UCLA
Department of Statistics Papers

Title
A Critical Assessment of the Burning Index in Los Angeles County, California

Permalink
https://escholarship.org/uc/item/4vv7v978

Authors
Schoenberg, Frederic P
Chang, Chien-Hsun
Keeley, J E
et al.

Publication Date
2007-01-29

eScholarship.org Powered by the California Digital Library
University of California

RH: Burning index in Los Angeles

A Critical Assessment of the Burning Index in Los Angeles County, California

Frederic Paik Schoenberg[A,E], Chien-Hsun Chang[A], Jon E. Keeley[C,D], Jamie Pompa[A], James Woods[B], and Haiyong Xu[A].

TOC Summary: The effectiveness of the Burning Index (BI) in predicting wildfire activity is assessed using 25 years of area burned data from Los Angeles, California. The BI predicts poorly compared to simple alternatives using just a few weather variables.

Key words: burn area, model evaluation, point process, wildfire, wind.

[A] Department of Statistics, 8142 Math-Science Building, University of California, Los Angeles, CA 90095.

[B] Department of Geography, California State University, Long Beach, CA 90840.

[C] USGS, Western Ecological Research Center, Sequoia-Kings Canyon National Parks, Three Rivers, CA 93271.

[D] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095.

[E] Corresponding author. Telephone: 310-794-5193; email: frederic@stat.ucla.edu

Abstract

The Burning Index (BI) is commonly used as a predictor of wildfire activity. An examination of data on the BI and wildfires in Los Angeles County, California from January 1976 to December 2000 reveals that although the BI is positively associated with wildfire occurrence, its predictive value is quite limited. Wind speed alone has a higher correlation with burn area than BI, for instance, and a simple alternative point process model using wind speed, relative humidity, precipitation and temperature well outperforms the BI in terms of predictive power. The BI is generally far too high in Winter and too low in Fall, and may exaggerate the impact of individual variables such as wind speed or temperature during times when other variables, such as precipitation or relative humidity, render the environment ill-suited for wildfires.

1 Introduction.

Fire danger is a concept that figures prominently in fire management planning. It is an assessment of the fire environment that determines the ease of ignition, rate of spread, difficulty of control and fire impact, and these are often referred to as the fire potential. Fire danger rating is an integration of weather elements, fuels and other factors affecting fire potential and is often used to direct fire management activities, scheduling prescribed fires, fire prevention activities, staffing for fire control, and forest closures (Pyne et al. 1996). It is expressed as a numerical index and weighting of the factors that make up this index vary regionally, due in part to differences in fire regimes. For example, Australia has long used the McArthur Forest Fire Danger Index (FFDI) (McArthur 1967), Canada the Fire Weather Index (FWI), and the USA the National Fire Danger Rating System (NFDRS).

Each fire danger rating uses weather observations at a fixed site and broad generalizations about fuels and other landscape characteristics to assess fire potential regionally, such as for entire national forests. Recently, inventories of fuel distributions have the potential for creation of more specific fire danger assessments (Woodall et al. 2005), and as specificity increases, fire behavior models provide an even finer scale prediction of fire danger (Andrews et al. 2003).

The NFDRS produces three main indices, the fire Spread Component (SC), the Energy Release Component (ERC), and the Burning Index (BI). SC and ERC are controlled largely by fuel structure, weather and antecedent drought, and both contribute to BI, the fireline intensity or flame length. This is of immense concern because it gives some indication of the potential for escape and eventual fire size as well as the possible destructiveness of a fire. Fire managers use this information in making decisions about the appropriateness of prescribed burning or alerts for increased preparedness, both in terms of fire suppression staffing and fire prevention activities. Since fireline intensity is an important factor in predicting fire containment and the likelihood of fire escape, the Burning Index is the rating of most interest to many fire managers. This is particularly true for natural crown-fire ecosystems such as southern California shrublands, where BI is commonly employed to assess fire danger (Mees and Chase 1991).

Evaluating the predictability of fire danger measures is important to improving and fine-tuning these indices for different regions. For example, the NFDRS system, initially designed for western systems, was revised after the first decade of use indicated it was not a good predictor of fire incidence in the more mesic eastern forests (Burgan 1988). The main fire danger indices have been developed for forested ecosystems and often they are less applicable to shrublands. For example the Australian McArthur FFDI was originally developed to describe fire danger in dry sclerophyll forests, but it is a poor predictor of area burned in mallee shrublands (Krusel et al. 1993). In these natural crown-fire ecosystems, a simple model utilizing a few meteorological variables offered vast improvement over the McArthur system (Krusel et al. 1993).

Evaluating the importance of fire danger rating systems to fire managers could be accomplished in a variety of ways, ranging from evaluating the cost-effectiveness of staffing decisions to historical patterns of burning associated with annual changes in these fire danger indices. Considering the extraordinary importance of fire danger in many parts of the world, relatively few studies have investigated the effectiveness of fire danger rating systems. Most of these studies have focused on relating fire danger rating systems to ultimate fire responses, including fire incidence and fire size, and since these rating systems are largely based on climate parameters, these studies have focused on historical relationships. Indeed, Andrews and Bradshaw (1997), whose work was instrumental in the current implementation of the BI, suggested that the value of a fire danger index be evaluated according to its relationship with fire activity, which may be defined as the incidence of large wildfires. The use of fire danger ratings by fire department officials for wildfire suppression and prevention may confound the empirical relationship between fire danger ratings and observed wildfire activity. Such empirical relationships have nevertheless been reported and used as support for the use of such rating systems for predictive purposes. For instance, Haines et al. (1983) and Haines et al. (1986) examined the NFDRS Ignition Component and Spread

Component (as well as a few other Fire Danger Rating Systems) in the northeastern US with respect to the number of fires per day, the total area burned per day, the occurrence of at least one wildfire on a day, and the occurrence of at least one fire of at least a certain size. They found statistically significant relationships between ranked indices and these four response variables, though regression equations typically explained less than a quarter of the variation. Mees and Chase (1991) observed a positive association between percentiles of BI with percentiles of burn area and number of fires in Southern California and concluded that the BI is successful as an indicator of potential fire workload. Mandallaz and Ye (1997a,b) and Viegas et al. (1999) assessed the performance of various European and Canadian fire danger indices in predicting wildfire activity in southern Europe and the former found that using weather variables and other covariates in addition to a fire danger index could produce substantial improvement in prediction of wildfires. Andrews et al. (2003) evaluated the US NFDRS indices and fire activity in the Tonto National Forest in Arizona, finding that a logistic function of ERC predicts wildfire incidence better than a logistic function of BI.

Here we evaluate the empirical relationship between the Burning Index and wildfire activity in Los Angeles County, California. This region is well-suited to such historical analysis because the Los Angeles County Fire Department and Department of Public Works have collected and compiled a wealth of data on the locations burned by large wildfires dating back over a century. The landscape in Los Angeles County is vulnerable to high intensity crown-fires due to the vast expanse of dense contiguous chaparral shrublands under the Mediterranean-type climate of summer and fall drought (Keeley 2000). An additional factor is that the dry season is followed in the Fall by high winds known locally as Santa Ana winds (Keeley and Fotheringham 2003). These offshore winds reach speeds exceeding 100 kph at a relative humidity below 10%, and are annual events lasting several days to several weeks, creating the most severe fire weather in the United States (Schroeder et al. 1964). We show that although the BI is positively associated with wildfire incidence in Los Angeles County, its empirical performance in predicting wildfire activity is poor relative to simple alternatives involving direct use of the weather variables incorporated by the Burning index.

2 Data

Since 1976, the Forest Service has implemented and monitored Remote Automatic Weather Stations (RAWS) across the United States. We focus here on data from 16 stations located within Los Angeles County, California. Each of the RAWS records information on a host of meteorological variables, including air temperature, relative humidity, precipitation, wind speed, and wind direction (Warren and Vance 1981). Daily summaries of these measurements are collected each afternoon at 1300 hr and transmitted via satellite to a central station for archiving. For some of the variables, such as relative humidity and temperature, not only is the current 1300 hr value used, but also its maximum and minimum values over the previous 24 hours. In what follows, for simplicity we refer to the maximum relative humidity from 1300 hr on day $t-1$ to 1300 hr on day t as simply the *maximum relative humidity for day t*.

Daily RAWS measurements are used to construct several wildfire indices, as part of the United States Fire Danger Rating System (NFDRS), which was developed by the USDA Forest Service in 1972 and revised in 1978 (Deeming et al. 1977; Bradshaw et al. 1983). Since then there have been some minor adjustments (Burgan 1988). The NFDRS relies on RAWS data, a manager-selected, stylized fuel model considered suitable for representing local conditions, estimates of live fuel moisture and four size classes of dead fuel moisture calculated based on a combination of weather data as well as prior fuel stick moisture readings, slope class and herbaceous fuel type information, in order to construct a Spread Component (SC) and an Energy Release Component (ERC). These two components are combined to produce the Burning Index, based on a suite of physics-based nonlinear dynamic equations. These equations describe heat transfer and moisture exchange, and were developed to predict flame length in a fire given that ignition has already occurred (see Pyne et al. 1996, Andrews and Bradshaw 1997).

Of the values produced by the NFRDS, the Burning Index is the most commonly employed, and is used by about 90% of station managers according to the USDA Forest Service, including those in Los Angeles County. Here, we examine the daily BI values, computed using FireFamily Plus software, which was made freely available from the Forest Service via www.fire.org. The local fuel model used in the computation of the BI was that for mature chaparral (fuel model B). For several of the 16 RAWS, data were missing on certain days in the time range considered here (January 1976 - December 2000); proportions and potential impacts of this missing data for certain stations and months are discussed in Peng et al. (2005). In what follows, for any particular day we consider the BI values and weather variables averaged across the number of RAWS which were available on that day, as in Mees and Chase (1991) and Peng et al. (2005). The distribution of the BI values is right-skewed: the mean BI recorded during this period is 69.7 and the median is 54. Mean daily BI records ranged from 0 to 351 during this 25-year period, with a standard deviation of 55.3. The

Los Angeles County Fire Department (LACFD) generally considers BI values above 150 to indicate a high likelihood for wildfire activity.

