Regional Predictions of Annual Area Burned for the U.S. Forest Service:

Analysis of climate-wildfire interactions and long lead forecast skill for regions 1 - 6.

A Report to:
RMRS-4802

In fulfillment of Contract #: 43-82FT-5-0119

October, 2005

Prepared by:

Anthony L. Westerling
8090 Regents Rd #301
San Diego, CA 92122

(858) 623 0298
awesterl@ucsd.edu
Abstract

Since 2001 I have provided seasonal wildfire area burned forecasts to federal land managers via the National Interagency Coordination Center’s Predictive Services and the National Seasonal Assessment meetings. For the last few years, I have also provided a seasonal forecast of total area burned on Forest Service lands in the contiguous western United States. At the request of NICC Predictive Services and the Forest Service, I have undertaken a project to assess the feasibility of providing both improved seasonal forecasts by forest service administrative region, and earlier forecasts. In the past, experimental forecasts have been issued in January or February, and then again at the end of March or the beginning of April. This first of two reports describes the climatic drivers of forest wildfire season severity as measured by area burned, number of large fires and suppression expenditures by Forest Service administrative region for regions 1 – 6. This report describes the skill of forecasts of average spring and summer temperatures for the western United States made on the preceding September 1, January 1, and April 1, and demonstrates an application of the April 1 temperature forecast to improving the skill of area burned forecasts by region for western forest service regions 1-6.
Climatic influences on forest wildfire in the western US are dominated by effects of summer dryness on fuel flammability, especially at higher latitudes and elevations (Westerling et al 2003, 2005). Surplus moisture in preceding seasons can also play an important role by promoting growth of fine fuels in some mountain forests, especially open pine forests of the southwest (Swetnam and Betancourt 1998). However, fire regimes in most western U.S. forests are generally not fuel-limited.


By way of illustration, we show average Palmer Drought Severity Index (PDSI) values for years large wildfires burned and for three preceding years for US climate divisions containing 1770 forest wildfires larger than 1000 acres, versus average PDSI values for 2873 non-forest wildfires > 1000 acres (Figure 1). These were fires reported by USF, NPS, BLM and BIA (see Westerling et al 2005 re. the fire data, NCDC 1994 re. climate division data). PDSI was weighted by the number of fires in each division per year. These weights were also used to generate 95% confidence intervals with 1000 random draws from 100 years of climate division data.

Large forest wildfires occurred on average during dry summers (when most of them ignited) and were not strongly associated with moisture conditions in earlier years. This is very different from the grass and shrubland fires, which were associated only with excess moisture in preceding years. These results demonstrate the relative importance of moisture availability on fuel flammability in many western forests.

Moisture availability is affected by cumulative precipitation and temperature. Warmer temperatures reduce soil moisture via an increased potential for evapotranspiration, a reduced snowpack, and an earlier snowmelt. Snowpack at higher elevations makes some winter precipitation available as runoff in late spring and early summer, and a reduced snowpack and earlier snowmelt consequently lead to longer, drier summer fire seasons in many mountain forests.

Skill in forecasting western fire season activity will in large part depend on whether temperature can be forecast with any skill. While long-lead (i.e., greater than one season lead time) precipitation forecasts would be useful, both dynamical and statistical methodologies demonstrate more skill with modeling temperature than precipitation. In consequence, this work will primarily focus on temperature forecasts.
The next section describes the data used for this analysis. Following that is an analysis of climatic influences on wildfire by USDA Forest Service Region for Regions 1 - 6. Subsequently, sources of predictive skill for long lead wildfire forecasts are described. Next is a statistical forecast methodology for temperature and its results for three different forecast lead times, an assessment of forecast skill prior to the fiscal year. Finally, we present models for September and April forecasts of fiscal year area burned by region (1-6).

**The Data** considered here center on fiscal year area burned and suppression costs for 1977-2003 provided by USFS Rocky Mountain Research Station. Because the Forest Service’s fiscal year changed in 1976, the aggregate data provided by RMRS were only useful from 1977.

The author supplemented these with a data set derived from NIFMID fire records in which the first October to September fiscal year begins in October 1970. These were USFS large fires greater than 500 acres (200 hectares, a lower size threshold than for the multi-agency data set used in Figure 1). These data were cross-checked for errors and combined by fiscal year and USFS region to allow consideration of large fire frequency and also to categorize fire-climate relationships by coarse vegetation type (forest versus grass and/or shrubland fires) and by subregions where useful (see Westerling et al 2003 and 2005 for a description of this fire history data).

