CSIRO Publishing

International Journal of Wildland Fire

Scientific Journal of IAWF

VOLUME 11, 2002

Editor-in-Chief
Dr M Flannigan, Canadian Forest Service, Edmonton, Canada.

IJWF is published for the International Association of Wildland Fire by:

International Association of Wildland Fire

CSIRO PUBLISHING
PO Box 1139 (150 Oxford Street)
Collingwood, Victoria 3066
Australia

Telephone: +61 3 9662 7644 (editorial enquiries)
+61 3 9662 7668 (subscription enquiries and claims)

Fax: +61 3 9662 7611 (editorial enquiries)
+61 3 9662 7555 (subscription enquiries and claims)

Email: publishing.ijwf@csiro.au (editorial enquiries)
publishing.sales@csiro.au (subscription enquiries and claims)

CSIRO PUBLISHING

Please submit all new manuscripts directly to **CSIRO** PUBLISHING, in electronic form only. See the *IJWF* Notice to Authors for more information.

www.publish.csiro.au/journals/ijwf

International Journal of Wildland Fire, 2002, **11**, 257–266

Long lead statistical forecasts of area burned in western U.S. wildfires by ecosystem province

Anthony L. Westerling[A,C], *Alexander Gershunov*[A], *Daniel R. Cayan*[A,B] *and Tim P. Barnett*[A]

[A]Climate Research Division, Scripps Institution of Oceanography, University of California, Mail Code 0224, 9500 Gilman Drive, La Jolla, CA 92093–0224, USA
[B]US Geological Survey, 9500 Gilman Drive, La Jolla, CA 92093–0224, USA.
[C] Corresponding author. Telephone: +1 858 822 4057; fax: +1 858 534 8561; email: leroy@ucsd.edu

This paper is derived from a presentation at the 4th Fire and Forest Meteorology Conference, Reno, NV, USA, held 13–15 November 2001

Abstract. A statistical forecast methodology exploits large-scale patterns in monthly U.S. Climatological Division Palmer Drought Severity Index (PDSI) values over a wide region and several seasons to predict area burned in western U.S. wildfires by ecosystem province a season in advance. The forecast model, which is based on canonical correlations, indicates that a few characteristic patterns determine predicted wildfire season area burned. Strong negative associations between anomalous soil moisture (inferred from PDSI) immediately prior to the fire season and area burned dominate in most higher elevation forested provinces, while strong positive associations between anomalous soil moisture a year prior to the fire season and area burned dominate in desert and shrub and grassland provinces. In much of the western U.S., above- and below-normal fire season forecasts were successful 57% of the time or better, as compared with a 33% skill for a random guess, and with a low probability of being surprised by a fire season at the opposite extreme of that forecast.

Additional keywords: climatology.

Introduction

Strong associations between observed climate anomalies at lags of one season to years in advance and normalized area burned in the western U.S. wildfire season have recently been described using a newly compiled comprehensive gridded western regional fire history (Westerling *et al.* 2001). Earlier studies of fire scar chronologies and local fire histories have demonstrated that large-scale climate patterns are linked to the severity of the wildfire season in various regions of the U.S. at similar lead times (Simard *et al.*1985; Swetnam and Betancourt 1990, 1998; Balling *et al.*1992; Jones *et al.* 1999; Veblen *et al.* 2000; Donnegan *et al.* 2001). These relationships and the availability of a comprehensive western wildfire history make possible this experimental statistical methodology forecasting area burned in western wildfires one season in advance.

Previous work (Westerling *et al.* 2001, 2003) has established that lags of the Palmer Drought Severity Index (PDSI) (Alley 1984) can be used to forecast area burned at lead times of a season to years in advance, using local PDSI as explanatory variables. While the statistical linkage does not account for effects of short term weather such as hot, dry winds and lightning, nor various human impacts such as fire suppression, land use, and arson, it does produce measurable prediction skill above that expected by chance by tapping into climate factors that affect fuel availability and flammability. Many authors have hypothesized that fuel availability is promoted by anomalously wet antecedent years through fuel production and carryover, while fuel flammability is enhanced in a dry year (Swetnam and Betancourt 1990, 1998; Veblen *et al.* 1999, 2000; Donnegan *et al.* 2001; Westerling *et al.* 2003).

Regional indices describing patterns of variability in PDSI—represented by leading principal components (PCs) of antecedent PDSI values—show similar skill in forecasting western wildfire season severity. Moreover, models based on these regional indices yield some predictive skill even in locations where strong associations between local PDSI values and area burned are lacking. For example, normalized area burned in the Northern Rockies shows little correlation with local PDSI at lead times early enough to be useful for forecasting the fire season, but may be forecast (correlations

 10.1071/WF02009 1049-8001/02/030257

> 0.5) using PCs of antecedent seasons' PDSI from 110 western climate divisions (Westerling 2001, 2003). Western climate divisions—often quite large and covering a diverse topography—sometimes rely on only a few weather stations for precipitation and temperature data, which may be unrepresentative of conditions pertinent to wildfires reported in the neighborhood of a particular division. Regional climate indices, however, record climate signals associated with atmospheric circulation patterns spanning a broad area, and may as a result be useful for generating area burned forecasts where local climate division information is inadequate.

Canonical Correlation Analysis (CCA) (Johnson and Wichern 1998) offers a method for constructing wildfire season severity models whose prediction skill derives from spatial and temporal patterns in climate spanning a broad region and several seasons. Here we estimate a forecast model for the entire western U.S. using CCA to calculate linear relationships between principal components of seasonal area burned aggregated by ecosystem province (Bailey *et al.* 1994) and principal components of U.S. Climatological Division PDSI values from several seasons (similar to the forecast methodology in Gershunov *et al.* 2000). In other words, the model is the sum of relationships between two sets of patterns—one describing PDSI, the other describing area burned. The result is an improvement in forecast skill over an earlier, gridded forecast using principal components regression that we report elsewhere (Westerling 2001). We use 'leave-one-out' cross-validation (von Storch and Zwiers 1999) to estimate robust measures of forecast skill over a range of choices for model complexity, and a Skill Optimization Surface (SOS) to select a parsimonious model maximizing forecast skill over the entire region.