Detailed data on Los Angeles County wildfires have been collected and compiled by several agencies, including the LACFD, the Los Angeles County Department of Public Works (LACDPW), the Santa Monica Mountains National Recreation Area, and the California Department of Forestry and Fire Protection. These data include the origin dates and polygons demarking locations burned in wildfires dating back to 1878. According to the LACFD, the wildfire data before 1950 is believed to be complete for fires burning at least 0.405 km^2 (100 acres), and though the LACFD has been mapping fires as small as 0.00405 km^2 (1 acre) since 1950, the data appear to be complete only for fires burning at least 0.0405 km^2 (Schoenberg et al. 2003b). We restrict our attention here to the 592 wildfires of at least 0.0405 km^2 (10 acres) recorded between January 1976 and December 2000. During this time period, there were 362 days with exactly one such wildfire recorded, and 66 days with two or more such wildfires. For further detail and images of the spatial locations of these wildfires, see Peng et al. (2005).

3 Methods.

Because BI presumably reflects on the probability of containment, particularly at initial attack, and this is tied to the number of large fires and ultimate fire size, one measure of the predictive value of the Burning Index is the historical relationship of variation in the index to area burned. The relationships between BI and daily outcomes such as number of wildfires ignited on that day or the area burned by those fires are initially inspected using basic exploratory analysis such as correlations and conditional means and medians. For comparison, corresponding summaries of the relationships between these outcome variables and individual weather variables or smoothed historical (pre-1976) wildfire activity for each date are reported.

In order to investigate ways in which the BI's relationship with burn area or number of wildfires may be improved, particular attention is paid to seasonal patterns in the relationship between BI and these outcome variables to see if there are months when the BI tends to over-predict or under-predict wildfire activity in Los Angeles County. Since fuel moisture is thought to be an important factor influencing wildfire activity, we compare the cumulative precipitation on preceding days (a proxy for fuel moisture) on days when wildfires occurred with all other days on which BI values were similar. In particular, we inspect the average daily precipitation over the k preceding days, for all n days on which wildfires occurred and on which the BI was in a particular range, for various values of k. Since the distribution of cumulative precipitation over prior days appears to be approximately normal, the results are then compared with 95%-confidence bounds obtained using $\bar{x} \pm 1.96 s/\sqrt{n}$, where $\bar{x}$ and s are the sample mean and standard deviation, respectively, of the average precipitation over the k preceding days for all days on which the BI was in the corresponding range.

One way to assess the association between the BI and wildfire activity is to investigate the fit and predictive performance of different marked point process models for wildfire incidence, where BI is a covariate in the model. In order to shed light on whether the BI uses weather information optimally in the sense of predicting wildfire activity, we compare these results with marked point process models where the weather variables recorded by the RAWS are covariates. Note that in point process modeling, there is no need to choose either number of wildfires or area burned as a response variable: one may readily model the entire process, including the origin dates, centroid locations, and areas burned for each fire, as a marked point process. That is, essentially the response variable is the number of wildfires of each possible size.

A point process N is a random measure defined on subsets of the observed space. For a bounded subset B of space-time, for instance, the random variable N(B) represents the number of points occurring in B. In a marked point process, to each point there corresponds a random variable whose value may depend on the time, location, and history of points of the process. We consider here the origin times and centroid locations of LACFD wildfires as points of the point process, with area burned in the fire as a mark. Equivalently, one may consider the domain of observation to be a portion of space-time-area burned, and thus each wildfire represents one point in this domain. A point process may be characterized by its conditional intensity or rate λ; here $\lambda(t, x, y, a)$ indicates the limiting expected rate of occurrence of wildfires of area a occurring on day t at location (x, y), conditioned on the historical information available prior to time t. We will use the term *rate* in place of intensity here in order to differentiate the limiting expected number of points per unit of space-time-area from a physical attribute of a wildfire. For more information on definitions, models, and methods related to point processes and conditional rates, see Daley and Vere-Jones (1988).

One way to address the fact that not all locations in Los Angeles County are equally likely to burn is to estimate a spatially-varying background rate of wildfires, $\mu(x, y)$. Such a function μ can be estimated by kernel smoothing the observed wildfires. However, there is an especially serious risk of over-fitting if all fires in the dataset are used for this purpose. Some authors estimate such a background rate using only the largest events (Ogata 1998) or using a historical catalog of prior events (Peng et al. 2005). For instance, in Los Angeles County, since wildfire histories are available dating back to 1878, one option would be to kernel-smooth the spatial locations of fires occurring before 1976 in determining the spatial background rate of wildfires occurring since 1976. Since there are thought to be severe missing data problems before 1950, we elected to smooth the centroids of fires occurring between 1950 and 1975 in estimating μ. As noted in Pompa et al. (2006), one problem with simply kernel-smoothing the centroids of historical fires using a standard (e.g. Gaussian) kernel is that the resulting estimate of $\mu(x, y)$ will be non-zero for locations (x, y) outside of the boundary of Los Angeles County, including locations in the Pacific Ocean. One way to address this while still giving each point in the historical dataset equal weight is to allow the kernel and/or bandwidth to vary spatially. For instance, a simple adjustment is to truncate the kernel used to smooth a historical fire whose centroid is at location (x, y), letting $k(u) = 0$ for $u > d$, where d is the shortest distance from (x, y) to the boundary of Los Angeles County. This suggests a background model of the form

$$\lambda(t, x, y, a) = \alpha\mu(x, y)g(a), \tag{1}$$

where α is a parameter to be estimated, $\mu(x, y)$ is the spatial background rate estimated by kernel smoothing the 1950-1975 fires as in Pompa et al. (2006), and $g(a)$ represents the distribution of wildfire areas. The tapered Pareto distribution may be chosen for $g(a)$ as suggested in Schoenberg et al. (2003b).

In order to incorporate BI measurements, one may also consider a model such as

$$\lambda(t, x, y, a) = f(B(t))\mu(x, y)g(a), \tag{2}$$

for some function f, where $B(t)$ represents the average BI on day t. For instance, for f one might select a linear function, so that $f(B(t)) = \alpha_1 + \alpha_2 B(t)$, where α_1 and α_2 are parameters to be estimated.

As an alternative to the model (2), one may consider a model using the RAWS weather variables directly. For instance, one might compare the model (2) with a model such as

$$\lambda(t, x, y, a) = f(R(t))\mu(x, y)g(a), \tag{3}$$

where $R(t)$ represents the maximum relative humidity for day t. In order to incorporate more weather variables, one might consider a model such as

$$\lambda(t, x, y, a) = f_1(R(t))f_2(W(t))f_3(P(t))f_4(A(t; k))f_5(T(t))f_6(D(t))\mu(x, y)g(a), \tag{4}$$

where $R(t), W(t), P(t), T(t)$ represent the maximum relative humidity, 1300 hr wind speed, 1300 hr precipitation, and 1300 hr temperature, respectively, for day t, averaged over all available weather stations, $A(t; k)$ represents the average precipitation on the k days prior to day t, and $D(t)$ is the date within the year corresponding to day t. The variables in this model represent a subset of the variables recorded at each RAWS station on each day for which the BI is computed, and hence the direct use of such information might be considered a suitable basis for comparison with the BI. The variables listed above were selected based on their apparent significance in the output of simple linear regressions of daily burn area on each of the weather variables.

The model (4) is entirely multiplicative, or *separable* in the terminology of Cressie (1993). Such a multiplicative form may be reasonable in light of the fact that high values of certain variables such as windspeed do not individually result in significantly increased wildfire activity unless the conditions as measured by other weather variables are also conducive to wildfires. As a result, the output of a model such as (4) will tend to be considerably higher on days when each of the variables is moderate than on days when one variable is very highly conducive to wildfires and the others are not. Tests for separability in point processes were proposed in Schoenberg (2004), based on comparing bivariate kernel estimates of the rate as a function of two variables or coordinates with the product of the two corresponding univariate kernel rate estimates. An example is the statistic S_3, which is defined as the integrated squared difference between these two kernel rate estimates. Confidence bounds for S_3 may be obtained by re-sampling from the separable (product) rate estimate, as described in Schoenberg (2004). Such tests are used here to assess whether the form in (4) appears to be appropriate.

In the case that the separability hypothesis is seriously violated for a pair of coordinates, an alternative type of model proposed in Schoenberg (2006) is to partition one of the coordinates and fit a separable model for each segment of the partition. For instance, one possible problem with the model (4) is that the distribution of the areas of wildfires may vary significantly with season, so that $D(t)$ and $g(a)$ are not separable. The relations between the weather variables and wildfire activity may similarly depend on the season in question. One possible solution is to consider a model such as

$$\lambda(t,x,y,a) = f_1(R(t)) f_2(W(t)) f_3(P(t)) f_4(A(t;k)) f_5(T(t)) \mu(x,y) g(a),$$

for each season. That is, for each season within the year, a different set of parameters govern the functions f_i and g. The full model then becomes

$$\begin{aligned}\lambda(t,x,y,a) &= f_1\{R(t);\beta(D(t))\}f_2\{W(t);\beta(D(t))\}f_3\{P(t);\beta(D(t))\} \\ &\times f_4\{A(t,k);\beta(D(t))\}f_5\{T(t);\beta(D(t))\}\mu(x,y)g\{a;\beta(D(t))\}, \end{aligned} \tag{5}$$

where β, the vector of parameters for the functions f_i and g, may vary depending on the date D corresponding to day t.

A convenient feature of separable point process models is that each component of the model may be estimated individually, and the functional forms in the model can be derived using non-parametric methods. For instance, one may use kernel regression of burn area or number of fires per day on maximum daily relative humidity in order to suggest a functional form for f_1. For the present paper, we use Gaussian kernels with bandwidths determined according to Silverman's formula (Silverman 1986). The parameter vector β may then be estimated using maximum likelihood estimation (MLE), using a standard optimization routine. Here, this estimation is performed using the Newton-Raphson minimization algorithm in the R statistical programming environment, which is freely available from www.r-project.org.

