Statistical Forecasts of the 2003 Western Wildfire Season Using Canonical Correlation Analysis

Anthony L. Westerling[1], Alexander Gershunov[1] and Daniel R. Cayan[1,2]

[1]*Scripps Institution of Oceanography, La Jolla, California*
[2]*US Geological Survey, La Jolla, California*

Experimental Long-Lead Forecast Bulletin

Volume 12, Numbers 1 and 2, pp 49 – 53

www.iges.org/ellfb

Statistical Forecasts of the 2003 Western Wildfire Season Using Canonical Correlation Analysis

contributed by Anthony L. Westerling[1], Alexander Gershunov[1] and Daniel R. Cayan[1,2]

[1]*Scripps Institution of Oceanography, La Jolla, California*
[2]*US Geological Survey, La Jolla, California*

Experimental forecasts for the 2003 fire season indicate low area burned in most western deserts and basins, high area burned in the southern Rocky Mountains and at higher elevations in Arizona and New Mexico, and mid to high area burned in the Sierra Nevada. This pattern—largely a continuation of that seen in 2002—is the result of persistent drought. If parts of the western U.S. experience significant additional precipitation before the start of the western fire season, high forecasts for mountain provinces could change, but low forecasts for basins and deserts are unlikely to be affected.

Statistical forecasts of seasonal area burned using observed Palmer Drought Severity Index (PDSI) values at lead times of a season to two years in advance of the fire season as predictors have successfully predicted many important features of the western United States' wildfire season (*Westerling et al 2001, 2002, 2003*). PDSI values observed for the month of March were particularly important for forecasting area burned at higher elevations along the West Coast in these statistical models. Consequently, our area burned forecasts for the fire season are produced in early April. This is inconvenient, since some fire management applications for which these forecasts might be useful require information earlier in the year. This is true, for example, for early requests for supplemental resources ("severity funds" in the parlance of fire managers) and for decisions about hiring and training temporary fire suppression personnel.

To accommodate requests for earlier forecasts of area burned for the western US' 2003 fire season, we produced experimental area burned forecasts in January and February 2003. The January forecast uses Pacific sea surface temperatures observed in December to forecast March PDSI by climate division. The February forecast uses persistence in the January 2001 PDSI to model March. (We will refer to these models below as the SST and Persistence models for convenience). We present both forecasts here, with a brief analysis of their retrospective skill for 1980 – 2002.

We use PDSI values for western US climate divisions for the two years preceding a fire season as predictors for area burned in the fire season. Moisture anomalies in preceding years are statistically important for forecasting wildfire in dry basins and deserts of the western US, where wildfire season severity is dominated by the availability of fine fuels with a short growth and curing cycle. Moisture in the winter before the summer fire season is important in areas like the Sierra Nevada, where fire regimes are dominated by the effect of winter precipitation on the flammability of heavy fuels later in the year. The six months through the end of March account for over three quarters of an average year's precipitation in climate divisions covering California and western Oregon and Washington, and half or more of an average year's precipitation in eastern Oregon and Washington and much of Arizona, Nevada, Idaho and Utah (*Figure 1*). March is typically the last chance for many western US climate divisions to experience significant precipitation during the water year, and March precipitation can be highly variable year to year. Thus, the PDSI by the end of March is an important indicator for fuel conditions in the upcoming fire season. Parts of Arizona, New Mexico, Colorado, Wyoming and Montana experience more precipitation in summer. While this moisture is important for fire regimes there, forecasts of summer precipitation are beyond the scope of this research.

The forecast model is estimated using Canonical Correlation Analysis (CCA) to calculate linear relationships between principal components of seasonal acres burned aggregated by Bailey's ecosystem provinces (see Bailey *et al* 1994) and leading values of PDSI for US climate divisions, as described in detail by Westerling *et al* (2002). Here we use the model coefficients derived in that study, but replace observed March PDSI with forecast March PDSI as a predictor.

Acres burned were summed for 17 Baileys ecosystem provinces for fires starting between May 1 and October 31 and scaled using a $\log_{10}$ transformation. These data were compiled (Westerling *et al* 2001) from fire reports from the USDA Forest Service and the Department of Interior's Bureau of Land Management, Bureau of Indian Affairs, and National Park Service for the period 1980 – 2000. These 17 time series were summarized by six leading principle components (PCs) accounting for over 85% of total variance.

For predictors, 110 western U.S. Climate Division PDSI series were used at five different leads: December and March immediately preceding, March and August one year previous to, and August two years prior to the fire season, for a total of 550 predictor variables. Six leading PC series comprising more than 70 percent of total variance for these data were used as the predictor series. PDSI is used here as a rough proxy for anomalous moisture available for growth and wetting of fuels.