Whenever confidence intervals are indicated where both these data sets are plotted in the same figure, the intervals given are for the shorter period data (1977-2003) for convenience.

Fires were categorized by coarse vegetation characteristics by using either vegetation type and/or fuel model as reported in the NIFMID records. For the less than 1 percent of large fires with neither vegetation type nor fuel model, the University of Maryland fractional vegetation classification scheme distributed by the North American Land Data Assimilation Systems website on a 1/8 degree grid was used. For an extensive description of the methodology, see Westerling et al (2005).

“Drought indices” and “moisture anomalies” are used interchangeably here to refer to the Palmer Drought Severity Index (PDSI). PDSI is used here as an indicator of both the moisture available to plants for the production of fuels and the moisture available to wet both live and dead fuels on monthly to seasonal time scales. PDSI indicates wet conditions when it is positive, and dry conditions when it is negative. These data are available by U.S. Climate division for 1895 – present, with typically half a dozen to ten divisions per state in the western U.S. (For a discussion of PDSI see Alley 1984 and 1985, and for climate divisions see Cayan et al 1986).

While PDSI is used here as a local index, matching USFS regions roughly to coincident climate divisions, a single regional temperature index for the western U.S. is employed. In recent work (Westerling et al 2005) the author demonstrates that average spring and summer temperature (March – August) for the western U.S. is a key indicator for forest wildfire risks in the region.

While precipitation is not explicitly used here, PDSI mostly reflects surpluses or deficits in cumulative precipitation, with temperature playing a lesser role. PDSI is highly correlated with cumulative precipitation, and is often a more useful metric for fire-climate studies since it does incorporate temperature and integrates moisture anomalies over several months.
In the Northern Rockies (Region 1) fire activity was strongly associated with summer drought, and was not significantly correlated with prior-year PDSI (Figure 2). The relationships between August PDSI and area burned and suppression costs were highly nonlinear, with area burned and suppression costs increasing sharply above a threshold (Figure 3 a,b,c).

Region 1 also exhibited a strong association with regional spring and summer temperature, and a similar non-linear response in fire activity above a temperature threshold (Figure 3 d,e,f). In recent work, Westerling et al. found that the length and intensity of summer droughts and fire seasons was strongly linked to the effect of regional temperature on snowmelt timing in the Northern Rockies and other areas where the length of the average snow-free season is relatively short and co-varies strongly with temperature.

Figure 2. Correlation between Northern Rockies (Region 1) average monthly PDSI and fiscal year area burned (black solid line), frequency of fires > 500 acres (black dashed line) and suppression costs (red line). Correlations (from left) are with monthly PDSI up to two years prior to the year of the fire, and extend through December of the year of the fire (far right).

Figure 3. Northern Rockies (Region 1) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature.
In the Rocky Mountain Region (Region 2), fire activity was once again strongly associated with summer drought. Area burned, suppression costs, and large fire frequency were all strongly correlated with drought the year of the fire (Figure 4). As in Region 1, moisture surpluses and deficits in preceding years were not significantly correlated with subsequent fire activity (Figure 4).

While there was a strong non-linear response in area burned and suppression costs to extreme drought (Figure 5a-c), the relationship with regional spring and summer temperature was much weaker than in Region 1 to the north (Figure 5d-f). Westerling et al (2005) found that temperature anomalies had a much weaker effect on the longer average snow-free season in the Southern Rockies than on the much shorter snow free season in the Northern Rockies above 5500 feet.

Rocky Mountain (R2)  

a) area burned & Aug. PDSI  
b) supp. costs & Aug PDSI  
c) fire freq. & Aug. PDSI  
d) area burned & MAMJJA T  
e) supp. costs & MAMJJA T  
f) fire freq. & MAMJJA T

Figure 4. Correlation between Region 2 average monthly PDSI and fiscal year area burned (solid black), frequency of fires > 500 acres (dashed black) and suppression costs (red). Correlations (from left) are with monthly PDSI up to two years prior to the year of the fire, and extend through December of the year of the fire (far right).

Figure 5. Rocky Mountain (Region 2) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature.
In the Southwest (Region 3), associations between fire activity and drought were less clear than in Regions 1 and 2 to the north. Suppression costs were strongly negatively correlated with drought during the fire year, while area burned showed no significant correlations with drought, and large fire frequency was weakly associated with moisture surpluses a year or more before the peak fire season (Figure 6).