Data

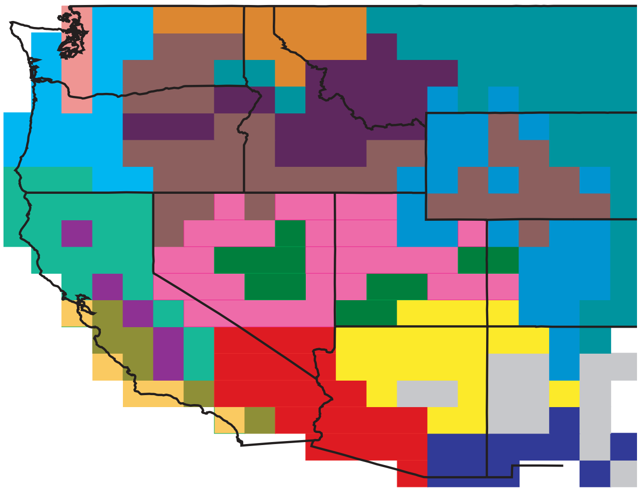
The fire history used here is composed of seasonal acres burned on a 1º × 1º grid extending from 31°N to 49°N latitude and from 101°W to 125°W longitude for 1980 through 2000 (Westerling *et al.* 2003). These data were compiled from 300 000 quality-controlled fire reports of the Bureau of Land Management (BLM), U.S. Forest Service (USFS), National Park Service (NPS) and Bureau of Indian Affairs (BIA). Data were sparse in the grid cells from 101°W to 103°W longitude, and the ecosystem provinces covering these longitudes are not characteristic of most of the rest of the western U.S., so data east of 103°W longitude were excluded from this analysis.

While some skill in forecasting area burned on a 1º × 1º grid has been previously demonstrated (Westerling *et al.* 2001), aggregating area burned over larger regions may help to reduce some of the temporal and spatial noise apparent in the gridded data. Simply put, the smaller the area of aggregation, the less the likelihood that there will be significant fire activity occurring within that area for any given fire season. The result of aggregation over too small an area is a time series with a couple of years of very large areas burned, and the rest close to zero. Given that only 21 years of data are available, it is difficult to detect a long-lead climate signal in such a time series against the background noise of management actions, errors in the data, and random variations in human- and lightning-caused ignitions. Aggregating over a larger area can average out some of these background effects and produce a time series more reflective of fuel conditions on the one hand, and more amenable to statistical analysis on the other. In order not to obscure any climate signal, however, it is important that any aggregation scheme group areas with wildfire regimes that respond to climate in similar ways.

An ecosystem classification scheme described by Bailey *et al.* (1994) and adopted by the USFS and others as a framework for ecosystem analysis and management offers a useful level of aggregation for this analysis. Bailey *et al.*'s ecosystem provinces are classified by coarsely generalized characteristics of climate, vegetation and elevation—all of which appear to be important for forecasting fire season severity (McKelvey and Busse 1996; Westerling *et al.* 2001). This classification scheme has the added benefit of being an accepted and familiar tool of many potential users of any operational long lead wildfire season severity forecast for the western U.S.

In many instances the location data for fires described in the original federal agency fire reports were sufficient only to locate these fires on a 1º resolution, and do not allow for finer resolutions. Application of Bailey *et al.*'s (1994) ecosystem provinces as a classification scheme for western wildfires was consequently constrained to the coarse 1º × 1º grid. Figure 1 shows the 17 ecosystem provinces used in this analysis projected onto the grid. Area burned was aggregated for a 6-month fire season (May through October) for 17 ecosystem provinces, and a $\log_{10}$ transformation was then used to normalize the data. The result was 17 time series, each of 21-years, of $\log_{10}$ area burned spanning most of the western US.

For explanatory variables, we use 110 western U.S. Climate Division PDSI series (Karl and Knight 1985). While alternative drought indices are available, the PDSI is an imperfect but readily available proxy for soil and fuel moisture broadly familiar to western wildfire managers, and produces long lead fire season severity forecasts with skill significantly better than that afforded by chance. The PDSI predictors were chosen to represent five different lead times: March and December immediately preceding the fire season, August and March 1 year previous to, and August 2 years prior to the fire season, for a total of 550 predictor variables (cf. Westerling *et al.* 2001). March and August PDSI in the preceding years were selected as indicators of moisture available in spring and summer for fuel growth. Excess



242 Pacific Lowland Mixed Forest Province
261 California Coastal Chaparral Forest and Shrub Province
262 California Dry Steppe Province
313 Colorado Plateau Semidesert Province
321 Chihuahuan Semidesert Province
322 American Semidesert and Desert Province
331 Great Plains-Palouse Dry Steppe Province
341 Intermountain Semidesert and Desert Province
342 Intermountain Semidesert Province
M242 Cascade Mixed Forest--Coniferous Forest--Alpine Meadow Province
M261 Sierran Steppe--Mixed Forest--Coniferous Forest--Alpine Meadow Province
M262 California Coastal Range Open Woodland-Shrub-Coniferous Forest--Meadow Province
M313 Arizona-New Mexico Mountains Semidesert-Open Woodland--Coniferous Forest--Alpine Meadow Province
M331 Southern Rocky Mountain Steppe--OpenWoodland--Coniferous Forest--Alpine Meadow Province
M332 Middle Rocky Mountain Steppe--Coniferous Forest--Alpine Meadow Province
M333 Northern Rocky Mountain Forest-Steppe--Coniferous Forest--Alpine Meadow Province
M341 Nevada-Utah Mountains Semidesert--Coniferous Forest--Alpine Meadow Province

Fig. 1. Bailey's ecosystem provinces projected onto a $1^\circ \times 1^\circ$ grid (Bailey *et al.* 1994).

moisture in these seasons may also contribute to fuel accumulation by suppressing antecedent years' area burned. December and March PDSI immediately prior to the fire season were selected as indicators of fuel moisture conditions leading into the fire season.