One way to compare two ore more competing point process models is using the Akaike Information Criterion (AIC), which is defined as $-2L(\beta)+2p$, where $L(\beta)$ is the log-likelihood and p is the number of fitted parameters. Lower values of AIC indicate better fit. The AIC thus rewards a model for having a higher likelihood, indicating greater agreement with the data, but penalizes a model for fitting more parameters, as a safeguard against over-fitting.

Competing point process models may also be compared by directly examining their power in predicting wildfires in the historical dataset. Consider issuing an *alarm*, or prediction of a wildfire occurrence, on any day when $\lambda(t)$ is above a certain threshold. For any particular threshold, there is a *success rate* (percentage of wildfires occurring at times when λ is above the threshold) as well as a *false alarm rate* (number of times per year when λ exceeded the threshold yet no wildfires occurred). For any fixed choice of false alarm rate, one may compare the corresponding success rates of the competing models, in order to determine which model appears to have greater efficacy in predicting wildfire activity.

For further evaluation of the model and to assess the sensitivity of the results to overfitting and to the fact that the same data were used in the model fitting and assessment, we use a jackknife procedure in which one year of the dataset is removed at a time. For each year i, the model is fit by MLE using only the data from all years other than year i, and the resulting variation in parameter estimates is inspected.

4 Results.

We find several shortcomings to the Burning Index and its use in predicting wildfire activity. Before detailing the problems with the BI and suggesting alternative ways of using weather variables to predict daily wildfire activity, we begin this Section by discussing the evidence in favor of the BI.

Evidence supporting the BI.

For the Los Angeles County data from 1976-2000 described in Section 2, the Burning Index is indeed positively associated with wildfire incidence. The correlation between the daily average BI score and the daily number of wildfires is 0.147; the correlation between the BI and the total area burned is 0.098. Note that since the BI was designed to indicate the potential for a large wildfire, rather than the potential for mere ignition, one might suspect the BI to correlate more highly with area burned than with the number of wildfires, so it may be somewhat surprising that in fact the opposite is observed. The correlation between BI and the area burned per fire is even lower ($r = 0.084$). Given the large day-to-day variability in wildfire activity and the high dependence of the BI on human activities such as land use, fire prevention and suppression efforts, and arson, a low correlation between wildfire incidence and the BI, which is dependent solely on weather variables, is to be expected.

Figure 1 confirms that the BI is higher on days when wildfires occur. The upward shift in the distribution of BI from days without wildfires to days when one wildfire occurs is readily apparent in Fig. 1, and a similar shift from days when one wildfire occurs to days when two or more occur can also be discerned. The right-most boxplot in Fig. 1 shows the distribution of BI on the days when the largest 50 wildfires during this 25-year period occurred. The median BI increases from 52.4 on days without wildfires to 81.9 on days when one fire occurred, jumping to 109.5 when one restricts one's attention to days when two or more wildfires occurred, and reaching 134.3 on the days of the 50 largest wildfires.

Discrepancies between BI and Wildfire Activity.

Although the BI is positively associated with wildfire incidence, the correlation is quite low, and it is readily apparent that much better prediction of wildfire activity can be achieved quite simply. In fact, while daily burn area has a correlation of 0.098 with BI, the correlation between burn area and wind speed alone is 0.153. Furthermore, one can obtain a correlation with daily burn area nearly half as large as that of the BI, using only the date and information on areas burned in past years as a guide, without any weather data at all. If for each calendar date, one takes a kernel smoothing of the average burn area on that date from the years 1951-1975, for example, as illustrated by the dashed curve in Fig. 2, the resulting daily estimate has a correlation of 0.046 with daily burn area during 1976-2000.

One obvious problem with the BI is that it is generally too high in Winter months (especially January) and too low in the Fall (especially September and October). The light grey curve in Figure 2 shows the daily averages of BI by calendar date, averaged over the years 1976-2000. One sees that the BI is typically highest in late summer, and takes on moderate values in December, January, and February. These values in Winter months are surprisingly high, given that the number and sizes of wildfires during these months are historically very low. Indeed, in the 25 years from 1976-2000, only 25 wildfires greater than 0.0405 km^2 (10 acres) were recorded in these three winter months, and these fires account for just 2.03% of the total area burned in all months.

Figures 3 and 4 illustrate the extent to which the BI overpredicts wildfire activity during Winter and underpredicts wildfire activity in Fall. The numbers above and to the left of the regression line in Fig. 3 indicate months in which the BI was too low: one sees that in July, November, and especially September and October, the average area burned was higher than one would have predicted using the BI. Numbers below and to the right of the line in Fig. 3 indicate months when the BI overpredicted wildfire potential: August, December and January are months with less wildfire activity than one would expect using the BI. Figure 4 focuses especially on October, November, December and January, four months when the problems with the BI seem to be among the most severe. One would expect the total area burned on days when the BI was in a certain range to be similar for these four months, but this is not the case. Among days with similar BI values, the area burned is typically higher in October and November than in December and January. The difference is especially prominent for days when the BI is above 200. When such days occur in October and November, wildfire activity tends to be quite high, but very little is burned on such days in December and January.

One possible explanation for the discrepancy between Fall and Winter is the impact of fuel moisture. The BI, which incorporates the drought severity index in addition to the weather on the particular day, does not seem to adequately take into account the cumulative effect of precipitation (or dry weather) on previous days. Figure 5 shows that, when comparing days when wildfires occurred with other days with comparable BI values, the average precipitation per day over preceding days is considerably lower for the days preceding wildfires. Consider, for instance, days when at least one wildfire occurred and when the BI was between 100 and 150. For such days, the mean daily precipitation over the preceding days is indicated by the dark, thin, solid curve in Fig. 5, and the corresponding 95%-confidence bounds are indicated by the dark, thin dashes. As detailed in Section 3, these confidence bounds indicate a range that would be expected to contain the solid curve if, among days when the BI was between 100 and 150, days when wildfires occurred were similar to days without wildfires. Hence the fact that the average precipitation on preceding days is outside the range for at least 5 days of prior precipitation indicates that the BI's use of precipitation and drought severity data is not optimal in terms of predicting wildfire activity: among days with comparable BI values, the days on which wildfires occur appear to have significantly lower fuel moistures than days on which wildfires do not occur. Similarly, comparison of the light, thin solid curve with its 95%-confidence bounds (light, thick dashes) indicates that for days when the BI is between 150 and 200, days with wildfires have significantly lower prior precipitation over the k preceding days than days without wildfires, for k greater than 20. Note that most of the weather variables are highly correlated, so the effect suggested in Fig. 5 might not be causal: it may be that most wildfires occur during relatively dry seasons due to other factors associated with dryness, such as high temperatures and low relative humidity.

Comparison using simple point process models

In order to evaluate how effectively the BI takes into account the weather information for the purposes of predicting wildfire activity, the fit of point process models using the BI (2) is compared with the fit of other models such as (4) and (5) which use the weather variables directly.

Kernel regression plots of burn area and number of fires per day versus BI suggest a linear function f for model (2); i.e.

$$\lambda(t, x, y, a) = \{\beta_1 + \beta_2 B(t)\}\mu(x, y)g(a), \tag{6}$$

where μ and g are estimated as described in Section 3. Similar kernel regression plots of burn area and number of daily fires against each of the weather variables suggested exponential forms for the functions f_i in the models (4)-(5), suggesting a point process model of the form $\lambda(t, x, y, a) =$

$$\beta_1 \exp\{\beta_2 R(t) + \beta_3 W(t) + \beta_4 P(t) + \beta_5 A(t; 60) + \beta_6 T(t) + \beta_7 [\beta_8 - D(t)]^2\}\mu(x, y)g(a). \tag{7}$$

Tests of separability of the components of model (7) showed significant departures from separability for the area burned distribution $g(a)$ and the date within the year $D(t)$. The statistic S_3 of Schoenberg (2004) had a value of 1.431 ($p < 0.01$), indicating that the distribution of burn area changes significantly with season. However, the distribution appears to be relatively constant within each of the following three periods: (a) May 1 to August 31; (b) September 1 - November 30; and (c) December 1 to April 30. This suggests a model such as (5) with exponential forms for f_i, so that

$$\lambda(t, x, y, a) = \beta_1^{(i)} \exp\{\beta_2^{(i)} R(t) + \beta_3^{(i)} W(t) + \beta_4^{(i)} P(t) + \beta_5^{(i)} A(t; 60) + \beta_6^{(i)} T(t)\} \mu(x, y) g(a), \quad (8)$$

for $i = 1, 2, 3$, corresponding to the three seasons listed above. The separability tests of Schoenberg (2004) did not reveal any other significant departures from the separability of the components of model (8). For instance, Figure 6 shows a comparison of the bivariate kernel smoothing and product of univariate kernel smoothings for temperature and maximum relative humidity, as suggested in Schoenberg (2004); differences between the two plots indicate departures from separability. The shading of each pixel is determined by smoothing the ratio of the number of wildfires occurring on days of the corresponding temperature and maximum relative humidity to the total number of such days in the dataset. While there are some noticeable differences for days with temperatures above 100^o F, the overall similarity of the two plots indicates a lack of significant deviation from the separability assumption in this case.

The maximum likelihood estimates of the parameters for the BI, weather, and date components of the models (6-8) are listed in Table 1. Table 2 shows the AIC values, relative to that of the background model (1), for the models (6), (7), and (8). For further comparison, the AIC for the model (3) with $f(x) = \beta_1 \exp(\beta_2 x)$ is also given. The entries reported in Table 2 are the AIC for the corresponding model minus the AIC for the background model

(1); lower numbers indicate better fit. Note that the fit of the simple model (3), which uses relative humidity data only and ignores all other weather variables, fits only slightly worse than the BI model. While the BI model (6) offers substantial improvement over the background model, the model (8) offers much further improvement over the BI model.