The apparent weak association with drought in the Southwest is likely the result of aggregating fires in very different fire regimes to produce one fiscal year total for the Region. First, Southwestern pine forests tend to be more open, with fine surface fuels playing more of a role in fire dynamics (Swetnam and Betancourt 1998). Growth of grasses and forbs that provide these surface fuels is responsive to antecedent moisture, while most years are adequately dry to cure this vegetation. Swetnam and Betancourt (1998) found, however, that fine fuels and antecedent moisture did not play the same role in mixed conifer forests of the Southwest. Furthermore, only about half the large fires in the Forest Services Region 3 fire history burned in forests. Half the fires burned in grass and shrubland fuel types where fire risks are typically strongly associated with antecedent moisture (Westerling et al 2003 and 2002).

This is readily apparent in Figure 7, where area burned and fire frequency for large forest and other wildfires are correlated separately with PDSI. Fire activity in Southwestern wildfires not burning in forests was significantly associated with moisture anomalies a year or more before the fire season. For forest wildfires, antecedent moisture was less important. PDSI during the fire season was more strongly associated with forest wildfire, but was still weaker than we would expect to see if we were to segregate fires by forest type.

Figure 6. Correlation between Region 3 average monthly PDSI and fiscal year area burned (solid black), frequency of fires > 500 acres (dashed black) and suppression costs (red). Correlations (from left) are with monthly PDSI up to two years prior to the year of the fire, and extend through December of the year of the fire (far right).

Figure 7. Correlation between Region 3 average monthly PDSI and fiscal year area burned (solid) and frequency of fires > 500 acres (dashed) for forest wildfires (green) versus all other wildfires (brown). Correlations (from left) are shown for PDSI from up to two years prior to the fire year, and extend through December of the year of the fire (far right).
Southwest (R3)
a) area burned & Aug. PDSI  
b) supp. costs & Aug PDSI  
c) fire freq. & Aug. PDSI  
d) area burned & MAMJJA T  
e) supp. costs & MAMJJA T  
f) fire freq. & MAMJJA T

Figure 8. Southwest (Region 3) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature. Area and Frequency are for wildfires in forest (green) and other (brown) vegetation types. Costs are aggregated for all fires.

Area burned in large fires (> 500 acres) was weakly associated with summer drought for forest wildfires (all acres correlation = 0.39, forest acres correlation = 0.41, non-forest acres = 0.24) (Figure 8a). Large wildfire frequency was not significantly associated with summer drought (all fires correlation = -0.18, forest fires correlation = -0.24, non-forest fires = -0.03). (The applicable critical value for statistical significance at the 95% confidence level—a correlation of +/- 0.36—is lower than that shown in Figures 7 & 8 because more data were available.)

Area burned in large forest fires was weakly associated with spring and summer temperature (Figure 8d). (all acres correlation = 0.19, forest acres correlation = 0.36, non-forest acres correlation = -0.12, critical value for 95% confidence level = +/- 0.36). Large fire frequency was also very weakly correlated with temperature (all fires correlation = 0.31, forest fires correlation = 0.35, non-forest fires correlation = 0.17).

Likewise, the correlation between suppression costs for Region 3 and temperature was weak (Figure 8e) (cor. = 0.35, p = 0.073). The only strong relationship for the Southwest was between suppression costs and summer drought (cor. = -0.69, p = 0.00).

While non-forest acres and fire starts were significantly correlated with PDSI a year or more before the fire season, this relationship was not strong enough to meaningfully forecast suppression costs or aggregate area burned.
In the Intermountain Region, (Region 4), fire activity was strongly associated with summer drought and with regional spring and summer temperatures. Unlike the Southwest, over 70% of large fires were forest fires. The only significant correlations between moisture surplus or deficit and aggregate fire frequency, area burned and suppression costs were in the year the fires burned, peaking in August and September (Figure 9).

Like the Northern Rockies (Region 1), area burned and large fire frequency showed a nonlinear response to both summer drought indices in the Intermountain Region (Figure 10a-c) and to regional temperature (Figure 10d-f). Suppression costs were strongly correlated with temperature (cor. = 0.62, p = 0.00) and with August PDSI (cor. = -0.67, p = 0.00).

Figure 9. Correlation between Region 4 average monthly PDSI and fiscal year area burned (solid black), frequency of fires > 500 acres (dashed black) and suppression costs (red). Correlations (from left) are with monthly PDSI up to two years prior to the year of the fire, and extend through December of the year of the fire (far right).