Since the divisional PDSI is an estimate of drought severity using a weighted sum of past and present monthly precipitation and temperature from a given climate division (Alley 1984, 1985; Guttman 1991), consecutive months of PDSI are usually highly correlated. Including PDSI values from additional spring and summer months would thus be unlikely to add much information. By limiting the model to data available up through March of the current year, it is possible to produce a forecast before budgeting and resource allocation decisions are taken for the western U.S. fire season.

Methods

Space-time varying patterns: *principal components*

To further reduce noise over what is accomplished by analysing wildfire by ecoregions, to avoid colinearity among both data sets, and to substantially diminish the number of variables included in the CCA, the predictor (PDSI) and predictand (area burned) data sets were formed from principal components (PCs) for each of the two data sets (Johnson and Wichern 1998). Intuitively, each PC describes the temporal evolution of a specific spatial pattern. Algebraically, the first PC of each data set is the linear combination of all the variables in the data set with maximum variance. Each subsequent PC is likewise a linear combination of the original variables with its variance maximized subject to the constraint that the PC is linearly independent of each other PC. Thus, each PC summarizes an independent mode of variability in its original data set, and taken together all the PCs summarize all the information contained in the original data set. For the predictor data, the first six principal components explain over 70% of total variance. Similarly for the predictand, the first six principal components explain more than 85% of total variance. So, relatively few PCs are needed to convey most of the information contained in these data.

Canonical correlation analysis

We use CCA to identify patterns in each of the two PC data sets that are optimally correlated with each other. A linear combination of the predictor PCs and a linear combination of the predictand PCs are calculated such that the correlation between the two is maximized. Each subsequent pair is similarly calculated to maximize their correlation subject to the constraint that they be uncorrelated with the other pairs. Since these canonical correlates (CCs) are linear combinations of PCs, which in turn are linear combinations of the original data, we can specify a set of CCs to be our linear forecast model and solve for standardized $\log_{10}$ acres burned, our predictand, given the appropriate lagged PDSI values. For a detailed CCA methodology see Barnett and Preisendorfer (1987), Johnson and Wichern (1998), and Gershunov *et al.* (2000). Intuitively and in fact, the CCs are a set of linear relationships between two sets of patterns describing the predictor and predictand fields.

Skill optimization surface

To select a parsimonious CCA model using the specified lags of PDSI, we calculated Skill Optimization Surfaces (SOSs) like that shown in Fig. 2. The x-axis denotes the number of PCs contributed from the predictor and from the predictand data. The number of each are constrained to be equal here to render the solution more tractable. The y-axis denotes the number of CCs included in the model. Note the triangular shape of the shaded area—the maximum number of CCs is limited to the number of PCs included. Thus, in the lower left corner we use only the first principal component, which explains the largest share of variance, from both the predictors and predictands, and as a result are constrained to estimating our model from only the first CC pair. As we move to the right, adding PCs in order of their share of the total variance of their data sets explained, we can choose to move up the y-axis, adding additional CC pairs in our model in order of decreasing strength of correlation.

The shading in the SOS denotes the overall skill of the forecast, calculated here as the sum of squared positive correlations between the cross-validated series of model estimates and the observed area burned for all 17 provinces (Fig. 2). Negative correlations indicate no skill and were excluded. This metric gives more weight to models with a small percentage of provinces with high forecast skill than the other metrics used for this analysis—the sum of correlations, including negative correlations (SCOR) and the percentage of provinces with significant correlations between forecasts and observations (PSIG). However, the model dimensions showing the greatest skill in this example—three canonical correlations composed of two sets of six principal components—were also ranked first for SCOR, while PSIG tended to rank nearby points first. This is approximately the model complexity needed to capture the effect of relevant climate forcing patterns on coherent fire patterns.* The SOS provides an alternative to arbitrarily selecting model complexity.

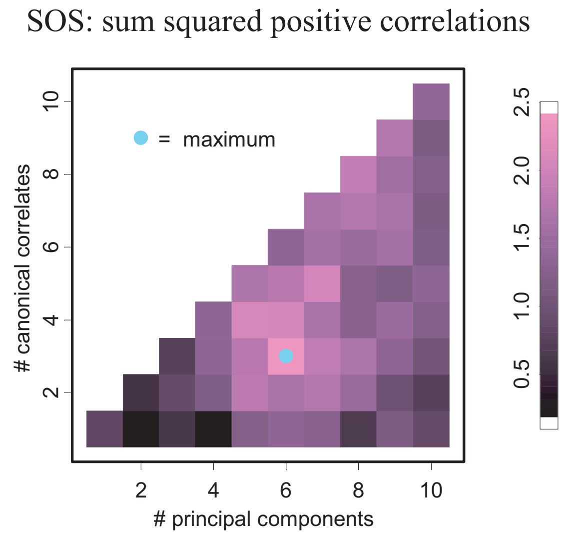


Fig. 2. CCA model skill optimization surface. Whole field metric: sum of squared positive correlations.

Model skill and cross-validation

Conditional and probabilistic forecasts may better reflect the uncertainty in prediction skill than simple correlations. This is achieved by 'tiling' of the data—i.e. expressing it in terms of categories as opposed to specific values—and presenting probabilistic outcomes contingent on forecasts for each category. In this case, after selecting the forecast model dimensions using an SOS, each province's observed and forecast time series were sorted into terciles. Thus, for example, each year's burn acreage for each region was assigned to be in one of three classes: Above normal (A), Below normal (B) or Normal (N), the occurrence of each class being equally likely (7 out of 21 years, or 33%). Because forecast model dimensions are selected as those

*Model complexity can also be optimized for geographic regions of interest (i.e. regional skill can always be improved at the expense of total predictand field skill).

yielding the highest skill from among the large number of models calculated in the SOS, even a forecast model derived using random predictors may show better skill than a random guess (33%). Consequently, the confidence intervals used in the discussion of results below were estimated for each class using the cross-validated forecast methodology described here, but with 21-year samples of consecutive PDSI values drawn from 1895–1976 as predictors. This allows us to compare the forecast skill we observe to the skill we would expect from using as predictors data which are apparently unrelated to area burned, but which have the same characteristics as observed PDSI.