Model	$\beta_1 \times 10^3$	$\beta_2 \times 10^4$	β_3	β_4	β_5	β_6	$\beta_7 \times 10^4$	β_8
BI	14.4	7.49						
(7)	1.492	-232.5	0.07493	-2.358	-0.4158	0.06591	1.376	211.1
(8a)	6.741	-252.2	0.06020	-7.079	-0.2097	0.04899		
(8b)	16.92	-276.7	0.1040	-2.845	-0.2976	0.02341		
(8c)	5.634	-298.8	0.08529	-1.004	-0.3068	0.03198		

Table 1: Parameter estimates for models (6), (7), and (8). *BI* refers to model (6). (8a) represents the parameters for the period May 1 - August 31. (8b) corresponds to the period from September 1 to November 30, and (8c) to the period from December 1 to April 30.

RH	*BI*	*(7)*	*(8)*
-218.3	-221.0	-381.5	-652.9

Table 2: Relative AIC values. Table entries are reported as AIC for the selected model minus AIC for the background model (1). *RH* refers to model (3), with $f(x) = \beta_0 e^{\beta_1 x}$, and *BI* refers to model (6).

We also examined an alternative form of (2) with f an exponential or logistic function of BI in analogy with Andrews and Bradshaw (1997), but the fit in each case was substantially worse than the fit of (6). Note that Andrews and Bradshaw (1997) also observed poor fit of a model similar to (2), with a logistic function $f(x) = [1 + \exp(\beta_1 + \beta_2 x)]^{-1}$, but without comparison to models based on weather variables.

An alternative way to compare the models (8) and (2) is by directly examining their power in predicting wildfire activity. Recall from Section 3 that, for a given model for λ and a fixed alarm threshold, we define the *success rate* as the percentage of the observed wildfires occurring at times when λ is above the threshold, and the *false alarm rate* as the number of days per year when λ exceeded the threshold yet no wildfires occurred. Figure 8 shows that the model (8) is uniformly more powerful at predicting wildfires for the years 1976-2000 than the BI model. That is, for any choice of false alarm rate, a higher percentage of the wildfires would have occurred on days when alarms were issued using model (8) rather than using model (2).

The discrepancy in predictive efficacy between the two models is in fact quite large. Consider, for instance, the case where an alarm threshold of $BI = 150$ is used. This value is commonly used by fire department personnel in Los Angeles County as a benchmark, and is indicated in Fig. 8 by the dotted vertical line. With this threshold, for the years 1976-2000, one would have issued 33 false alarms per year; 88.5% of the alarms would have been false alarms, and the success rate would have been 23.1% (i.e. 23.1% of the observed wildfires would have occurred on days when alarms were issued). By comparison, if one were to use the model (8) and to choose a threshold corresponding to 33 false alarms per year, the success rate would be 44.3%. This amounts to a 91.8% increase in the success rate. An alarm threshold of $BI = 200$ would yield 13 false alarms per year, for a false alarm rate of 90.0% and a success rate of 9.3%. By comparison, using the model (8) with a threshold corresponding to 13 false alarms per year, 27.2% of wildfires would have occurred on the days of alarms. This represents a 192% improvement in the success rate, compared to the BI model.

Sensitivity Analysis.

The sensitivity of our results is investigated using a jackknife procedure in which each year of data is alternately set aside as a *test* year, and the model (7) is then fitted using the remaining years. Comparison of the mean jackknife estimates in Table 3 with the overall estimates in Table 1 shows that the estimates are very stable. Table 3 also indicates the root-mean-square (RMS) deviation and maximum absolute deviation between the jackknife estimates and the overall estimates in Table 1. The results show that the discrepancies in any particular year are typically very small. For parameters β_2 β_3, β_5, β_6, β_7, and β_8, the maximum deviation in any year was less than 10% of the parameter value itself, and the RMS deviations, indicating the deviations in a typical year, ranged from 0.7% to 2.3% of the overall parameter estimate for these parameters. For β_1 and β_4 the deviations were larger, however, indicating the high variablility in both the overall rate of wildfire incidence from year to year as well as the volatility of the relationship between wildfire activity and precipitation. The parameter β_1 changed most upon removal of the year 1980, the year with the highest number of wildfires in the dataset. The parameter β_4 deviated most when of the year 1997 was removed; 1997 was the year in the dataset with 44.1 cm in total precipitation and contains the day with the highest precipitation record in the dataset (14.5 cm).

	$\beta_1 \times 10^4$	$\beta_2 \times 10^4$	$\beta_3 \times 10^3$	β_4	$\beta_5 \times 10^3$	$\beta_6 \times 10^3$	$\beta_7 \times 10^6$	β_8
Mean	14.98	-232.6	74.88	-2.379	-415.9	65.95	137.6	211.4
RMS	1.533	4.841	1.910	0.1042	9.504	1.522	3.540	0.5569
Max Dev	5.660	13.99	5.943	0.5206	29.67	5.901	8.610	1.512

Table 3: Parameter estimates for model (7) with one year's data removed in each iteration. *Mean* indicates the average and of the estimate over all 25 iterations. *RMS* is the root-mean-squared difference between each of the 25 jackknife estimates and the overall estimate in Table 1. *Max Dev* represents the maximum absolute difference between each of the 25 jackknife estimates and the overall estimate from Table 1.

5 Discussion.

The potential for accurate prediction of future wildfire activity given only daily weather variables is inherently severely limited. Weather is only one of many factors relating to wildfire occurrence and spread, and human interaction variables such as the propensity for arson or, conversely, for fire prevention, obviously play a huge role. Indeed, the use by fire department officials of the BI for allocating fire suppression resources and implementing fire prevention policies may suggest that, even if the BI were a perfect predictor of wildfire susceptibility, its correlation with wildfire incidence or burn area might be limited. Because of such human interactions, if the alternative models discussed here were used by fire department officials in place of the BI, the correlation of such models with wildfire incidence and burn area might decrease.

On the other hand, other studies that have claimed to validate the use of the BI as a predictor of wildfire activity have done so based on its correlation with wildfire incidence (Haines et al. 1983; Mees and Chase 1991). Further, there is some evidence suggesting that fire suppression activities, though critically important in reducing damages and deaths in wildfires, might not dramatically alter the actual amount of area burned per fire (Moritz 1997, Keeley et al. 1999, Johnson et al. 2001, Keeley 2002). There is substantial need for empirical validation of the BI, especially since it is an adaptation not only from a predictor of flame length to an indicator of wildfire activity but also from measurements in well-controlled laboratory experiments to actual wildland fires. Since such indices are used for various purposes including insurance and urban planning as well as fire department resource management (Irby et al. 2000; Pyne et al. 1996), it is important that existing methods be modified in order to predict wildfire potential as accurately as possible.

It is clear that the BI does not make optimal use of the daily weather variables recorded by the RAWS in predicting wildfire activity in Los Angeles County. Wind speed alone has a higher correlation with burn area than the BI. Further, a model using exclusively maximum relative humidity over the previous 24 hours fits nearly as well as the BI. The BI does not make optimal use of drought severity records and precipitation readings, and generally seems to overpredict wildfire activity on days when individual variables such as wind speed or temperature are high, and other variables, such as fuel moisture or precipitation, render the conditions relatively poorly suited for the ignition and spread of wildfires.

A word should be said about the spatial and temporal scales of the analyses conducted here. The problem addressed here is that of predicting wildfire activity on a given day, given the current weather. The BI appears to be commonly used for exactly this purpose. Note, however, that the relationships between certain weather variables and burn area might depend critically on the temporal scale of the analysis. For instance, Schoenberg et al. (2003a) found that burn area increases steadily with temperature up to a plateau, rather than increasing steadily as suggested here by the exponential term in (8). The analysis in Schoenberg et al. (2003a) dealt with data aggregared on a *monthly* scale; when one employs a similar kernel regression to the *daily* burn area versus temperature, no such plateau is readily apparent.

Similarly, the trends observed here might not extend to data on very different spatial scales than that considered here. Los Angeles County is somewhat heterogeneous in its vegetation, climate, topography, and particularly land use, though the spatial region studied here is highly localized compared to other analyses of wildfire prediction methods, such as those mentioned in the Introduction. Averaging BI across all available stations on a given day, as was done here as in the work of previous authors (e.g. Mees and Chase 1991), may introduce bias, since in addition to the problems of spatial heterogeneity and differences in elevation, the data are not missing at random, but rather certain stations have higher rates of missing data during months when wildfires are less frequent. A related source of bias is missing data from the LACFD wildfire records. Since the data are thought to be complete for fires burning at least 0.0405 km^2 (10 acres), this threshold is used as a lower threshold for the present analysis, and as a result the wildfires focused on here may occur predominantly on days with especially high windspeeds and/or with high numbers of wildfire outbreaks so that one or more of the wildfires could not be contained by the initial attack response by LACFD personnel. If a lower minimum fire size threshold were used, the parameters (especially the background rate μ in Tables 1 and 2) would be expected to increase, since the overall rate of wildfires would be higher. For information on missing data and spatial patterns in the dataset considered here as well as suggestions on the optimal spatial interpolation of the BI records, see Peng et al. (2005). Further study is needed on the effects of missing data, lower thresholds for wildfire sizes and study area boundaries on parameter estimates for models such as those considered here.

The simple model proposed in (8) appears to be much more accurate than the BI, predicting 91.8% to 192% more wildfires than the BI given the same number of false alarms per year. However, this should be viewed as a demonstration of the extreme shortcomings of the BI rather than the success of the weather model in (8). The model in (8) is quite simplistic and its fit could no doubt be improved by using more complicated functional forms for each of the terms, as well as considering different interactions between the variables in the model. In the model (8), the rate of wildfires of a given size occurring in a given region on a certain day is assumed to depend on various weather variables but not on prior wildfire activity; hence according to the model, wildfire records on different days are conditionally independent given the weather records on those days. It seems unlikely that inclusion of terms involving clustering or inhibition between wildfires would substantially improve the fit of the model, and indeed our initial efforts at including such terms had little noticeable impact and were removed from this analysis for the sake of simplicity, but since wildfires very rarely occur in locations that have burned recently, the inclusion of spatial-temporal inhibition terms in the model is likely to offer some modest potential for improvement in the fit of the model. Furthermore, many important variables are excluded from the model: not only is it limited to just four of the daily weather variables recorded by the RAWS, but also other variables such as fuel age, fuel moisture, vegetation, land use and other human interaction variables, and numerous other important covariates are omitted from the model. Spatial variations in the weather variables are entirely omitted, as only spatial averages of the variables are considered. Despite all these shortcomings, the model (8) seems rather impressive in its ability to predict wildfires, when compared to the BI. Our intention is for this to speak to the inadequacy of the Burning Index for the purpose of predicting wildfire activity and the need for substantial improvements, rather than to promote the model (8) in particular. The investigation of more complex point process models that may offer superior predictive performance and comparison of our results with other study regions outside of Southern California are important areas for future research.