Intermountain (R4)

a) area burned & Aug. PDSI  b) supp. costs & Aug PDSI  c) fire freq. & Aug. PDSI

d) area burned & MAMJJA T  e) supp. costs & MAMJJA T  f) fire freq. & MAMJJA T

Figure 10. Intermountain (Region 4) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature.
In California, (Region 5), fire activity was strongly associated with summer drought and to a lesser extent with surplus moisture a year or more before the fire season. Northern and Southern California were considered separately (NorthOps v.s. SouthOps) using our data set of USDA Forest Service wildfires > 500 acres and drought indices constructed from climate divisional PDSI in northern and southern California. The PDSI indices were nearly identical (cor = 0.9), and of course the suppression costs were the same, since only fiscal year suppression costs aggregated for all of Region 5 were available.

The frequency of large fires and the area burned in large fires in Northern California were strongly associated with drought indices throughout most of the year the fire burned, and were not strongly associated with moisture anomalies in earlier years (Figure 11). Very similarly, in Southern California, large fires and large acres were strongly associated with drought conditions through much of the calendar year containing most of the respective fiscal year (Figure 12). However, Southern California large fires and acres burned in large fires were also correlated with excess moisture the preceding year (exceeding the 95% confidence level) (Figure 12). This reflects the fact that two thirds of the large fires in the North were forest fires, while only one fifth of the fires in the South were forest fires.

Comparing scatter plots of August PDSI and area burned for Northern and Southern California (Figure 13a), area burned in both regions show a nonlinear response to summer drought. Fire starts in the south are not as strongly associated with summer drought as in the north (Figure 13c).
California (R5)

Figure 13. California (Region 5) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature. Area burned (in a & b) and fire frequency (in c & f) are for large fires (> 500 acres) aggregated for Northern (points denoted by blue ‘N’) and Southern (points denoted by purple ‘S’) California. Suppression costs are for Region 5 as a whole. Area, Frequency and Cost are all for fiscal years. Acres burned in panel a and fire frequency in panel c are plotted against different drought indices by subregion… Northern area/frequency are plotted against northern PDSI, Southern area/frequency are plotted against southern PDSI.

Fire activity in both northern and southern California is not as strongly associated with the regional temperature signal as with summer drought (Figure 13). Westerling et al (2005) found that regional temperature anomalies were inversely correlated with cumulative water year precipitation in California (warm years were low precipitation years). Temperature also had a strong impact on the timing of the spring melt and the length of the snow-free season at higher elevations in the Sierra Nevada, but most of the forest area and fires were at lower elevations where the precipitation deficit was the controlling factor in determining the severity of spring and summer drought.

Finally, while fiscal year area burned in large Region 5 forest fires was not significantly correlated with antecedent years’ moisture anomalies, it is the case that all of the largest area burned years occurred a year following at least one winter month with moderately wet to very wet conditions (Figure 14c). Conversely, in fire seasons where this condition was not met (ie, when winter months a year before the fire season were all dry), forest fire area burned was always subsequently low (Figure 14c).
California (Region 5)

a) Forest Acres vs August PDSI

b) Non-forest Acres vs August PDSI

c) Forest Acres vs Max Oct – May PDSI (one year or more prior to the fire season)

d) Non-forest Acres vs Max Oct – May PDSI (one year or more prior to the fire season)

Figure 14. a) Fiscal year acres burned for all forest fires > 500 acres in Region 5 vs August PDSI. b) Fiscal year acres burned for all fires > 500 acres and not in forest vs August PDSI. c) forest and d) non-forest acres versus maximum PDSI (wettest winter monthly PDSI) for October – May twelve to five months prior to the start of the fiscal year.
In the Northwest, (Region 6), fire activity was strongly associated with drought concurrent with the fiscal year that fires burned, and more weakly with extended drought in prior years (Figure 15). And, the response to extreme drought was highly non-linear, with area burned, suppression costs and large fire frequency increasing sharply below a threshold drought level (Figure 16a – c).

As in other more northerly regions, fire activity was strongly associated with temperature (Figure 16d – f). (area cor. = 0.76, p=0.0, cost cor. = 0.67, p=0.0, frequency cor. = 0.66, p=0.0)

Figure 15. Correlation between Region 6 average monthly PDSI and fiscal year area burned (solid black), frequency of fires > 500 acres (dashed black) and suppression costs (red). Correlations (from left) are with monthly PDSI up to two years prior to the year of the fire, and extend through December of the year of the fire (far right).