To avoid an artificially inflated estimate of the skill achieved in this exercise, model diagnostics here are all for results using 'leave-one-out' cross-validation. That is, for each time step (year) of the model, a forecast is constructed using model coefficients estimated on the subset of the data excluding that time step. This reduces the potential for false statistical skill in the diagnostic measures reported below using Pearson's correlations and conditional probabilities.* In the context of our CCA models, 'leave-one-out' cross-validation requires not only that at each time-step the coefficients of our forecast model be estimated on the subset of the data excluding that time step, but also that the loadings on the principal components and canonical correlations be recalculated at each time-step on that subset as well. Thus, for each time-step and number of PCs and CCs, we are calculating a different model.

The cross-validation procedure also generates a separate SOS for each time step. The SOS does not help us to select an exact, fixed model, but rather the approximate dimension or model complexity that gives the best result. In this example, SOS plots showed a maximum in skill at 6 PCs and 3 CCs for each set of cross-validated models, so we further examine the properties of a cross-validated series of models using these dimensions in the remainder of the paper. When we refer to the forecast model hereafter, we are actually referring to 21 separate forecast models, one for each year, each using 6 PCs and 3 CCs estimated withholding information about that year's fires. Since the model estimating each year's fire activity is trained on the remaining years' observations, the resulting skill is representative of the forecast skill to be expected in an operational application of our forecasting technique.

Results and Discussion

Forecast skill

Forecast skill, expressed as the correlation between cross-validated model estimates and observed normalized area burned, is generally highest in the Sierra Nevada and the Intermountain West (Fig. 3). The model appears to perform best in the Northern and Southern Rocky Mountain Forest-Steppe Provinces, the Cascade Mixed Forest Provinces, the Sierran Steppe and California Dry Steppe Provinces, the Intermountain Semidesert and Desert Province, the American Semidesert and Desert Province, the Arizona–New Mexico Mountains Semidesert–Open Woodland–Coniferous Forest–Alpine Meadow Province, and the Nevada–Utah Mountains Semidesert–Coniferous Forest–Alpine Meadow Province, where cross-validated correlations between observed and forecast area burned are greater than 0.5.

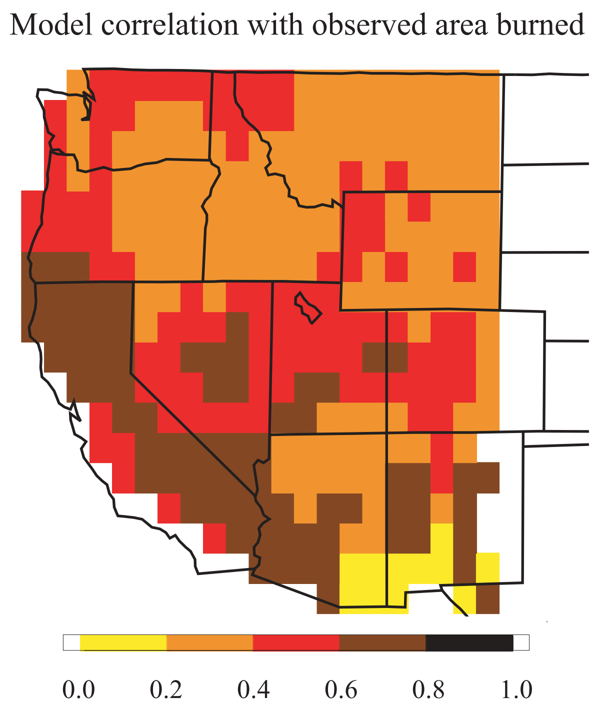


Fig. 3. Correlation between CCA model and observed area burned by ecosystem province.

The probability of a correct 'A' forecast—i.e. that above-normal area burned is observed given an above-normal forecast—is 71% (or 5 out of 7) in the Sierra Nevada and Great Basin, and 57% (4 out of 7) in the Northern Rockies, Cascades, California Central Valley and parts of New Mexico and Arizona (Fig. 4*A*). Using a bootstrap, models were found to be significant at the 90% confidence level if they were associated with correct A or B forecasts 71%, or more, of the time (corresponding to 5 or more correct forecasts out of 7 in a tercile). A 70% confidence level pertained to models associated with 4 or more correct out of 7 A or B forecasts.

*Because the model specification includes explanatory variables from several years, we considered whether autocorrelation in the area burned indices could result in artificially high skill even with 'leave-one-out' cross-validation. We found that the effects of leaving out an additional 2 years before or after the year being forecast were indistinguishable from the effects of reducing our short sample size (21 years) by an additional 2 years. While a more extensive cross-validation–leaving out 3 or 5 years–might be considered more complete, it is impracticable given the short length of the fire history currently available.