As described in Westerling *et al* (2002), a canonical correlation analysis (CCA) was used to find optimally correlated patterns in the area burned and PDSI PCs. (For a detailed CCA methodology see also Barnett and Preisendorfer 1987, Johnson and Wichern 1998, Gershunov *et al* 2000). A linear forecast model was constructed using the first three canonical correlation pairs (CCs) calculated for the six area burned and six PDSI PCs. All results shown here are for a series of cross-validated models; retrospective forecasts for each year from 1980 to 2000 were calculated using "leave-one-out" cross-validation.

We use two different methods to estimate March PDSI. First, March PDSI for western US climate divisions is estimated with the same CCA methodology as above, using eight CCs derived from eight PC series describing March PDSI for 1951 - 2000 and eight PC series describing December Pacific SST for 1950 – 1999. Second, January PDSI is persisted to model March PDSI. Forecasts are presented as terciles (low, medium, and high area burned).

Probabilities for observing high or low area burned conditional on having forecast high or low area burned are similar for forecasts using March PDSI derived from Pacific SST, persistence in January PDSI, and observed March temperatures and precipitation (*Figure 2*). The chief benefit of using observed March PDSI versus forecasts of March PDSI is improved skill for high area burned forecasts in the Sierra Nevada, Southern Cascades and Coast Ranges in Northern California, and at higher elevations in Nevada and Utah. Skill for high area burned forecasts using persistence in January PDSI is very similar to that for models using March PDSI estimated from Pacific SST. Skill for low area burned forecasts using persistence is greater over much of the intermountain West than that for forecasts using Pacific SST.

Both of the experimental forecasts for 2003 fire season indicate low area burned in most western deserts and basins, and high area burned in the southern Rocky Mountains and at higher elevations in Arizona and New Mexico (*Figure 3)*. This pattern is largely a continuation of that seen in 2002 (Westerling 2002). The statistical models for desert grass- and shrub-land provinces put the most weight on moisture anomalies occurring a year before the fire season, with positive (negative) moisture anomalies producing higher (lower) area burned in subsequent years. Conversely, models for higher elevations with heavier fuels tend to put the greatest weight on the most recent moisture anomalies, with negative (positive) anomalies producing higher (lower) area burned forecasts for the next fire season. Since much of the West has been quite dry for the last few years, the prediction for the basins and deserts is for a low severity fire season, and for a high severity season in dry forest lands.

The forecasts differ with respect to the Sierra Nevada and the Colorado Plateau. The SST model predicts a mid-tercile year for both, but with no skill. SST and Persistence model forecasts for neutral years in both locations perform worse than random chance (ie, less than 33% probability of observing mid-tercile area burned given a forecast of mid-tercile area burned). The Persistence model forecasts a high area burned fire year for the Sierra Nevada, with skill better than random chance (4 out of 7 correct, or 57%), but not as high as for forecasts using observed March PDSI (where top tercile forecasts have been correct 5 out of 7, or 71% of the time). California climate divisions receive a higher percentage of annual precipitation in March on average than is true elsewhere in the West, and March precipitation in California is highly variable, so it is not surprising that the skill of these early forecasts for the Sierra

Nevada is lower than that obtained using March PDSI calculated from observed precipitation.

In Figure 4, probabilities for observing low, mid and high area burned conditional on the tercile of this year's Persistence forecast are presented. These probabilities were calculated for 21 cross-validated retrospective forecasts for 1980 – 2000. Note the low or zero probabilities of observing a "high" year when a "low" year has been forecast in the Great Basin (*Figure 4a – c*), Mojave (*Figure 4d*), and Colorado Plateau (*Figure 4e*). For the mountain provinces with high Persistence forecasts (*Figure 4f – h*), high forecasts are less skillful then the low forecasts in the Intermountain West, but the probability of a surprise (observing the opposite of the forecast extreme) is still quite low. The greater uncertainty for the mountain province forecasts is likely a result of fire regimes in these areas being more strongly affected by the preceding winter's precipitation and by summer drought concurrent with the fire season. (See Westerling 2002 and 2003 for a detailed analysis of these fire – climate interactions).

The low forecast for basins and deserts is unlikely to change by the end of March, since the main effect of March moisture anomalies in these models would be on the 2004 fire season forecast. The high forecast for the Sierra Nevada, southern Rockies, and mountains in Arizona and New Mexico might still be affected by weather experienced over the next few weeks.

Updated forecasts for the 2003 fire season using March PDSI will be made available in April at http://www.metora.ucsd.edu/cap.