Northwest (R6)

- a) area burned & Aug. PDSI
- b) supp. costs & Aug PDSI
- c) fire freq. & Aug. PDSI

- d) area burned & MAMJJA T
- e) supp. costs & MAMJJA T
- f) fire freq. & MAMJJA T

Figure 16. Northwest (Region 6) area burned, suppression costs, and frequency of fires > 500 acres versus August PDSI and average western US March through August temperature.
The ability to forecast forest wildfire activity at long lead times is limited by the relative importance of climate influences during and immediately prior to the fire season. By long lead-time, we mean forecasting at the end of September (the start of the fiscal year) for wildfire activity in summer the following year. The major sources of “memory” that can be exploited for such a long lead wildfire forecast are the moisture reservoir in soils, accumulations of fine fuels, and sea surface temperatures.

Soil moisture is highly auto-correlated (i.e., wet or dry conditions tend to persist for a time). However, the longer the lead-time considered, the more likely it is that unusually wet or dry conditions will revert to the mean. The correlation between August PDSI and PDSI a year earlier is generally not significantly different from zero (Figure 17).

Antecedent moisture surpluses can affect subsequent fire seasons by enabling the growth of vegetation that fuels wildfires, especially fine surface fuels (grasses, forbs, etc). However, many of the large wildfires that account for the majority of the area burned and suppression costs of the USDA Forest Service occur in forests where the availability of fine surface fuels is not a limiting factor for the ignition and spread of fires. This is reflected in the lack of any clear pattern in antecedent PDSI for forest wildfires in Figure 1. This is particularly an issue for forecasting fire activity in the Northern Rockies, which accounts for nearly one quarter of USFS fires over 500 acres in the western U.S., and about a third of the area burned in those larger fires. In Western forests at lower elevations and latitudes, antecedent moisture appears to play a larger, albeit still minor, role that may afford some forecast skill. In the Northwest, forest fire activity appears to be associated with multiyear drought, (Figure 15), so there may be some forecast skill to be derived from moisture deficits observed in antecedent years.

The Pacific Ocean can be thought of as a heat reservoir, with variations in water temperature driving variability in climate processes over the western region. The Pacific Decadal Oscillation and El Nino/La Nina are well-known indices that describe spatial patterns in Pacific Ocean sea surface temperatures (sst’s) that oscillate on decadal and multiyear cycles respectively. Together these modes of variability in ocean temperatures are the strongest source of interannual variability in climate for much of the western U.S. We explore the use of Pacific sst’s below to forecast for the fire season using observations through three lead times: August before the start of the USFS fiscal year, and the following December and March. (While the fiscal year starts in October, data for September is not always available promptly by the start of the following month, so the earliest forecast was made using observed SSTs and PDSI from August.)
**The Statistical Methodology** employed is a variation on that used to forecast wildfire area burned by western U.S. ecosystem provinces (described by Westerling et al 2002, Alfaro et al 2005). To summarize: A large number of predictors are prefiltered with principle components analysis (PCA). This procedure produces a tractable number of linearly independent time series that summarize most of the information contained in the original data set. Likewise, numerous variables to be explained are prefiltered with PCA.

The new predictor and predictand principal component series are then related to each other using canonical correlation analysis (CCA). Whole-field metrics that summarize the model skill for all the predictand time series with one number are calculated for each possible combination of principle components and canonical correlation pairs using jackknife cross-validation. Metrics employed here were the sum of squared positive correlations and the sum of correlations between observed and modeled average March to August temperature by climate division. The number of principle components and canonical correlation pairs is selected that yields a concise solution with the highest forecast skill over the entire region, as measured by these summed correlation coefficients.

In this case, North Pacific sea surface temperatures and western US climate division PDSI observed in August, December or March were used as predictors of average March to August temperatures for 110 western US climate divisions covering Washington, Oregon, California, Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico, North Dakota, South Dakota and Nebraska. The SST and PDSI data sets were prefiltered with PCA separately, and then the resulting PC were combined into one data set and prefiltered with PCA a second time (Alfaro et al 2005). The forecast March – August average temperatures for 110 western Climate Divisions were then averaged to obtain a regional average temperature forecast. This regional average temperature has been shown to be highly correlated with forest wildfire activity throughout much of the western US (see preceding discussion of climate-fire interactions and Westerling et al 2005).

North Pacific SSTs and western US PDSI were chose as predictors here based on work done by Alfaro et al (2005). Alfaro et al successfully used these variables to predict maximum temperatures in the western US on monthly to seasonal time scales with the same methodology described above. Patterns in the Pacific SST field may give some indication of near future variability in atmospheric circulation, temperature, and precipitation over western North America. Local soil moisture is thought to influence the evolution of summer temperatures, with higher temperatures occurring over dry soils. Moist soils regulate temperature via evaporation until the moisture reservoir is exhausted.