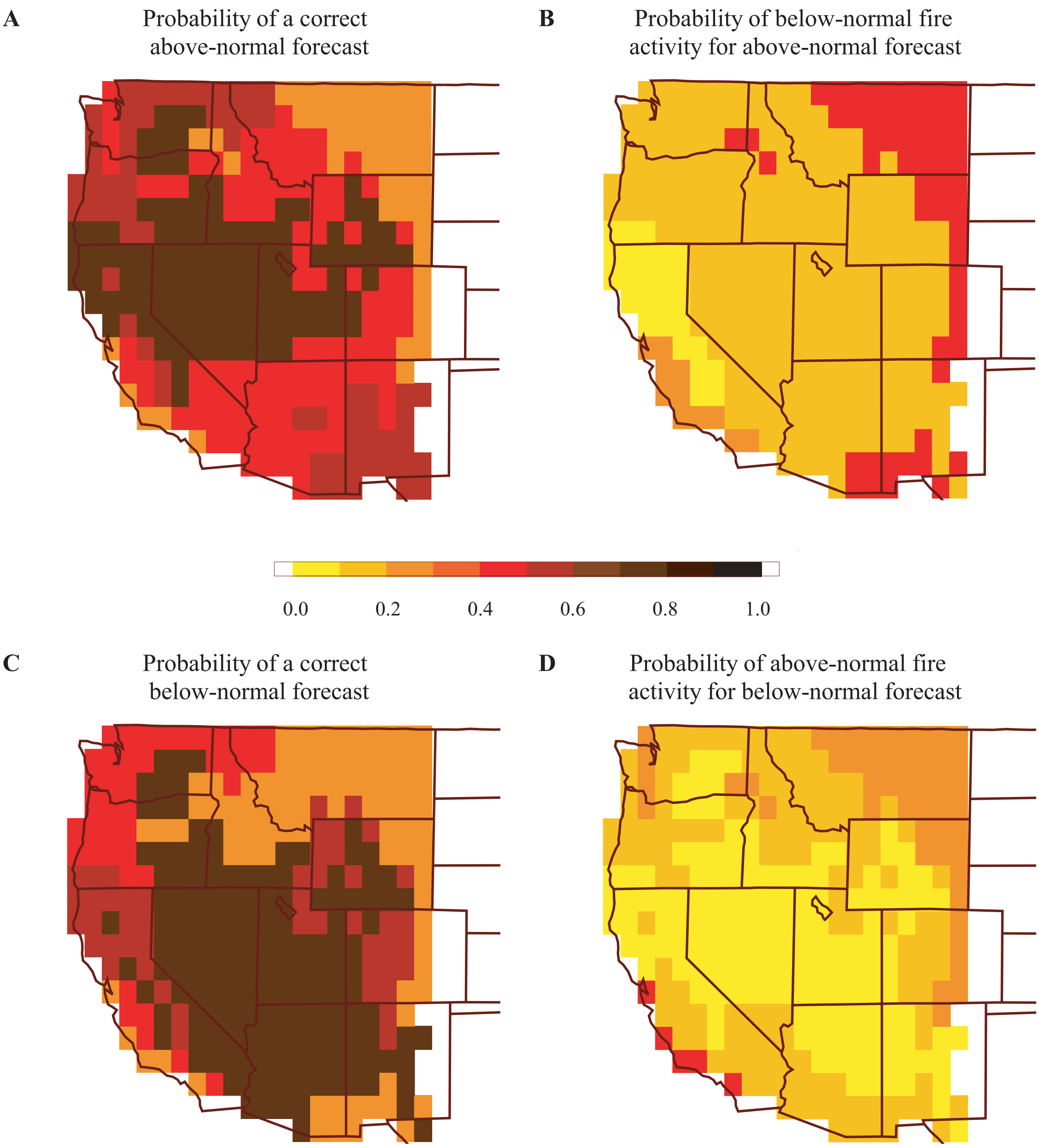


Fig. 4. Probability of (*A*) above-normal fire season given an above-normal forecast; (*B*) below-normal fire season given an above-normal forecast; (*C*) below-normal fire season given a below-normal forecast; and (*D*) above-normal fire season given a below-normal forecast. Probabilities are all for cross-validated forecasts.

The probability of a surprise—i.e. the probability of observing below-normal area burned given an above-normal forecast—is low (14%, or 1 out of 7) for much of this region (Fig. 4*B*). Figures 4*C* and 4*D* show the corresponding probabilities for a 'B' (below-normal) forecast. Note that in many provinces these probabilities are not symmetric. 'A' forecasts in the Northern Rockies, for example, are more skillful than 'B' forecasts. The probability of correctly forecasting a 'normal' fire season is not shown, but in general these models show the greatest skill in forecasting above- or below-normal seasons, and little skill when forecasting a normal fire season. The Sierra Nevada and Great Basin show the greatest forecast skill overall.

Forecast interpretation

The weights placed by the CCA model on the various lags of divisional PDSI for ecosystem provinces in the Great Basin and the Sierra Nevada reveal mechanisms that enable meaningful seasonal forecasts of area burned (Fig. 5). The weights shown in Fig. 5 for the Intermountain Semidesert and Desert Province are typical for all the Great Basin provinces and for coastal Southern California provinces as