References

Bailey RG, Avers PE, King T, McNab WH, eds. (1994) 'Ecoregions and subregions of the United States (map).' Washington, DC: USDA Forest Service.

Barnett, T. P., and Preisendorfer, R, 1987, Origins and levels of monthly and seasonal forecast skill for the United States surface air temperatures determined by canonical correlation analysis, *Mon. Wea. Rev.*, 115, 1825-1850.

Gershunov, A, T. P. Barnett, D. R. Cayan, T. Tubbs, L. Goddard, 2000, Predicting and Downscaling ENSO impacts on Intraseasonal Precipitation Statistics in California: The 1997/98 Event: *J. of Hydrometeorology*, 1, 201:210.

Johnson, R. A. and Wichern D. W., 1998, *Applied Multivariate Statistical Analysis*, Prentice Hall, 816 pp.

Osmond, C. B., Pitelka, L. F. and Hidy, G. M., (Eds.), 1990, *Plant Biology of the Basin Range*. Ecological Studies 80, Springer-Verlag, 375 pp.

Westerling, A. L., A. Gershunov, T. J. Brown, D. R. Cayan, M. D. Dettinger, 2003: "Climate and Wildfire in the Western United States," *Bulletin of the American Meteorological Society*. In Press.

Westerling, A. L., A. Gershunov, D. R. Cayan, T. P. Barnett, 2002: "Long Lead Statistical Forecasts of Western U.S. Wildfire Area Burned," *International Journal of Wildland Fire* 11(3,4): 257 – 266.

Westerling, A. L., D. R. Cayan, A. Gershunov, M. D. Dettinger, T. J. Brown, 2001: "Statistical Forecast of the 2001 Western Wildfire Season Using Principal Components Regression," *Experimental Long-Lead Forecast Bulletin*, March 2001, 10: 71-75.

This research is supported by the NOAA Office of Global Programs through the California Applications Program and the Experimental Climate Prediction Center, and from a National Fire Plan grant to the USDA Forest Service's Southern Research Station.

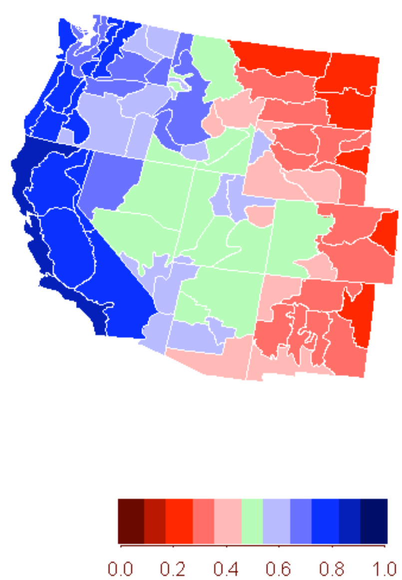

Figure 1: Percent of annual precipitation occurring October through March.