Acknowledgements. Our sincerest gratitude to Larry Bradshaw at the USDA Forest Service for so generously providing us with the weather station data and helping us process it and generate the corresponding BI values. We also thank Roger Peng who provided valuable help and advice, as well as members of the Los Angeles County Fire Department and the Los Angeles County Department of Public Works (especially Mike Takeshita, Herb Spitzer, and Frank Vidales) for sharing their data and expertise.

References

Andrews, P.L., and L.S. Bradshaw. 1997. FIRES: Fire Information Retrieval and Evaluation System - a program for fire danger rating analysis. Gen. Tech. Rep. INT-GTR-367.

Ogden, UT; U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 64 p.

Andrews, P.L., D.O. Loftsgaarden, and L.S. Bradshaw. 2003. Evaluation of fire danger rating indexes using logistic regression and percentile analysis. International Journal of Wildland Fire 12: 213-226.

Bachmann, A. and B. Allgöwer. 2000. The need for a consistent wildfire risk terminology. In: Proceedings from: the Joint fire Science conference and workshop, "Crossing the millennium: integrating spatial technologies and ecological principles for a new age in fire manage." (eds L.F. Neuenschwander & K.C. Ryan), pp. 67-77. University of Idaho, Moscow, Idaho.

Bradshaw, L.S., J.E. Deeming, R.E. Burgan, and J.D. Cohen. 1983. The 1978 National Fire-Danger Rating System: Technical Documentation. United States Department of Agriculture Forest Service General Technical Report INT-169. Intermountain Forest and Range Experiment Station, Ogden, Utah. 46 p.

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.

Cressie, N.A. 1993. Statistics for Spatial Data, revised ed. Wiley, New York.

Cruz, M.G., M.E. Alexander, and R.H. Wakimoto. 2003. Assessing the probability of crown fire initiation based on fire danger indices. Forestry Chronicle 79(5): 976–983.

Daley, D. and D. Vere-Jones. 1998. An Introduction to the Theory of Point Processes. Springer, NY.

Deeming, J.E., R.E. Burgan, and J.D. Cohen. 1977. The National Fire-Danger Rating System – 1978. Technical Report INT-39, USDA Forest Service, Intermountain Forest and Range Experiment Station.

Garcia Diez, E.L., L. Rivas Soriano, F. de Pablo, and A. Garcia Diez. 1999. Prediction of the daily number of forest fires. International Journal of Wildland Fire 9(3): 207–211.

Gilless, J.K. and J.S. Fried. 1998. Stochastic representation of fire behavior in a wildland fire protection planning model for California. Forest Science 45(4): 492–499.

Haines, D.A., W.A. Main, J.S. Frost, and A.J. Simard, A.J. 1983. Fire-danger rating and wildfire occurrence in the Northeastern United States. Forest Science 29(4): 679–696.

Haines, D.A., W.A. Main, and A.J. Simard, 1986. Fire-danger rating observed wildfire behavior in the Northeastern United States. Res. Pap. NC-274. St. Paul, MN: U.S. Department of Agriculture, Forest Service, 23 p.

Irby, R., F.C. Beall, B. Barrette, and M. Frago. 2000. Wildland Fire Hazard Assessment. Final Report FEMA 1005-47. UC Forest Products Laboratory.

Johnson E.A., K. Miyanishi, and S.R.J. Bridge. 2001. Wildfire regime in the boreal forest and the idea of suppression and fuel buildup. Conservation Biology 15: 1554-1557.

Keeley, J.E. 2000. Chaparral. P. 203–253 in North American Terrestrial Vegetation, 2nd Edition, M.G. Barbour and W.D. Billings, (eds.). Cambridge University Press, N.Y.

Keeley, J. 2002. Fire management of California shrubland landscapes. Environmental Management 29: 395-408.

Keeley, J.E. and Fotheringham, C.J. (2003). Impact of past, present, and future fire regimes on North American Mediterranean shrublands. In: Fire and climatic change in temperate ecosystems of the western Americas, T.T. Veblen, W.L. Baker, G. Montenegro and T.W. Swetnam, editors. Springer, NY, pp. 218–262.

Keeley, J.E., C.J. Fotheringham, and M. Morais. 1999. Reexamining fire suppression impacts on brushland fire regimes. Science 284: 1829–1832.

Krusel, N., D. Packham, and N. Tapper. 1993. Wildfire activity in the mallee shrubland of Victoria, Australia. International Journal of Wildland Fire 3: 217-227.

Mandallaz, D. and R. Ye. 1997a. A new dryness index and the non-parametric estimate of forest fire probabilities. Schweiz. Z. Forstwes. 148(10): 809–822.

Mandallaz, D. and R. Ye. 1997b. Prediction of forest fires with Poisson models. Can. J. For. Res. 27: 1685–1694.

McArthur, A.G. 1967. Fire behavior in eucalypt forest (Leaflet No. 107). Commonwealth of Australia, Canberra, Australian Capital Territory, Forestry and Timber Bureau.

Mees, R. and R. Chase. 1991. Relating burning index to wildfire workload over broad geographic areas. International Journal of Wildland Fire 1: 235-238.

Moritz, M.A. 1997. Analyzing extreme disturbance events: fire in Los Padres National Forest. Ecological Applications 7(4): 1252–1262.

Peng, R. D., F.P. Schoenberg, and J. Woods. 2005. A space-time conditional intensity model for evaluating a wildfire hazard index. Journal of the American Statistical Association

100(469): 26–35.

Pompa, J.L., F.P. Schoenberg, and C. Chang. 2006. A note on non-parametric and semi-parametric modeling of wildfire hazard in Los Angeles County, California. Journal of Environmental and Ecological Statistics, Thematic Issue on Wildfire Modeling, in review.

Pyne, S.J., P.L. Andrews, and R.D. Laven. 1996. Introduction to Wildland Fire, 2nd ed. John Wiley & Sons, Inc., New York, NY.

Rothermel, R.C. 1991. Predicting behavior of the 1988 Yellowstone fires: projections versus reality. Int. J. Wildland Fire 1(1): 1–10.

Schoenberg, F.P. 2004. Testing separability in multi-dimensional point processes. Biometrics 60: 471-481.

Schoenberg, F.P., R. Peng, Z. Huang, and P. Rundel. 2003a. Detection of nonlinearities in the dependence of burn area on fuel age and climatic variables. Int. J. Wildland Fire 12(1): 1–10.

Schoenberg, F.P., R. Peng, R., and J. Woods. 2003b. On the distribution of wildfire sizes. Environmetrics 14(6): 583–592.

Schroeder, M.J., M. Glovinsky, V. Hendricks, F. Hood, M. Hull, H. Jacobson, R. Kirkpatrick, D. Krueger, L. Mallory, A. Oertel, R. Reese, L. Sergius, and C. Syverson. 1964. Synoptic weather types associated with critical fire weather. Institute for Applied Technology, National Bureau of Standards, U.S. Department of Commerce, Washington, D.C.

Silverman, B. W. 1986. Density Estimation. London: Chapman and Hall.

Simard, A.J., J.E. Eenigenburg, and S.L. Hobrla. 1987. Predicting extreme fire potential. P. 148–157 *in* Ninth Conference on Fire and Forest Meteorology, Apr. 21-24, San Diego, CA. American Meteorological Society, Boston, MA.

Turner, M.G. and W.H. Romme. 1994. Landscape dynamics in crown fire ecosystems. Landscape Ecology 9(1): 59–77.

Viegas, D.X., G. Bovio, A. Ferreira, A. Nosenzo, and B. Sol. 1999. Comparative study of various methods of fire danger evaluation in Southern Europe. Int. J. Wildland Fire 9(4): 235–246.

Warren, J.R. and D.L. Vance. 1981. Remote Automatic Weather Station for Resource and Fire Management Agencies. United States Department of Agriculture Forest Service Technical Report INT-116, Intermountain Forest and Range Experiment Station.

Woodall, C.W., J.J. Charney, G.C. Liknes, and B.E. Potter. 2005. What is the fire danger now? Linking fuel inventories with atmospheric data. Journal of Forestry 103: 293-298.

Figure captions:

Figure 1: Boxplots of average daily BI for days with 0 wildfires, days with 1 wildfire, days with 2 or more wildfires, and days when the largest 50 wildfires occurred, respectively. Within each box, the solid horizontal line indicates the median, and small circles indicate outliers.

Figure 2: Smoothed average burn area (BA) and BI per day by date, smoothed using a Gaussian kernel smoother with bandwidths of 11.8 for BA and 15.3 for BI as determined by the formula of Silverman (1986). The dark solid curve corresponds to wildfires occurring between January 1951 and December 1975, and the dashed curve to those from January 1976 to December 2000. The light curve shows the unsmoothed BI values for each calendar date, averaged over the years 1976-2000.

Figure 3: Area burned versus average BI for each month, from 1976-2000. Each number indicates a month (1 = January, 2 = February, etc.).