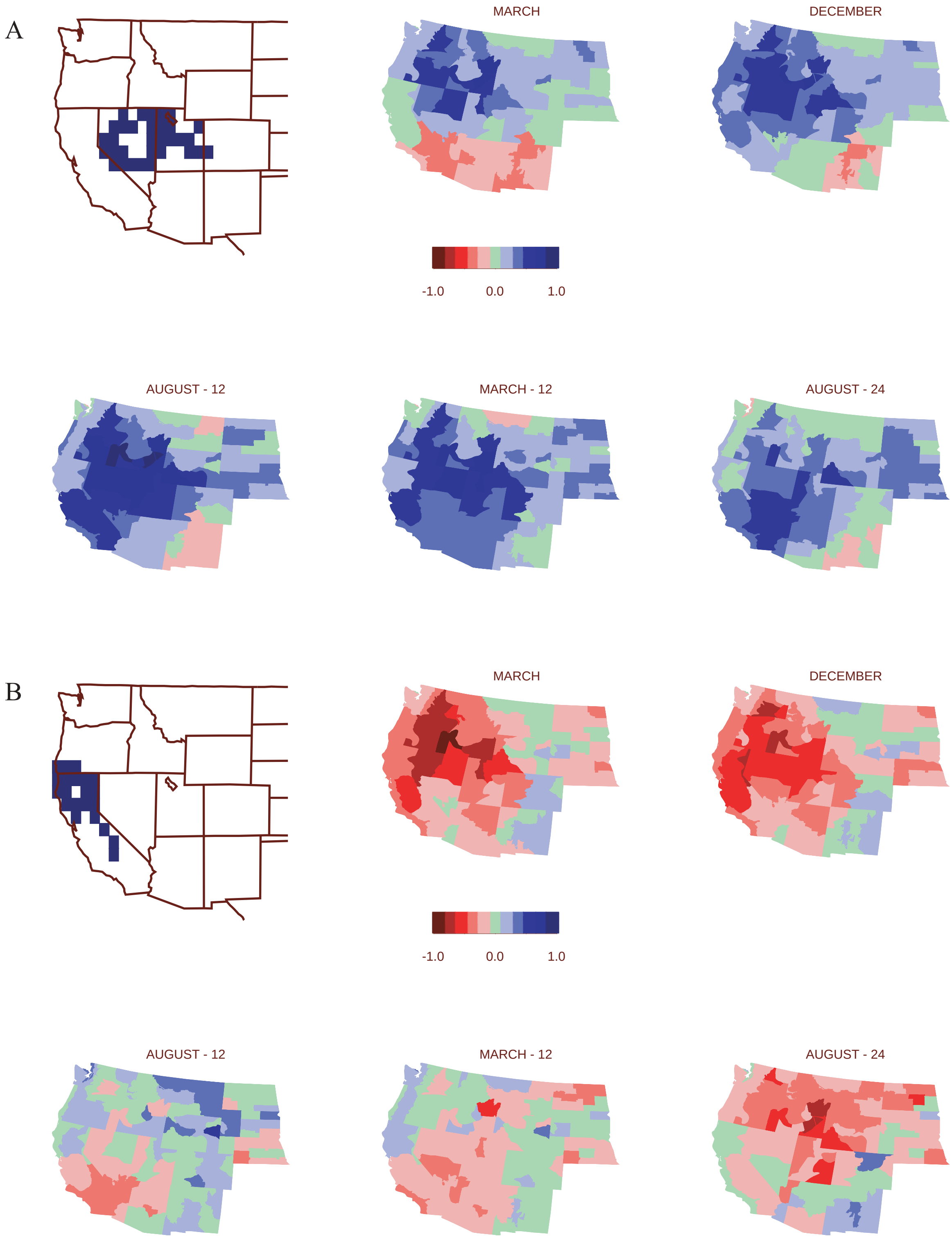


Fig. 5. (*A*) CCA model loadings on lags of PDSI for Intermountain Semidesert and Desert Province (341), and (*B*) CCA model loadings on lags of PDSI for Sierran Steppe Province (M261). Positive PDSI indicates excess moisture anomalies, and negative PDSI indicates deficit moisture anomalies, so positive loadings on lagged PDSI (shown in blue) indicate a positive association between area burned anomalies and moisture anomalies. Negative loadings on lagged PDSI (in red) indicate a negative association between area burned anomalies and moisture anomalies.

well. They show that area burned in the Great Basin is associated with positive soil moisture index anomalies over the region, with the strongest weights on positive soil moisture index values in summer a year before the fire season. CCA model weights in Fig. 5 for the Sierran Steppe province show a strong association between area burned and anomalous dryness in the winter and spring immediately preceding the fire season, and a weak positive or near-neutral association with positive soil moisture index values over northern California a year before.

The weights that the CCA model places on the various lags of PDSI are very much like the relationships between area burned and PDSI observed in Westerling *et al.* (2003) using the same data as in this analysis. They found a few characteristic patterns repeated in multiple locations across the western U.S. for which they hypothesize links to dominant vegetation types. In particular, in shrub and grassland areas such as the Great Basin dominated by fine fuels, they found strong positive correlations between area burned and anomalous soil moisture index values a year before the fire season, but no significant correlation with current soil moisture index values. They hypothesized that fire season severity in these areas was limited by fuel availability and that positive moisture anomalies a year earlier were associated with fine fuel production and holdover into the following year's fire season (cf. Kipfmueller and Swetnam 2000).

In areas with open-canopy forests interspersed with parcels of shrub and grassland like the Sierra Nevada and Colorado Rockies, they found strong negative correlations with anomalous soil moisture index values during and immediately preceding the fire season, and weaker positive correlations with anomalous soil moisture index values a year before the fire season. Fine fuel production and carryover are probably less important in these heavier-fuel areas than the effect of anomalous moisture on fuel flammability (Agee 1993; Swetnam and Betancourt 1998; Veblen *et al.* 2000; Donnegan *et al.* 2001).

Forecasts for 2000–2002

The area burned forecast for the 2000 wildfire season (Fig. 6*A*) indicated an above normal fire season throughout much of the western U.S., with the areal extent of the above-normal forecast very close to that observed (Fig. 6*B*). The Coast range in central and southern California, the California central valley and the Colorado Plateau saw normal or below-normal conditions instead of the above-normal forecast, and the Mojave area also saw below-normal conditions instead of the normal forecast. Only in the Chihuahuan Semidesert Province in the extreme south-west and the Pacific Lowland Mixed Forest Province in the north-west did the observed area burned exceed the forecast category. The 2000 season forecast model coefficients were calculated using the previous 20-year record, so this forecast represents information available by the end of March 2000. Despite the extraordinary nature of the observed 2000 fire season compared with this recent history, the forecast model very successfully characterizes the season.