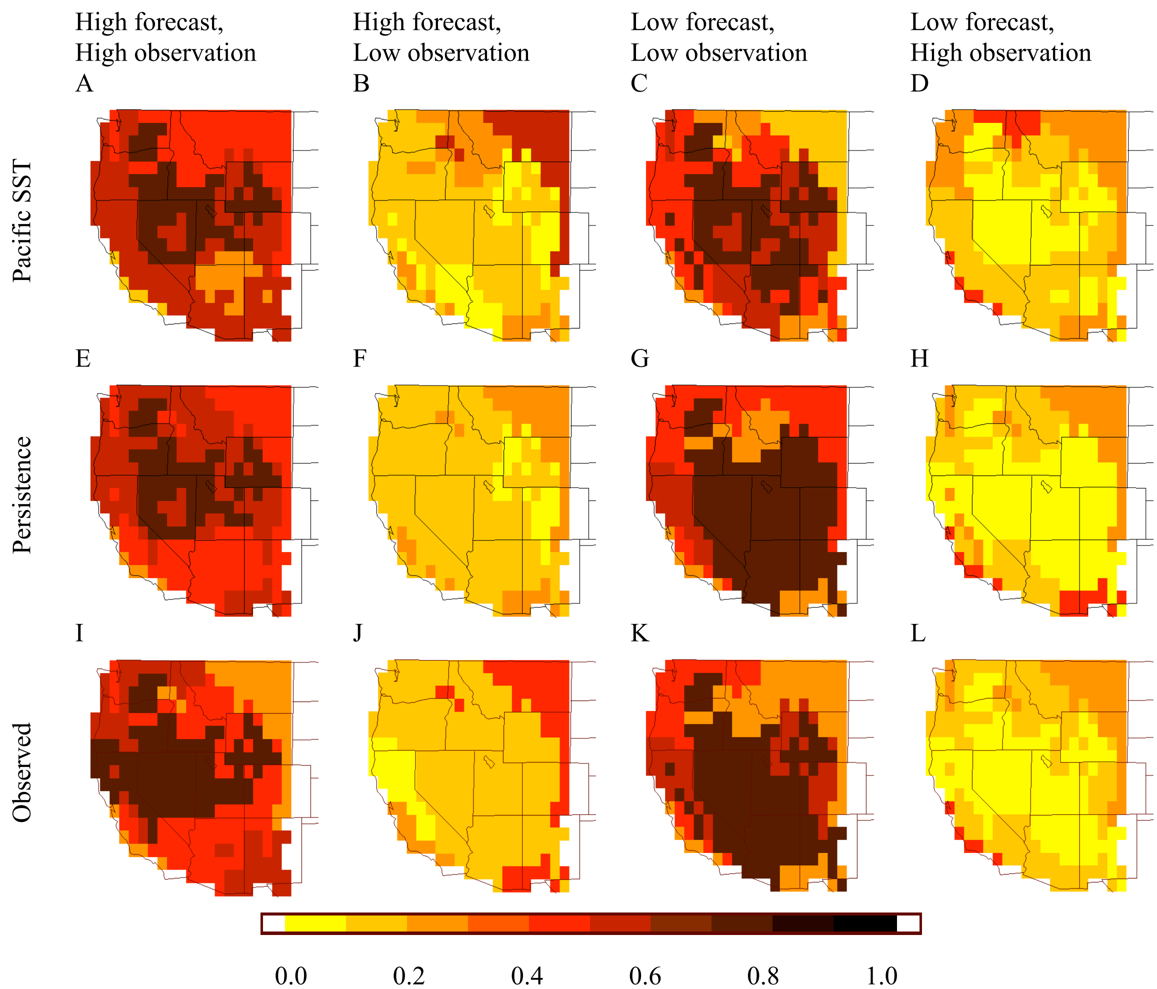

Figure 2: Probability of observing top or bottom tercile area burned given a forecast for top or bottom tercile area burned, where the March PDSI used as a predictor is derived from (A - D) Pacific SST, (E - H) persistence of January PDSI and (I - L) observed March PDSI. The darkest shading for correct forecasts (A, E, I, C, G, K) mapped above corresponds to a probability of a correct forecast of 71%, or 5 out of 7 correct forecasts in a tercile. Using a bootstrap (see Westerling et al 2002), results with this rate of correct forecasts were found to be significant at a 90% confidence level.

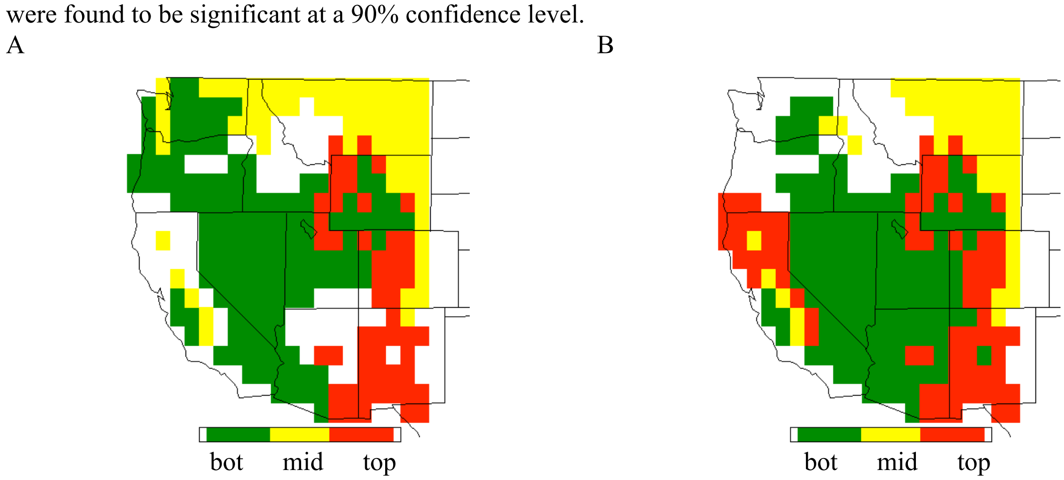

Figure 3: Area burned forecasts for 2003 fire season where March 2003 PDSI is estimated by (A) CCA model using Pacific SST and (B) persistence in January PDSI. Forecasts are presented as top, middle, or bottom terciles. Clear areas are where skill for the given tercile forecast is worse than random chance.