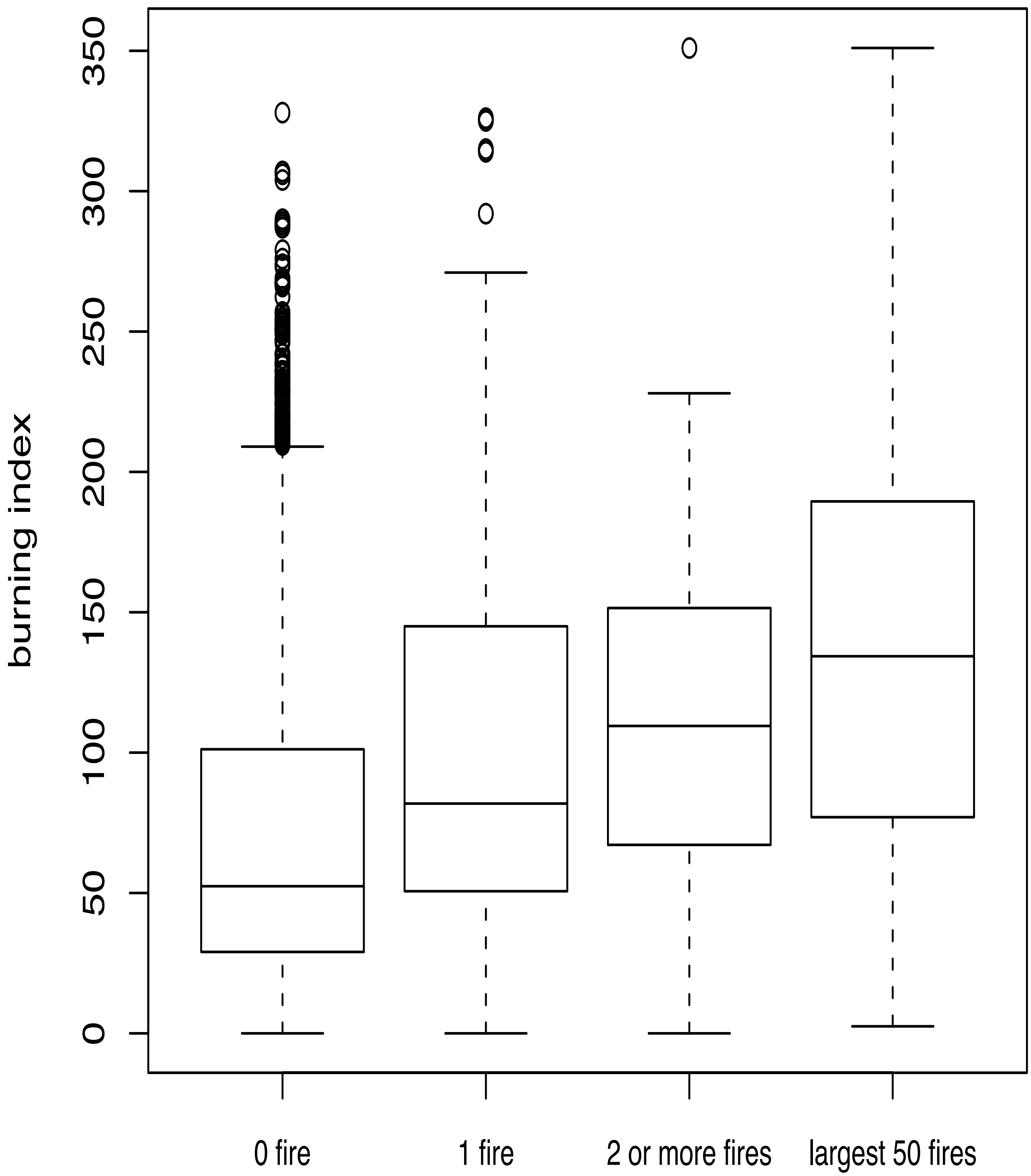
Figure 4: Area burned versus BI for different classes of BI, for fires occurring between October through January, from 1976-2000.

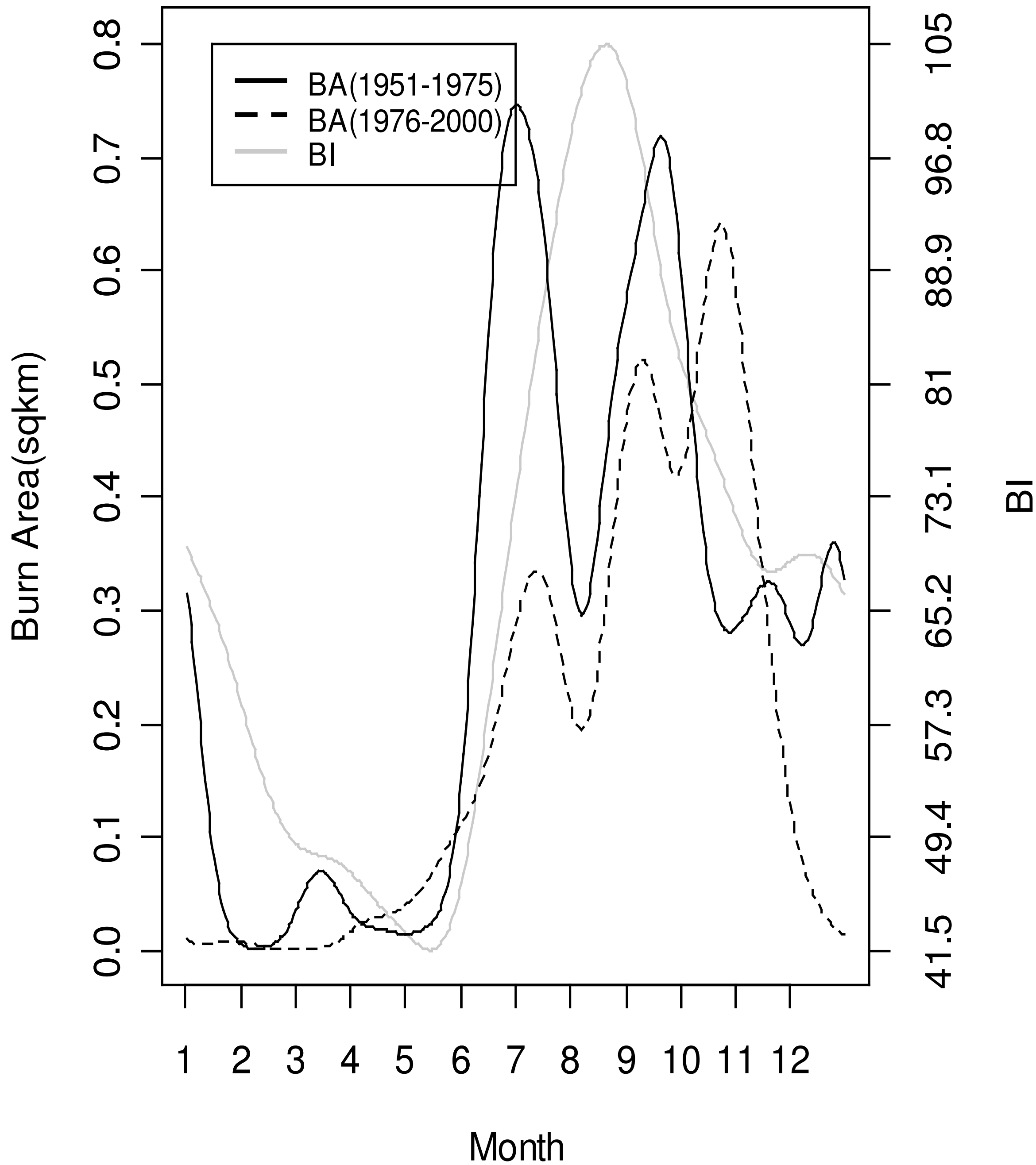
Figure 5: Average daily precipitation over the days preceding wildfires. The dark, thin curves corresponds to days when the BI was in the range 100-150 and at least one wildfire occurred; for such days, the dark, thin, solid curve indicates the mean daily precipitation over the preceding days, and the dark, thin dashed curves represent the corresponding 95%-confidence bounds. As described in Section 3, these confidence bounds are constructed by considering all days with BI in the range 100-150: for each such day, the mean daily precipitation over the preceding days is calculated, and the sample mean $\pm$ 1.96 times the

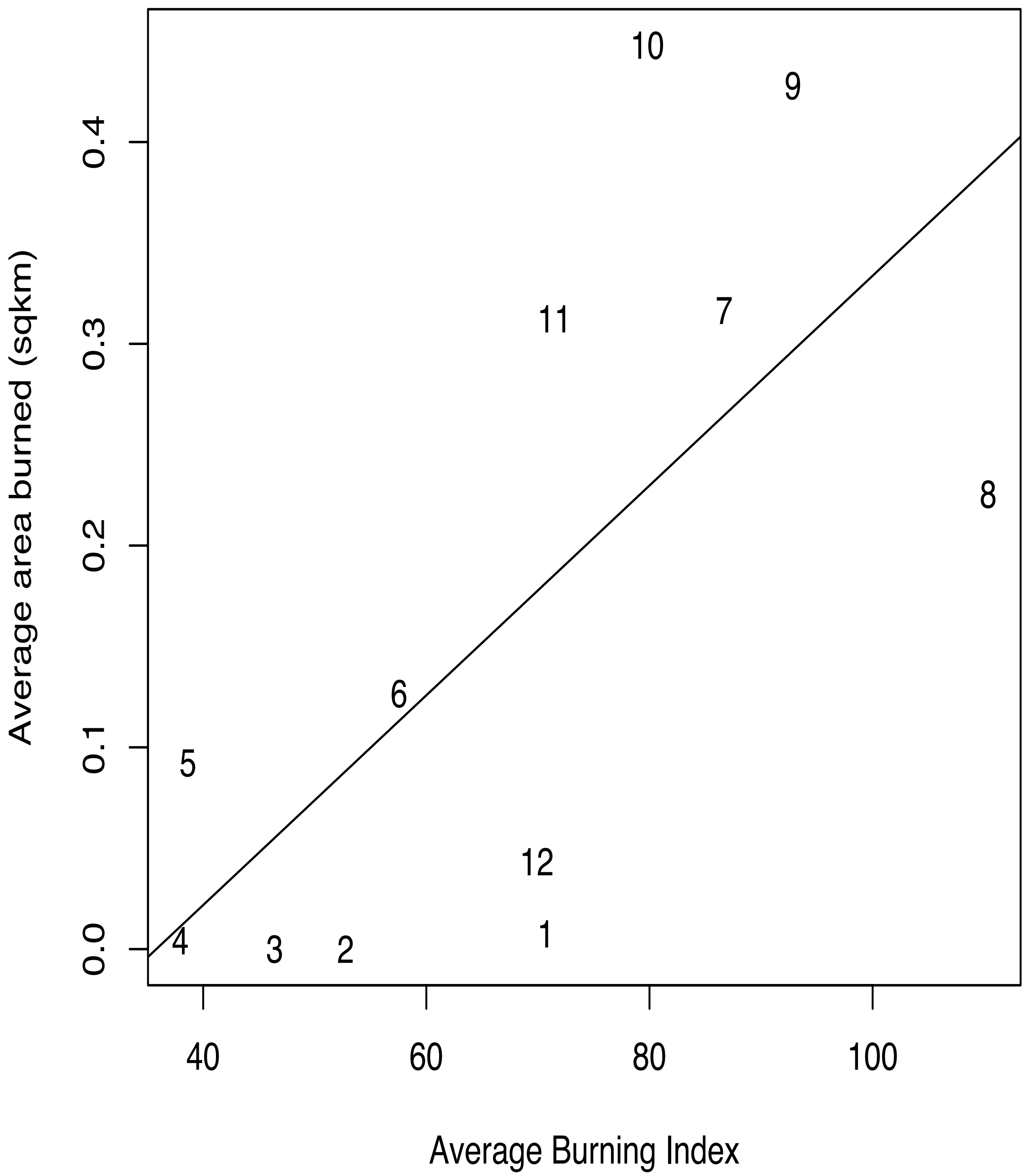
sample standard deviation is used to construct the confidence bounds. The light, thick solid curve corresponds to days when the BI was in the range 150-200 and at least one wildfire occurred: for such days, this curve indicates the mean daily precipitation over preceding days. The light, thick dashed curves are the 95%-confidence bounds corresponding to the light, thick solid curve.