A forecast for the 2001 fire season was prepared in December 2001, before 2001 observed area burned data were available (Fig. 6*C*). The model continued to perform well in areas where it had shown skill in the 1980–2000 training period. The 2001 forecast predicted a less severe season on the whole than in 2000, with a normal fire season forecast in much of Southern California, the Intermountain West and the Northwest, and above-normal area burned in the Sierra Nevada and Rocky Mountains. 'Normal' forecasts using our methodology showed less skill than 'A' or 'B' forecasts for the model training period, and 2001 was no exception: all the 2001 'N' forecasts were contradicted by observations for their respective ecosystem provinces (Fig. 6*D*). 'Above-normal' forecasts in the Sierran Steppe and Arizona–New Mexico Mountains Semidesert Provinces showed the highest skill (5 out of 7 correct) in the model training period, and again successfully depicted the 2001 season. 'A' forecasts in the Northern Rocky Mountain Forest Steppe, California Dry Steppe, Chihuahuan Semidesert, and Arizona–New Mexico Mountains Semidesert Provinces showed the next highest skill during the model training period (4 out of 7 correct). In 2001, 'A' forecasts in the first two provinces successfully depicted observations, while in the latter two provinces an 'N' season was observed, similar to the expected success ratio. There were no 'B' forecasts for 2001 for areas that showed significant skill in forecasting below-normal area burned in the training period. In areas where, based on performance for 1980–2000, we might reasonably expect skillfull 'A' forecasts, the 2001 forecast performed well.

A preliminary forecast for the 2002 wildfire season (Fig. 6*E*) was presented in March 2002 at *Fire in the West*: *A Climate/Fuels Assessment*, *Outlook and Research Symposium* hosted by the Climate Assessment for the Southwest in Tucson, Arizona, and distributed via the internet. The 2002 forecast anticipates an above-normal fire season in the California Dry Steppe Province and in a region extending from the Chihuahuan Semidesert Province in New Mexico and eastern Arizona up through the Colorado Plateau and Arizona–New Mexico Mountains Provinces to the Southern Rocky Mountain Steppe Province. Below-normal forecast area burned for the Intermountain Semidesert provinces reflect the fact that these areas—where severe wildfire seasons tend to follow wet years—have been anomalously dry for the past few years. While it is still too early for a formal assessment of the 2002 forecast, it does appear to capture some features of the current wildfire season, where the largest fires on record in Colorado and eastern Arizona and a very active fire season in New Mexico are likely to result in above-normal area burned for the

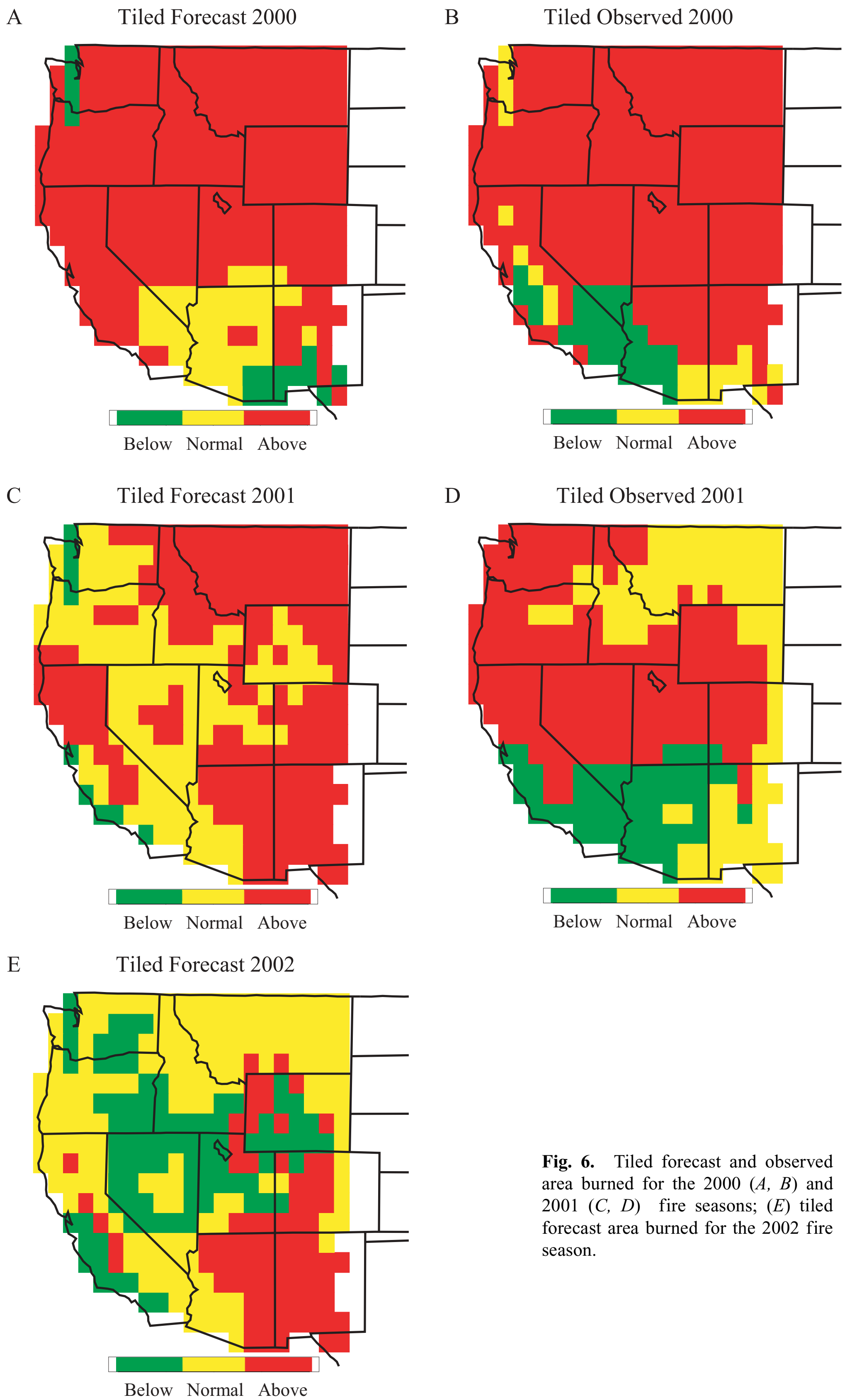


Fig. 6. Tiled forecast and observed area burned for the 2000 (*A, B*) and 2001 (*C, D*) fire seasons; (*E*) tiled forecast area burned for the 2002 fire season.