Using August, December and March PDSI and SST values allowed for forecasts to be made both at the start of the fiscal year in October, and in January and April for comparison. All of the CCA model results presented here were obtained using leave-one-out cross-validation (see Westerling et al 2002), and are indicative of the true forecast skill to be expected.

**The Results** are not encouraging for a long lead forecast issued in September before the fiscal year start. Cross-validated correlations for the three forecasts with the observed regional temperature average are: for the September forecast, cor=0.27 (p=0.13), for the January forecast, cor=0.41 (p=0.02), for the April forecast, cor=0.44 (p=0.01).
While the January and April forecasts offer some real, albeit modest, forecast skill, the September forecast is not even minimally significant. In order to allow for a more direct comparison between the forecasts, the April forecast did not incorporate observed March temperatures, so it could be further improved. Interestingly, the improvement in forecast skill for temperature from a forecast made in January to a forecast made in April is only incremental, indicating that the influence of Pacific SSTs on spring and summer temperatures is largely determined by mid-winter.

This analysis would offer some hope that forecasts can be made in mid-winter for the peak summer fire season with modest skill. For much of the Forest Service' western regions, however, this analysis has demonstrated that fire activity is strongly associated with factors like spring and summer temperature that are not readily forecast with any skill three or more seasons prior to the peak summer months of the fire season.

Forecast models for fiscal year area burned in each of the six western USFS regions in the contiguous US in September and April are presented below, incorporating local PDSI observed in August or March and the regional temperature forecast made in September or April as predictors in cross-validated linear regression models for each western region. Forecast skill in April is quite high, explaining over two thirds of variance in fiscal year area burned for the western region as a whole, while forecasts in August do not show any skill outside of Region 6.

For each region’s forecast model, the analysis of forecast skill is based on leave-one-out cross-validation using data for 1977-2003. This means that for each of the 27 fiscal years up through 2003, a model was estimated using the observations from the other 26 fiscal years. This model was then applied to the observed PDSI and forecast temperature for the fiscal year in question to produce a prediction, or retrospective forecast, for that fiscal year. The result is 27 sets of model coefficients for each forecast model in each region.

Correlation and R-squared statistics for these cross-validated retrospective forecasts are lower than would be the case for a single model estimated with all 27 years’ data at once, and are a conservative measure of the forecast skill. The accuracy of the retrospective forecasts for 1977-2003 is indicative of the true forecast skill—the result for each year is not affected by knowing the area burned for that year.
The models are lognormal: the variable forecast is the logarithm of area burned. Without the log transformation, area burned is not normally distributed.

Forecasts for 2004 and later use the same model specifications as for the cross-validated models, estimated on the 1977-2003 period’s data.

R-squared for the April forecasts was highly significant for all but Region 3, while only Region 6 showed any skill for the September forecast (Table 1). Note that the skill for Region 6 in September was entirely from local observed PDSI… the September forecast of spring and summer average temperature added nothing to forecast skill for Region 6. April forecast skill in Region 3 is likely affected, at least in part, by the effects of the monsoon, which are not likely to be captured by the predictors used here.

<table>
<thead>
<tr>
<th>Region</th>
<th>Sep Forecast R2, rho, (rho p-value)</th>
<th>Apr Forecast R2, rho, (rho p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.11, -0.03, (0.87)</td>
<td>0.49, 0.71, (0.00)</td>
</tr>
<tr>
<td>2</td>
<td>-0.14, -0.13, (0.58)</td>
<td>0.43, 0.67, (0.00)</td>
</tr>
<tr>
<td>3</td>
<td>0.03, 0.25, (0.21)</td>
<td>0.06, 0.33, (0.10)</td>
</tr>
<tr>
<td>4</td>
<td>-0.10, 0.05, (0.82)</td>
<td>0.27, 0.55, (0.00)</td>
</tr>
<tr>
<td>5</td>
<td>-0.02, 0.22, (0.28)</td>
<td>0.43, 0.67, (0.00)</td>
</tr>
<tr>
<td>6</td>
<td>0.20, 0.46, (0.02)</td>
<td>0.57, 0.76, (0.00)</td>
</tr>
<tr>
<td>West</td>
<td>-0.06, 0.13, (0.53)</td>
<td>0.65, 0.82, (0.00)</td>
</tr>
</tbody>
</table>

Table 1. Forecast skill for area burned by region and westwide, 1977-2003. Cross-validated R-squared and correlation (rho), with p-values for correlation t-statistics.