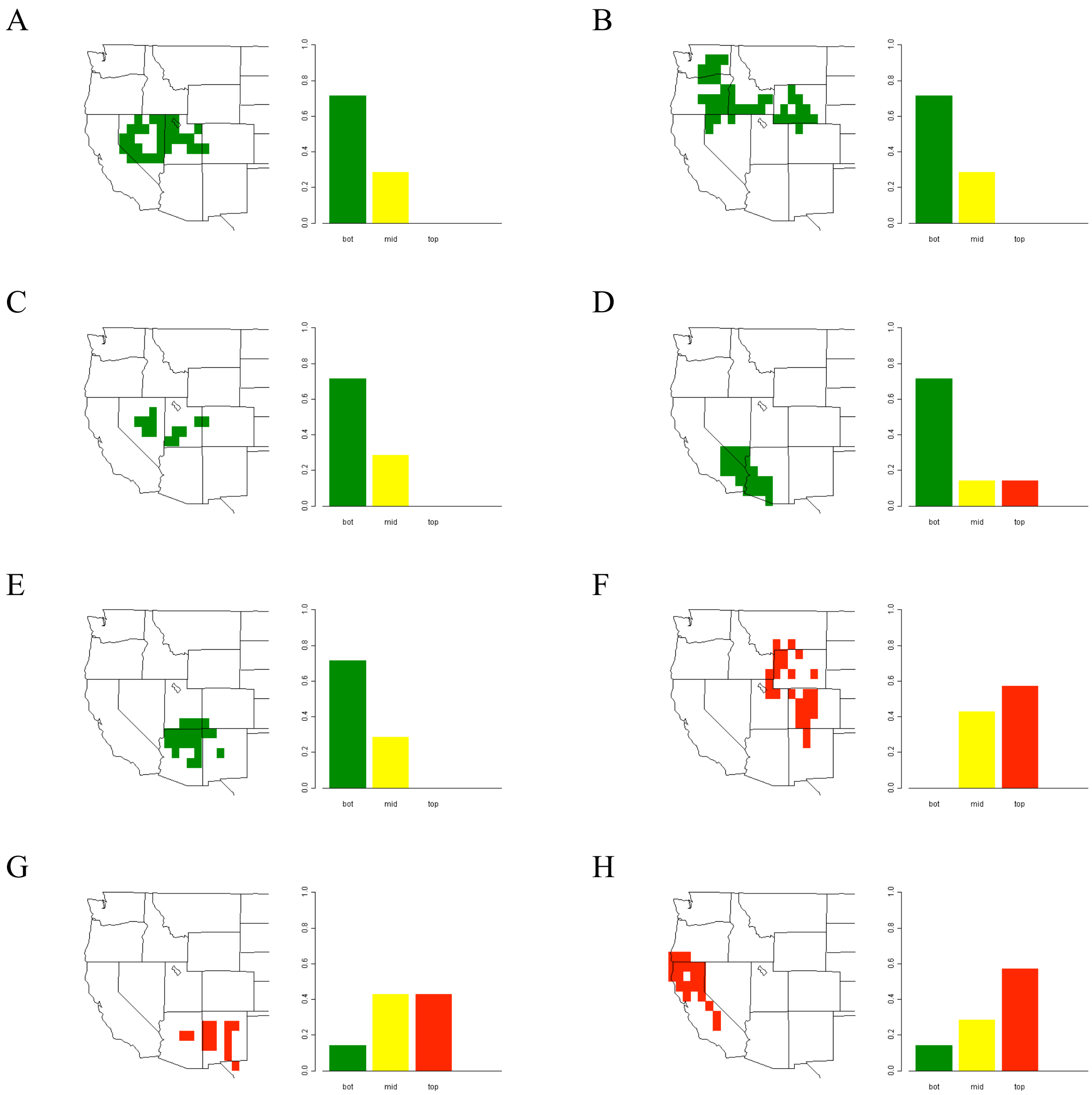

Figure 4: Panels A – D show probabilities of observing low, mid, and high (green, yellow and red, respectively) tercile area burned conditional on this years Persistence forecast of low area burned for four desert shrub and grassland ecosystem provinces in the Great Basin and the Southwest. Panel E shows the same thing for the Colorado Plateau, which, while containing more forested areas, shows the same positive correlation between area burned and anomalous moisture one year prior to the fire season as seen in the desert provinces (A – D). Panels F – H show conditional probabilities associated with this year's high forecast for the Southern Rockies, Sierra Nevada and mountains in Arizona and New Mexico.