J4.6 DEVELOPMENT OF MODEL OUPUT STATISTIC (MOS) PRODUCTS FOR PREDICTIVE SERVICES

Hauss J. Reinbold*, Timothy J. Brown, Beth L. Hall

Desert Research Institute, Reno, Nevada

1. INTRODUCTION

In 2004, Southern California Predictive Services developed the 7-day significant fire potential product, which by 2006 is anticipated to be operational nationwide. Necessary input for this product includes Remote Automated Weather Station (RAWS) point 7-day forecasts of fire weather, time-lag fuel moisture and National Fire Danger Rating System (NFDRS) indices for selected locations. In the first phase of the project, model output statistic (MOS) equations were developed for over 200 RAWS in California and Rocky Mountain Predictive Services areas using National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model output. In a second phase of the project currently underway, MOS equations are being developed for four more Predictive Services geographic areas. MOS output includes twice daily (00 and 12 UTC) tabular 10-day forecasts of fire weather and fire danger elements (e.g., temperature, energy release component, 100-hour fuel moisture).

After successful completion of the first phase, a second phase was implemented to create similar products for the Northern Rockies, Southwest, Western Great Basin and Eastern Area Predictive Services (Alaska is expected to be added in 2006). The overall project goal was to extract and add value to model forecasts from the National Weather Service for use by fire weather meteorologists and fire management. Primary objectives included 1) developing computing software that will extract relevant meteorological elements from numerical weather models; 2) performing a regression analysis and developing model output statistic (MOS) type equations that relates model output to a specific set of Remote Automatic Weather Stations (RAWS); and 3) developing and providing value-added products and information from the MOS equations. These objectives primarily support the 7-day significant fire potential product, but may be useful for other purposes. This paper describes the development of the MOS equations and their use in a value-added product for fire management.

*Corresponding author address: Hauss Reinbold, Desert Research Institute/DAS, 2215 Raggio Parkway, Reno, NV 89512; email: hauss.reinbold@dri.edu

2. DATA

The two primary data sets used in this project are the NCEP GFS model forecasts and RAWS observations. The spatial domain for all datasets covers the continental United States. All data sets are inclusive of the period from January 1, 2001 through December 31, 2003.

2.1 NCEP GFS FORECAST DATASET

Every six hours, NCEP produces GFS forecasts of atmospheric elements from 0-384 hours in 3-hour increments. The zero hour forecast, or initialization grid, of each 00z and 12z run (from January 1, 2001 through December 31, 2003) was evaluated. The spatial resolution of this forecast model changed in December 2002. From January 2001 through November 2002, the model output was produced on a lambert conformal projection in a 85x129 grid with a resolution of 81-km. From December 2002, through the end of 2003, the GFS forecasts we received were in global grids with one-degree resolution. For this project, only the grid nodes close to specific RAWS were utilized.

2.2 RAWS

Land and fire management agencies retain an observational network of RAWS for fire weather related measurements (see http://www.fs.fed.us/raws). They are typically located in wilderness, forest and rangeland areas where it is desired to monitor fire danger. The hourly observations are transmitted to the National Interagency Fire Center (NIFC) using a geostationary operational stationary satellite (GEOS) operated by the National Oceanic and Atmospheric Administration (NOAA). These data are forwarded to the Weather Information

Management System (WIMS) for agency use distribution, and to the Western Regional Climate Center (WRCC) for historical archiving. The RAWS data used in this study were obtained from WRCC. From this dataset, RAWS surface measurements of daily minimum temperature, daily maximum temperature, wind speed, daily minimum relative humidity, daily maximum relative humidity, were extracted and used in the analysis. WIMS provides site descriptors (fuel model, slope, climate class, etc.) for each RAWS, which, combined with the above atmospheric elements, allows for the calculation of fire danger indices (BI – burning index, IC – ignition component, ERC – energy release component, SC – spread component). Larry Bradshaw, U.S. Department of Agriculture Missoula Fire Sciences Laboratory, provided the original NFDRS computer software code for calculating the indices that was then adapted to fit the RAWS data format used in this project.

There were 278 RAWS sites total in the four Predictive Services areas for which the project has been completed (Northern California, Southern California, Rocky Mountain Area and Northern Rockies) so far that had sufficient quality data for the years 2001-2003. Sites were chosen based on three criteria regarding completeness of the data set – 1) no more than four months of missing data in two six-month periods, 2) availability of year-round operational measurements, and 3) availability of historical data for the period January 2001 through December 2003. Measurements that were clearly in error (e.g., relative humidity over 100 percent, negative wind speed) were considered missing and excluded from the analysis.