September Cross-validated Linear Regression Forecast Model Specifications:

Region 1: \[ \ln(\text{Area}) \sim T_{\text{forecast}(S)} \]

Region 2: \[ \ln(\text{Area}) \sim T_{\text{forecast}(S)} \]

Region 3: \[ \ln(\text{Area}) \sim T_{\text{forecast}(S)} \]

Region 4: \[ \ln(\text{Area}) \sim T_{\text{forecast}(S)} + \text{PDSI}_{\text{max}(-1)} \]

Region 5: \[ \ln(\text{Area}) \sim T_{\text{forecast}(S)} + \text{PDSI}_{\text{max}(-1)} \]

Region 6: \[ \ln(\text{Area}) \sim \text{PDSI}_{\text{Aug}(-2)} \]

\[ \ln(\text{Area}) \] is the fiscal year total area burned by region. \[ T_{\text{forecast}(S)} \] is the average March through August temperature for the western US climate divisions forecast in September prior to the fiscal year start. \[ \text{PDSI}_{\text{max}(-1)} \] is the maximum monthly PDSI value for the winter one year before the fire season (i.e., October two years before the fire season through May one year before the fire season) for selected climate divisions within the particular region being forecast. In Region 5 this \[ \text{PDSI}_{\text{max}(-1)} \] is for climate divisions in and around southern California. \[ \text{PDSI}_{\text{Aug}(-2)} \] in region 6 is the average PDSI value for climate divisions in the Pacific Northwest for the month of August two years prior to the peak summer fire season (thirteen months prior to the start of the fiscal year).

April Cross-validated Log-Linear Regression Forecast Model Specifications:

Region 1: \[ \ln(\text{Area}) \sim T_{\text{forecast}(A)} \times \text{PDSI}_{\text{Mar}(0)} \]

Region 2: \[ \ln(\text{Area}) \sim T_{\text{forecast}(A)} \times \text{PDSI}_{\text{Mar}(0)} \]

Region 3: \[ \ln(\text{Area}) \sim T_{\text{forecast}(A)} \times \text{PDSI}_{\text{Mar}(0)} \]

Region 4: \[ \ln(\text{Area}) \sim T_{\text{forecast}(A)} \times \text{PDSI}_{\text{Mar}(0)} + \text{PDSI}_{\text{max}(-1)} \]

Region 5: \[ \ln(\text{Area}) \sim \text{PDSI}_{\text{Mar}(0)} + \text{PDSI}_{\text{max}(-1)} \]

Region 6: \[ \ln(\text{Area}) \sim T_{\text{forecast}(A)} + \text{PDSI}_{\text{Mar}(0)} + \text{PDSI}_{\text{Aug}(-2)} \]

\[ \ln(\text{Area}) \] is once again the logarithm of fiscal year total area burned by region. \[ T_{\text{forecast}(A)} \] is the average March through August
temperature for the western US climate divisions forecast in April six months after the fiscal year start, but prior to the start of most of the western forest wildfire season. PDSI\textsubscript{max(-1)} and PDSI\textsubscript{Aug(-2)} are the same as before, while PDSI\textsubscript{Mar(0)} is the average PDSI observed over the previous month (March) for selected climate divisions in the region for which area burned is being forecast. Model specifications denoted here with “x” indicate the inclusion of an interaction term (i.e., $T_{\text{forecast(A)}} \times \text{PDSI}_{\text{Mar(0)}}$ means $T_{\text{forecast(A)}} + \text{PDSI}_{\text{Mar(0)}} + T_{\text{forecast(A)}} \times \text{PDSI}_{\text{Mar(0)}}$). In Region 5, PDSI\textsubscript{max(-1)} is for climate divisions in southern California as before, while PDSI\textsubscript{Mar(0)} is for northern California.

While the correlation and R\textsuperscript{2} for the April forecast models’ is highly significant, it is important to acknowledge the limitations of these kinds of models. The forecast variable is the log of area burned, and while the forecast skill is good, in very active fire years the forecast can greatly understate the observed area burned. For example, in Region 1 in 1988, forecast log area burned is 11.8, while observed log area burned is 14.4. The difference between these two numbers is over 1.5 million acres, or nearly 12 times the level forecast for that year.

These forecast models should not be relied upon to produce an explicit point forecast for area burned. Rather, they are best employed as indicators of whether the coming fire season is going to result in “high” or “low” area burned. For example, forecasts and observations can be binned into terciles corresponding to the lowest third, middle third and highest third of values for the twenty-seven years from 1977 to 2003, and compared in contingency tables. (e.g. Region 1 in Table 2).

