K. Agee, 2001: Factors controlling spatial variation in historical fire regimes: A multiscale example from the Interior West, USA. *Ecology,* **82,** 660–678.

Hosmer, D. W., and S. Lemeshow, 1989: *Applied Logistic Regression.* John Wiley and Sons, 307 pp.

Karl, T. R., and R. W. Knight, 1985: *Atlas of Monthly Palmer Hydrological Drought Indices (1931–1983) for the Contiguous United States*. Historical Climatology Series Vol. 3-7, National Climatic Data Center, 319 pp.

Keeley, J. E., 2004: Impact of antecedent climate on fire regimes in coastal California. *Int. J. Wildland Fire,* **13,** 173–182.

Kipfmueller, K. F., and T. W. Swetnam, 2000: Fire–climate interactions in the Selway–Bitterroot wilderness area. *Wilderness Science in a Time of Change Conference,* D. N. Cole et al., Eds., Vol. 5, *Wilderness Ecosystems, Threats, and Management,* USDA Forest Service, 270–275. [Available online at http://tree.ltrr.arizona.edu/~tswetnam/tws-pdf/Kipfmueller_Swetnam.pdf].

Larsen, C. P. S., and G. M. MacDonald, 1995: Relations between tree ring widths, climate, and annual area burned in boreal forests of Alberta. *Can. J. For. Res.,* **25,** 1746–1755.

NCDC, 1994: Time bias corrected divisional temperature–precipitation–drought index. Documentation for Dataset TD-9640, 12 pp. [Available from DBMB, NCDC, NOAA, Federal Building, 37 Battery Park Ave., Asheville, NC 28801-2733.]

Preisler, H. K., and J. W. Benoit, 2004: A state space model for predicting wildland fire risk. *Proc. Joint Statistical Meeting,* Toronto, ON, Canada, American Statistical Association, 2972–2977.

——, and A. L. Westerling, 2005: Estimating risk probabilities for wildland fires. *Proc. Joint Statistical Meeting,* Minneapolis, MN, American Statistical Association, 2482–2490.

——, D. R. Brillinger, R. E. Burgan, and J. W. Benoit, 2004: Probability based models for estimating wildfire risk. *Int. J. Wildland Fire,* **13,** 133–142.

Roads, J. O., F. M. Fujioka, S.-C. Chen, and R. E. Burgan, 2005: Seasonal fire danger forecasts for the USA. *Int. J. Wildland Fire,* **14,** 1–18.

Swetnam, T. W., and J. L. Betancourt, 1990: Fire-Southern Oscillation relations in the Southwestern United States. *Science,* **249,** 1017–1020.

——, and ——, 1998: Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest. *J. Climate,* **11,** 3128–3147.

Veblen, T. T., T. Kitzberger, and J. Donnegan, 2000: Climatic and human influences on fire regimes in ponderosa pine forests in the Colorado Front Range. *Ecol. Appl.,* **10,** 1178–1195.

Westerling, A. L., 2007: Climatology for wildfire management. *Economics of Forest Disturbances—Wildfires, Storms, and Invasive Species*, T. P. Holmes, J. P. Prestemon, and K. L. Abt, Eds., Springer-Verlag, in press.

——, and T. W. Swetnam, 2003: Interannual to decadal drought and wildfire in the western United States. *Eos, Trans. Amer. Geophys. Union,* **84,** 545–560.

——, A. Gershunov, D. R. Cayan, and T. P. Barnett, 2002: Long lead statistical forecasts of western U.S. wildfire area burned. *Int. J. Wildland Fire,* **11** (3–4), 257–266.

——, T. J. Brown, A. Gershunov, D. R. Cayan, and M. D. Dettinge, 2003a: Climate and wildfire in the western United States. *Bull. Amer. Meteor. Soc.,* **84,** 595–604.

——, A. Gershunov, and D. R. Cayan, 2003b: Statistical forecasts of the 2003 western wildfire season using canonical correlation analysis. *Experimental Long-Lead Forecast Bulletin,* Vol. 12, No. 1–2, Institute of Global Environment and Science, 71–75.

——, D. R. Cayan, T. J. Brown, B. L. Hall, and L. G. Riddle, 2004: Climate Santa Ana Winds and autumn wildfires in southern California. *Eos, Trans. Amer. Geophys. Union,* **85,** 289–296.

——, H. Hidalgo, D. R. Cayan, and T. Swetnam, 2006: Warming and earlier spring increases western U.S. forest wildfire activity. *Science,* **313,** 940–943.

Wilks, D. S., 1995. *Statistical Methods in Atmospheric Sciences*. Academic Press, 467 pp.