3. METHODS OF ANALYSIS

There were three main components to the analysis. The first was to extract and prepare the relevant model and RAWS data for the regression. The second was to perform the regression analysis for the forecast elements at each station. The third was the development of value-added products for each station and forecast.

A quality control (QC) analysis was performed on the RAWS data for stations that met the initial acceptance criteria described in section 2.2. The data were checked for suspicious values (spikes) through a visual inspection of the time series for each variable and RAWS site. Measurements that were clearly in error (e.g., relative humidity over 100 percent, negative wind speed) were considered missing and excluded from the analysis. The GFS forecast files did not require a similar QC process.

In order to compare the RAWS observations to the GFS forecast data, it was necessary to match the data sets spatially. With the GFS grids at one-degree resolution, the data from the four grid nodes surrounding each RAWS site were used in the regression analysis for each station. The 81-km resolution model grids were interpolated to one-degree regional grids for the purposes of this study.

The RAWS observations include the 1300 local time observations of temperature, humidity and wind speed needed for the calculation of the NFDRS indices, as well as the daily values for maximum and minimum temperature and relative humidity. These observations are compared to both 00z and 12z model values in the regression analysis. Data from 2001-2002 was used in the actual computation of the regression equations, and data from 2003 was used to verify the equations generated and in the computation of some bias estimates.

The regression analysis was computed by a statistical software package (S-plus). For each station, two sets of equations were generated, one based on 00z initialized GFS and the other based on 12z initialized GFS. A forward step-wise regression procedure was used to generate the regression equations for 12 elements at each RAWS: maximum and minimum temperature, maximum and minimum relative humidity (RH), wind direction, wind speed, burning index (BI), energy release component (ERC), ignition component (IC), spread component (SC), 100-hr fuel moisture and 1000-hr fuel moisture. For the Northern Rockies area (and future analysis areas) 10-hr fuel moisture was added as a forecast element.

In order to improve the accuracy of some forecast elements, persistence values were included as input to the regression analysis for maximum temperature, minimum RH, BI, ERC, IC, 100-hr fuel moisture, 1000-hr fuel moisture, and 10-hr fuel moisture. Also, some coastal meteorological gradients were included in the regression analysis for a few coastal stations in the California Predictive Service areas in order to increase the reliability of their regression models. In order to improve computational speed, predictor variables for the regression analysis was limited to lower level temperature, RH, and wind components. An example typical regression equation is:

MaxTemp = .1149*pMaxTemp + .5576*850Temp4 +.3321*700Temp2,

where pMaxTemp is the persistence maximum temperature observed at this RAWS the previous day, 850Temp4 is the 850 mb model temperature value of temperature at a neighboring grid cell, and 700Temp2 is the 700 mb model temperature value at another neighboring grid cell.

With the regression equations finally generated, operational forecasts are then made twice a day based on current runs of 00z and 12z GFS. Each forecast is made for ten days (though the potential product only requires seven days). Additional forecast elements include, maximum and minimum dew point (computed based on forecast temperature and RH) and the high, medium and low levels of the Haines Index (Haines, 1998) computed from model forecast values.

The third stage of the analysis involved creating several value-added products. After each forecast is completed, the departure from climatology is computed for maximum and minimum temperature, maximum and minimum RH, maximum and minimum Dew Point, and wind speed. The percentile rank based on climatology for each of these same forecast elements is also produced. For one area, the Rocky Mountains, a set of meteograms for each station forecast is also generated. These completed products are then made available to users via ftp and the CEFA web site.

4. ANALYSIS

Estimation of how well the regression equations may actually predict future values based on the same model data is analyzed in two ways. First, the R^2 value for each equation give an idea of just how good the regression model is for that variable. Values close to 1 imply a perfect forecast equation and values close to zero imply a poor regression model. Second, after running forecasts for a certain amount of time, one can simply compare to actual measurements for the same time period.

The quality of the regression equations as indicated by the R^2 values varies from element to element. Those indicators that rely on persistence generally have very high R^2 values. For instance, the R^2 values for maximum temperature are mostly in the .9 to .95 range (Figure 1). The NFDRS elements that rely on persistence also have very high R^2 values (Figures 2 and 3). RH R^2 values are much more variable (Figure 4), usually demonstrating values anywhere from .5 to .8, although min RH includes persistence and thus has higher R^2 values than max RH overall. R^2 values for wind are usually below .5. R^2 Values for the remaining forecast elements such as maximum RH (Figure 5), minimum temperature (Figure 6), and IC (Figure 7) are on average lower and more variable than their counterparts as the figures show. The forecasts for these elements should therefore be trusted a little less. Extensive validation work has not been done on the BI, IC and SC equations.

A short comparison of the maximum temperature forecast for one station in the Rocky Mountain Predictive Services area to the observed values for that station showed positive results (Figure 8). One should remember that the accuracy of the GFS model forecast will also affect the accuracy of the MOS forecasts.