Note that the observed area burned was in the highest third in six out of the nine years in which forecasts were in the highest third for Region 1 area burned. The probability of a ‘surprise’, where area burned was high when the forecast was low, was only one in nine years. In the example from 1988 discussed above, both the observation and the forecast were in the highest tercile. The forecast was accurate in that it predicted a “high” area burned that year, but the simple statistical model did not anticipate the extreme conditions experienced that year.

The April forecast skill for the contiguous western U.S. as a whole is higher than for any individual Forest Service Region (Table 1, Figure 20); the forecast errors in individual regions tend to cancel each other out over the period of record. The contingency table for the western U.S. (Table 3) demonstrates that the aggregate forecast is particularly good at forecasting low area burned years: eight out of the nine years when forecasts were in the bottom third, the observed area burned was also in the bottom third. Likewise, there were no high area burned surprises following a low April forecast. Judging by reports that 2004 was a particularly low area burned year in the contiguous western US, the low 2004 forecast (Figure 20) would appear to confirm this forecast skill. The forecast skill for high area burned years is slightly better than for Region 1.

<table>
<thead>
<tr>
<th>Observed</th>
<th>Low</th>
<th>Mid</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Mid</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Hi</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

**Table 2:** Region 1 Terciled April Forecast vs Observed Fiscal Year Area Burned

<table>
<thead>
<tr>
<th>Forecast</th>
<th>Low</th>
<th>Mid</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mid</td>
<td>0</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Hi</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

**Table 3:** Western Terciled April Forecast vs Observed Fiscal Year Area Burned
Figure 19. April forecast models versus observations for fiscal year area burned, Regions 1-6.
Figure 20. April forecast model versus observations for fiscal year area burned aggregated over Regions 1-6.

**In Conclusion,** April forecast skill—six months after the start of the fiscal year, but prior to most of the summer fire season—is very good. September forecast skill—prior to the start of the fiscal year—is nonexistent using the methods described in this analysis. The one possible exception shown here was in the Northwest, where persistent drought explained about twenty percent of variance in wildfire a year or more later.

The key point is that USFS area burned in the western U.S. is very sensitive to temperature, and temperature is hard to forecast a year or so in advance.

Looking to the future, are there other possibilities to explore that might allow for some useful skill forecasting fiscal year area burned prior to the fiscal year start? It seems reasonable that we should be able to exploit persistence in PDO and forecasts of ENSO to forecast fire activity, since these indices are thought to be associated with temperature anomalies, particularly in the Northwest and Southwest. In practice, however, spring and summer temperatures in the western US since the 1970s do not appear to be highly correlated with these indices.

It might be the case that a relationship between PDO, ENSO and spring and summer temperatures in the western US in recent decades may have been obscured by a strong secular trend in spring temperatures in the western US that is not a result of the interannual to decadal variability described by these indices. The author recommends that any future work explore this issue. It is possible that by filtering out the trend in temperature, associations between PDO, ENSO might be more easily detected. Using these associations in combination with persistence in the temperature trend might provide a basis for making long lead forecasts of fiscal year area burned by region.

In addition to examining other sources of long lead predictive skill prior to the fiscal year, some improvements could be made to the April forecast. Since the forecast is made in six months into the forecast season, estimated area burned to date as of the end of March for each region could be incorporated as an additional predictor in the cross-validated linear regressions for each region. Similarly, the temperature forecasts used as predictors in the regression models could also incorporate observed March temperature as an additional predictor field in the CCA. Also, the specification might be improved for some or all regions by using an inverse gamma specification for the error terms in the regression instead of a logarithmic transformation of the predictand (area burned).

The cross-validated forecast skill for temperature in January was nearly as good as for temperature forecasts made in April. This forecast skill for temperature could be exploited to provide a January forecast for area burned. This area burned forecast would not be as good as the April area burned forecast, since observed PDSI as of
the start of January is not as strong a predictor of area burned as is March PDSI, but the model would still provide a forecast with significant skill three months earlier than what is presently available.

Finally, we are in the process of compiling detailed large fire histories for Alaska and parts of the Southeastern U.S. This data project will be complete by early 2006, and would be invaluable for any analysis of long lead forecast skill for USFS in these regions. Any future analysis could address long lead forecast skill in the Southern, Eastern and Alaska regions, in addition to the contiguous western US (regions 1-6).

References