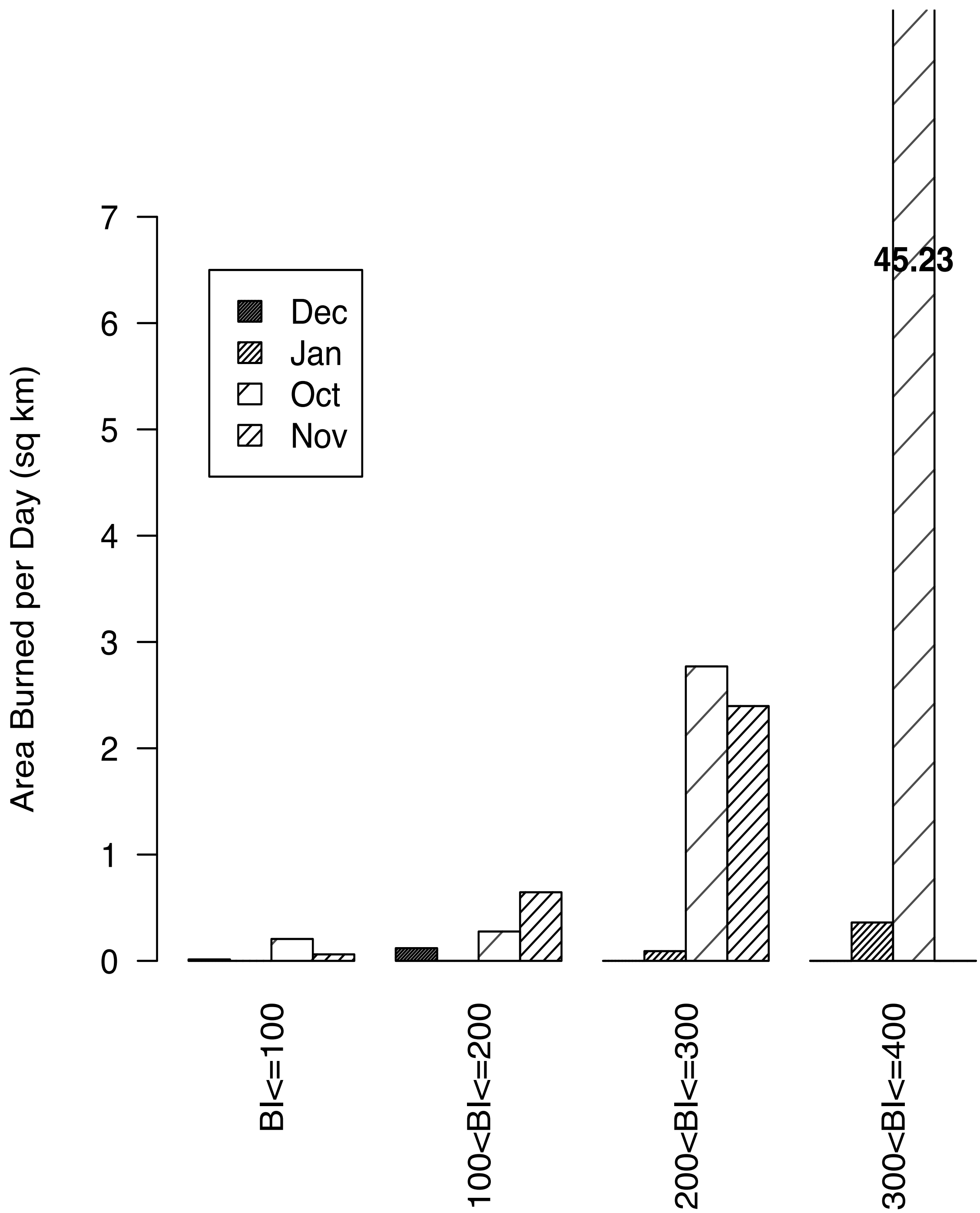
Figure 6: Bivariate kernel smoothing (top) and product of univariate kernel smoothings (bottom) for temperature and maximum relative humidity. The shading of each pixel indicates the smoothed ratio of the number of wildfires occurring on days of the corresponding temperature and maximum relative humidity to the total number of such days. The result is a smoothed rate of occurrence of wildfires, in number of fires per day. Points on the plot indicate wildfires. In order to restrict attention only to rates that may be stably estimated, pixels corresponding to values of relative humidity and temperature are omitted if the denominator in the ratio described is in the bottom 10th percentile.

Figure 7: Success rate (percent of fires occurring on days of warnings) versus false alarm rate (number of alarms per year), for the BI model (6) and Weather model (8). Dashed curve = Weather model (8); solid curve = BI model (6).

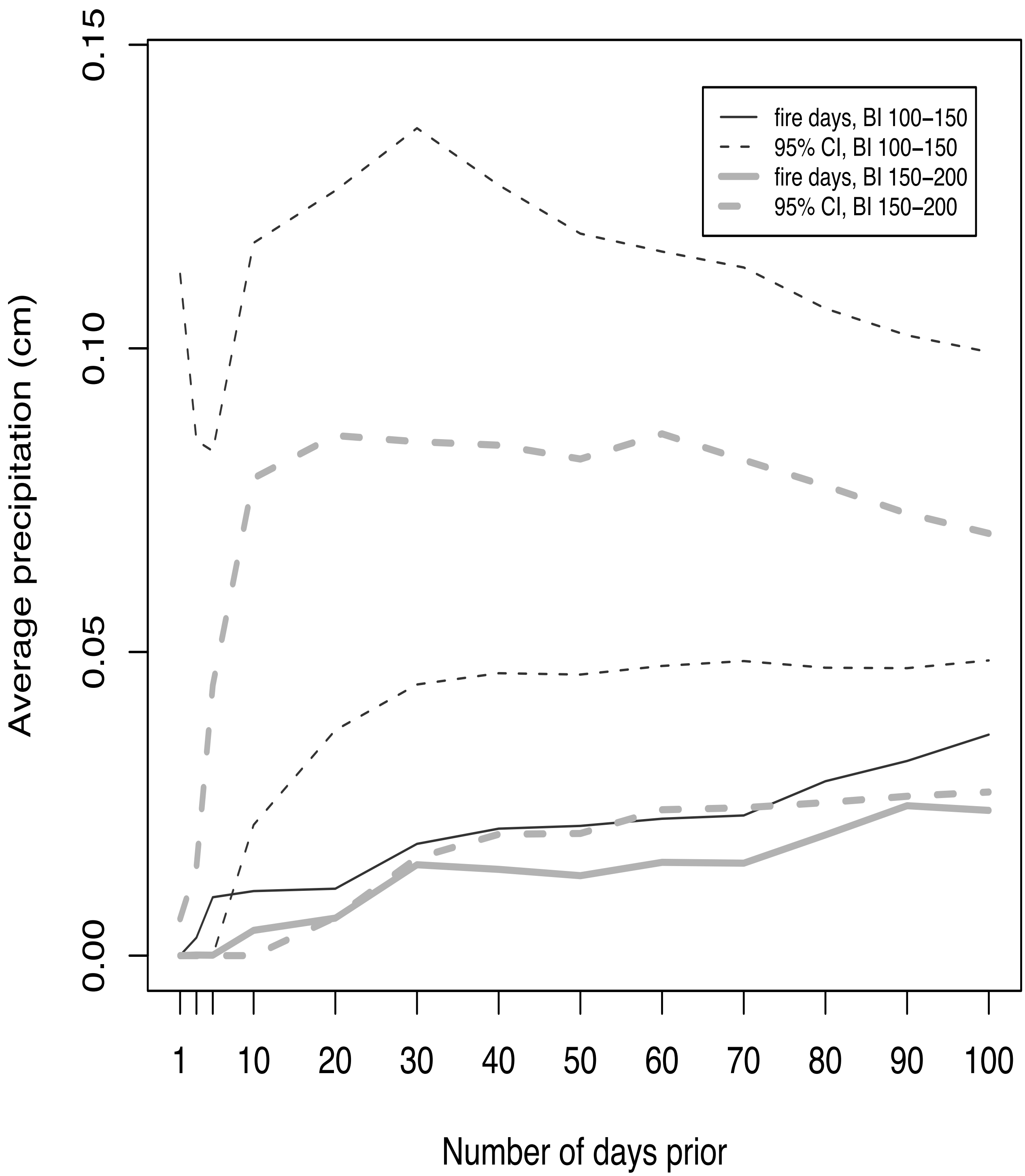

burning index
0
50
100
150
200
250
300
350
0 fire
1 fire
2 or more fires
largest 50 fires