5. DISCUSSION AND CONCLUSION

The goal of creating and delivering MOS forecasts has been successfully completed so far. High priority elements such as ERC, maximum temperature, minimum RH and 100- and 1000-hour fuel moisture appear to have high forecast skill. These forecasts provide users with important forecasts for use in managing public lands with respect to fire danger. Southern California Predictive Services incorporates these forecasts into a table indicating significant fire potential for various sub-regions within their area (http://www.fs.fed.us/r5/fire/south/fwx/Fire_Potential.html). The raw text forecast products and meteograms are available at http://cefa.dri.edu/Operational_Products/MOS/txtmosfcsts.php.

6. ACKNOWLEDGEMENTS

This work was supported by the National Predictive Services Group and the Bureau of Land Management national Office of Fire and Aviation under cooperative Assistance Agreement number 1422RAA000002. We very much appreciate the efforts of Tom Rolinski (Southern California Predictive Services, Tom Wordell (National Interagency Fire Center), Terry Marsha (Pacific Northwest Predictive Services) and Larry Bradshaw (Missoula Fire Sciences Laboratory) for their various important contributions to this project.

7. REFERENCES

Haines, D. A., 1988: *A Lower Atmospheric Severity Index for Wildland Fire*. National Weather Digest, **13**, No. 2:23-27.