Chihuahuan Semidesert, Arizona–New Mexico Mountains Semidesert, and the Southern Rocky Mountain Steppe Provinces.

Conclusion

The CCA methodology uses large-scale patterns in soil moisture anomalies over a wide region and several seasons to produce regional forecasts of wildfire seasonal area burned aggregated by ecosystem province a season in advance. While this forecast model is based on a fire history of only 21 years, it produces significant results with stringent cross-validation. In much of the western U.S., upper and lower tercile fire seasons can be forecast with a probability of success of 50% or greater (as compared with 33% by chance alone), and with a low probability of a fire season at the opposite extreme. Operational use of this or a similar forecast methodology would seem to offer benefits for strategic planning for wildfire management, as well as scheduling prescribed fire and other fuel treatment measures.

Acknowledgements

We acknowledge the invaluable assistance received from Bill Mitchell (BLM) and Howard Roos (USFS) at the National Interagency Fire Center, Mike Warren at the National Park Service, Tim McGrath at the Cleveland National Forest, Russ Davis at Scripps Institution of Oceanography, and our reviewers. We are particularly grateful for the unstinting support provided by Sharon Sheppard and Suz Rittenhouse at the Fire Applications Support help desk at the National Interagency Fire Center. Support for this research came from the NOAA Office of Global Programs through the California Applications Program and the Experimental Climate Prediction Center, from the DOE Accelerated Climate Prediction Initiative, and from a National Fire Plan grant to the USDA Forest Service's Southern Research Station.

References

Agee JK (1993) 'Fire ecology of Pacific Northwest forests.' (Island Press: Washington D.C.) 493 pp.

Alley WM (1984) The Palmer Drought Severity Index: limitations and assumptions. *Journal of Climate and Applied Meteorology* **23**, 1100–1109.

Alley WM (1985) The Palmer Drought Severity Index as a measure of hydrologic drought. *Water Resources Bulletin* **21**, 105–114.

Bailey RG, Avers PE, King T, McNab WH (Eds) (1994) 'Ecoregions and subregions of the United States (map)'. 1:7,500,000. [With supplementary table of map unit descriptions, compiled and edited by W. H. McNab and R. G. Bailey.] (USDA Forest Service: Washington, D.C.)

Balling RC, Meyer GA, Wells SG (1992) Relation of surface climate and burned area in Yellowstone National Park. *Agricultural and Forest Meteorology* **60**, 285–293.

Barnett TP, Preisendorfer R (1987) Origins and levels of monthly and seasonal forecast skill for the United States surface air temperatures determined by canonical correlation analysis, *Monthly Weather Review* **115**, 1825–1850.

Donnegan JA, Veblen TT, Sibold SS (2001) Climatic and human influences on fire history in Pike National Forest, central Colorado. *Canadian Journal of Forest Research* **31**, 1527–1539.

Gershunov A, Barnett TP, Cayan DR, Tubbs T, Goddard L (2000) Predicting and downscaling ENSO impacts on intraseasonal precipitation statistics in California: the 1997/98 Event. *Journal of Hydrometeorology* **1**, 201–210.

Guttman NB (1991) Sensitivity of the Palmer Hydrologic Drought Index. *Water Resources Bulletin* **27**, 797–807.

Johnson RA, Wichern DW (1998) 'Applied multivariate statistical analysis.' (Prentice Hall: Upper Saddle River, NJ)

Jones CS, Shriver JF, O'Brien JJ (1999) The effects of El Niño on rainfall and fire in Florida. *The Florida Geographer* **30**, 55–69.

Karl TR, Knight RW (1985) 'Atlas of monthly Palmer Hydrological Drought Indices (1931–1983) for the contiguous United States.' Historical Climatology Series 3–7. (National Climatic Data Center: Asheville, NC)

Kipfmueller KF, Swetnam TW (2000) Fire–climate interactions in the Selway-Bitterroot Wilderness Area. USDA Forest Service, Rocky Mountain Research Station Proceedings RMRS-P-15-Vol.-5.

McKelvey KS, Busse KK (1996) Twentieth Century fire patterns on Forest Service lands. Sierra Nevada Ecosystems Project: Final report to Congress, Vol. II, Chapter 41. (Centers for Water and Wildland Resources: University of California, Davis)

Simard AJ, Haines DA, Main WA (1985) Relations between El Niño/Southern Oscillation anomalies and wildland fire activity in the United States. *Agricultural and Forest Meteorology* **36**, 93–104.

Swetnam TW, Betancourt JL (1990) Fire–Southern Oscillation relations in the Southwestern United States. *Science* **249**, 1017–1020.

Swetnam TW, Betancourt JL (1998) Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest. *Journal of Climate* **11**, 3128–3147.

Veblen TT, Kitzberger T, Donnegan J (2000) Climatic and human influences on fire regimes in ponderosa pine forests in the Colorado Front Range. *Ecological Applications* **10**, 1178–1195.

Veblen TT, Kitzberger T, Villalba R, Donnegan J (1999) Fire history in northern Patagonia: the roles of humans and climatic variation. *Ecological Monographs* **69**, 47–67.

von Storch H, Zwiers FW (1999) 'Statistical analysis in climate research.' (Cambridge University Press: New York)

Westerling, AL, Cayan DR, Gershunov A, Dettinger MD, Brown TJ (2001) Statistical forecast of the 2001 western wildfire season using Principal Components Regression. *Experimental Long Lead Forecast Bulletin* **10**.

Westerling AL, Brown TJ, Gershunov A, Cayan DR, Dettinger MD (2003) Climate and wildfire in the western United States. *Bulletin of the American Meteorological Society*, in press.