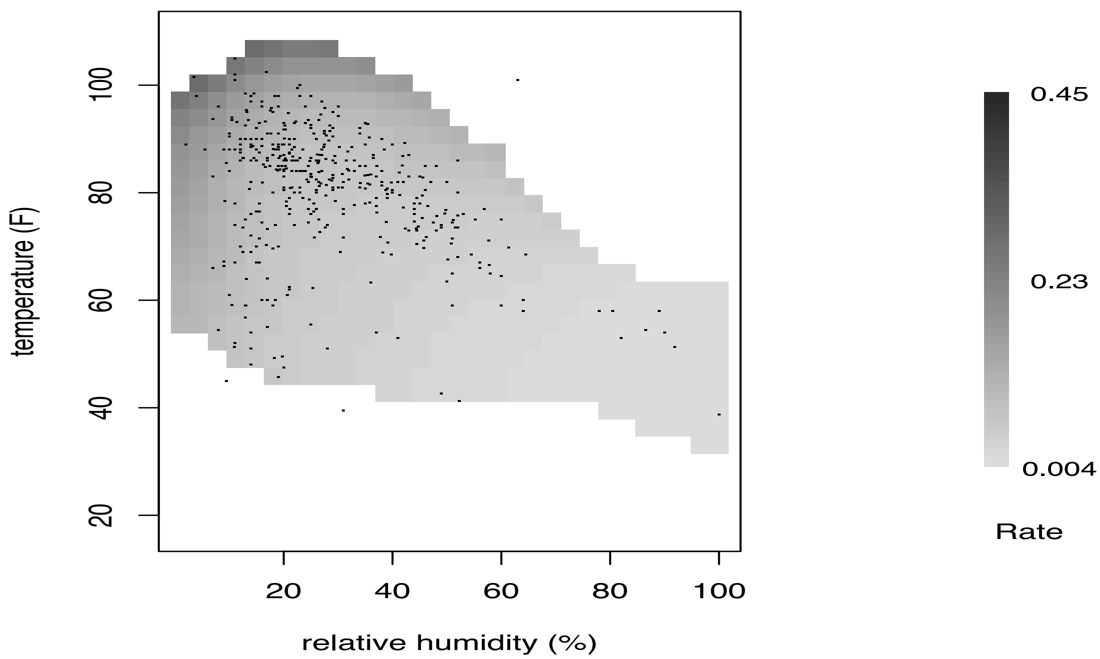

BA(1951-1975)
BA(1976-2000)
BI
Burn Area(sqkm)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
BI
41.5
49.4
57.3
65.2
73.1
81
88.9
96.8
105
1 2 3 4 5 6 7 8 9 10 11 12
Month


Average area burned (sqkm)
0.0
0.1
0.2
0.3
0.4
40
60
80
100
Average Burning Index
1
2
3
4
5
6
7
8
9
10
11
12

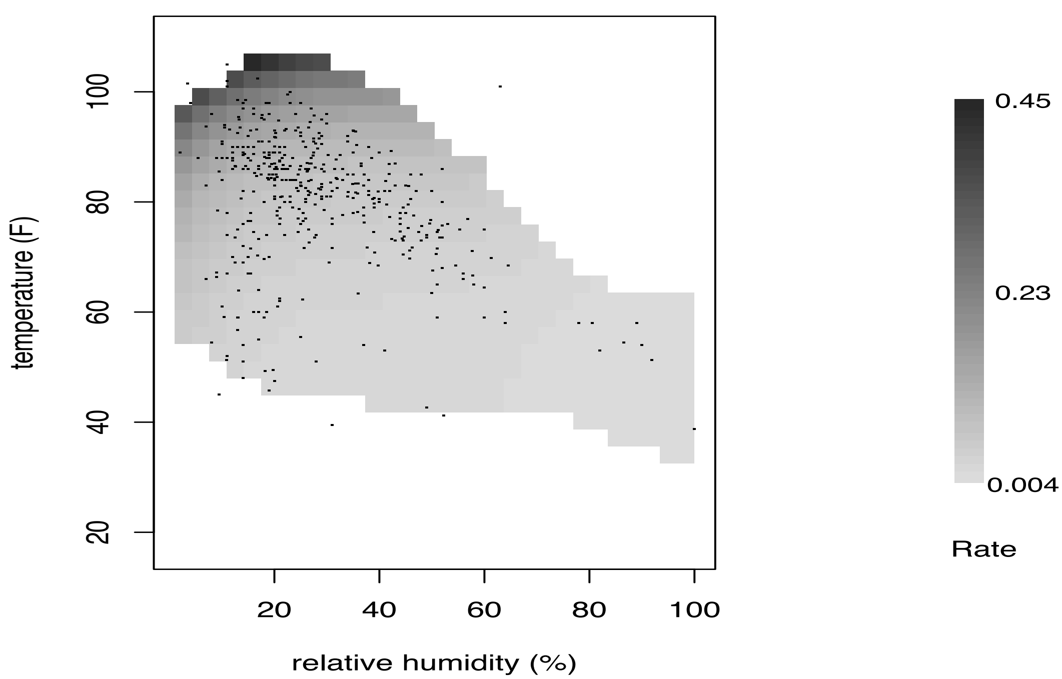


Area Burned per Day (sq km)
0
1
2
3
4
5
6
7
Dec
Jan
Oct
Nov
45.23
BI<=100
100<BI<=200
200<BI<=300
300<BI<=400

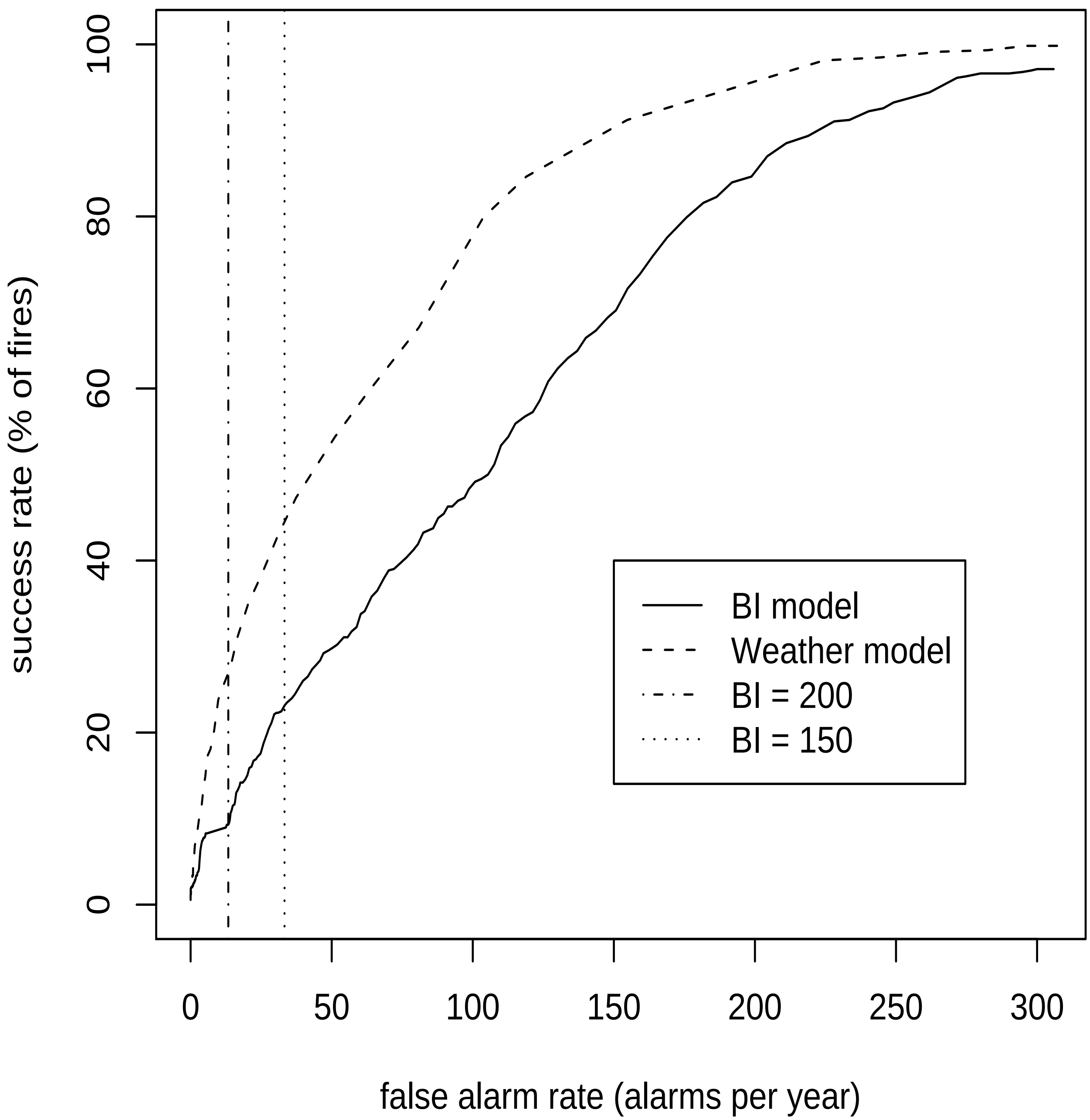

fire days, BI 100–150
95% CI, BI 100–150
fire days, BI 150–200
95% CI, BI 150–200
Average precipitation (cm)
0.00
0.05
0.10
0.15
1
10
20
30
40
50
60
70
80
90
100
Number of days prior



temperature (F)
20
40
60
80
100
20
40
60
80
100
relative humidity (%)
0.45
0.23
0.004
Rate



temperature (F)
20
40
60
80
100
20
40
60
80
100
relative humidity (%)
0.45
0.23
0.004
Rate


success rate (% of fires)
0
20
40
60
80
100
0
50
100
150
200
250
300
false alarm rate (alarms per year)
BI model
Weather model
BI = 200
BI = 150