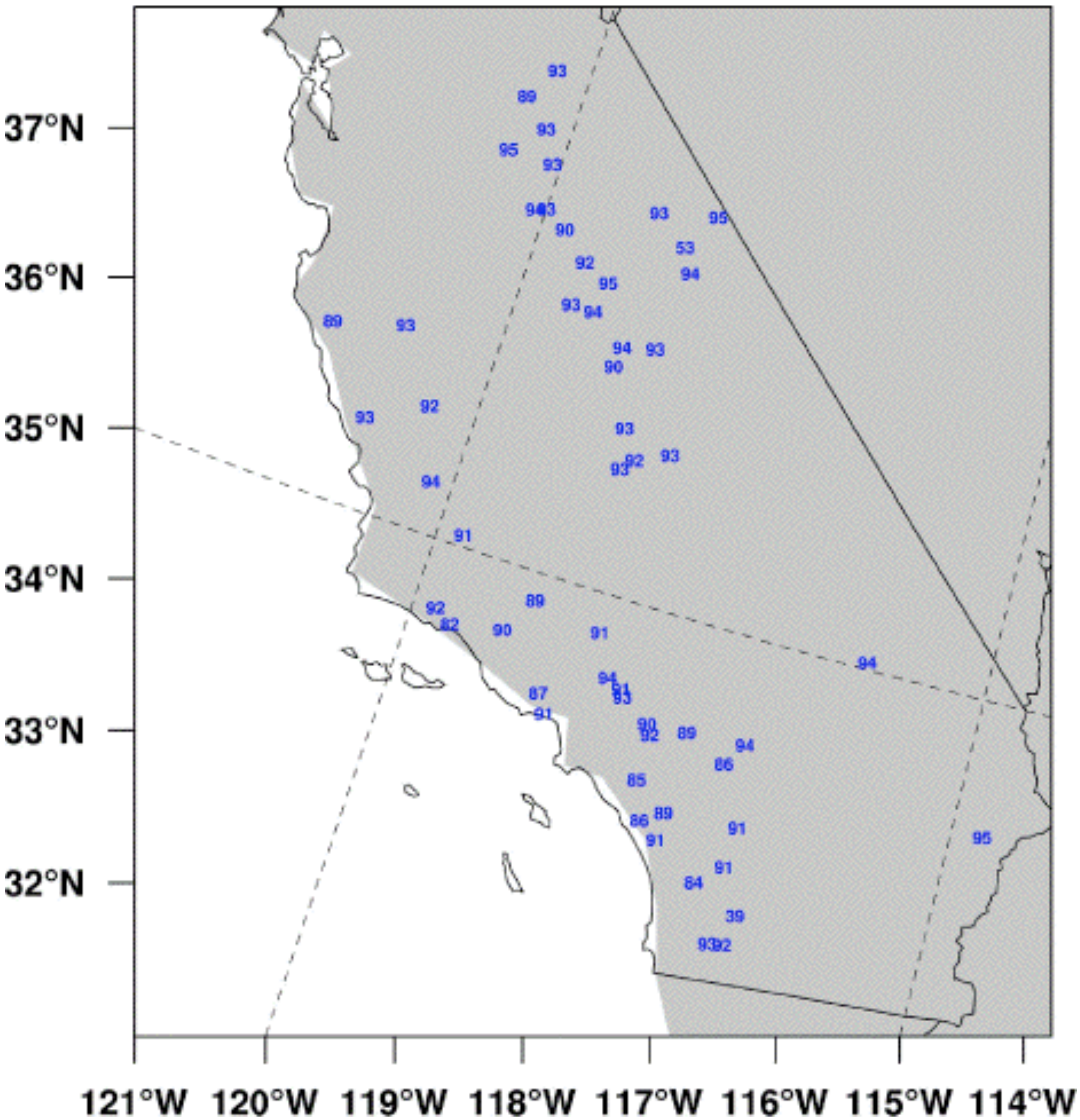

Figure 1. Southern California Predictive Services area R^2 values for maximum temperature.

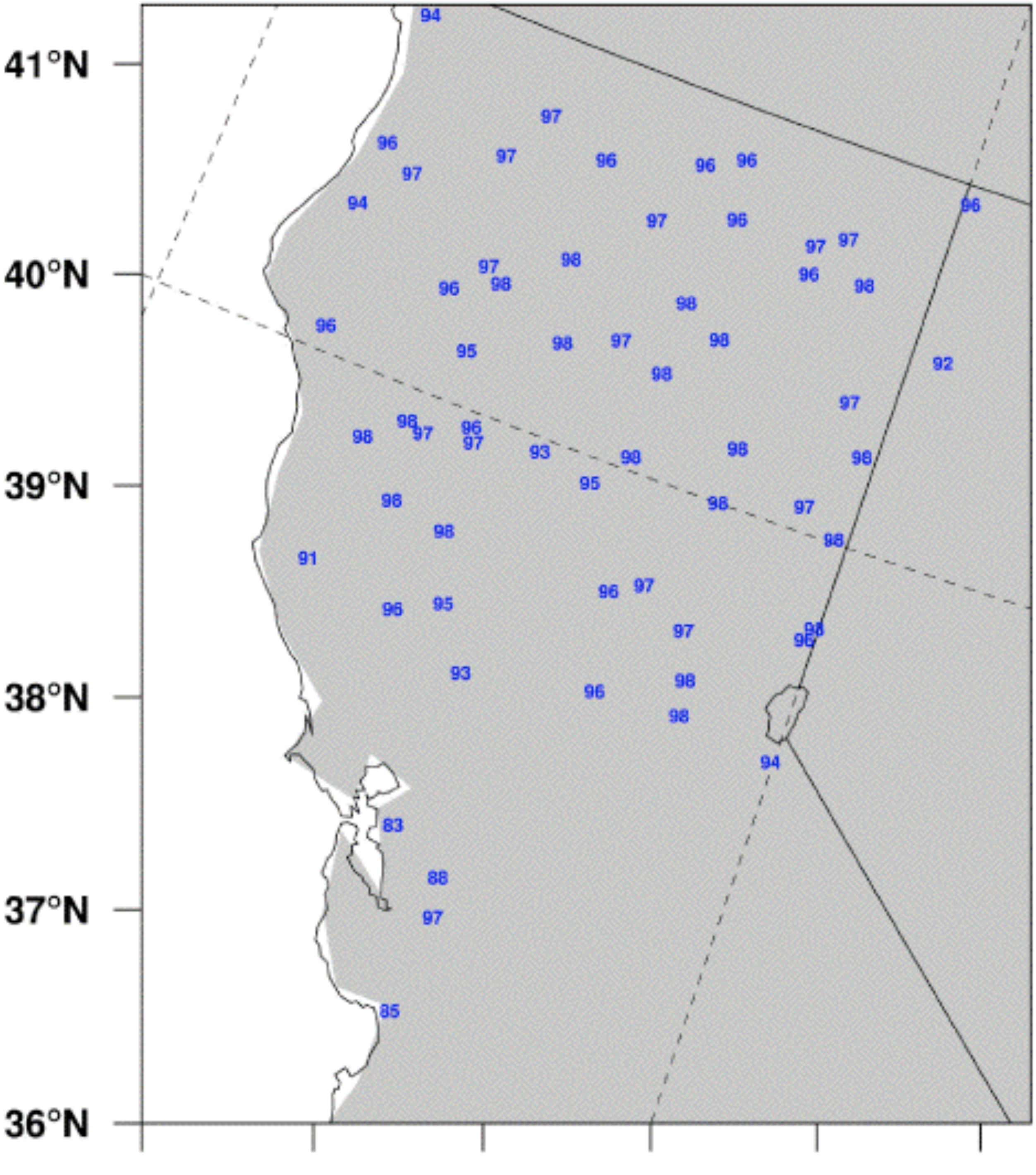

Figure 2. Northern California Predictive Services area R^2 values for maximum temperature.

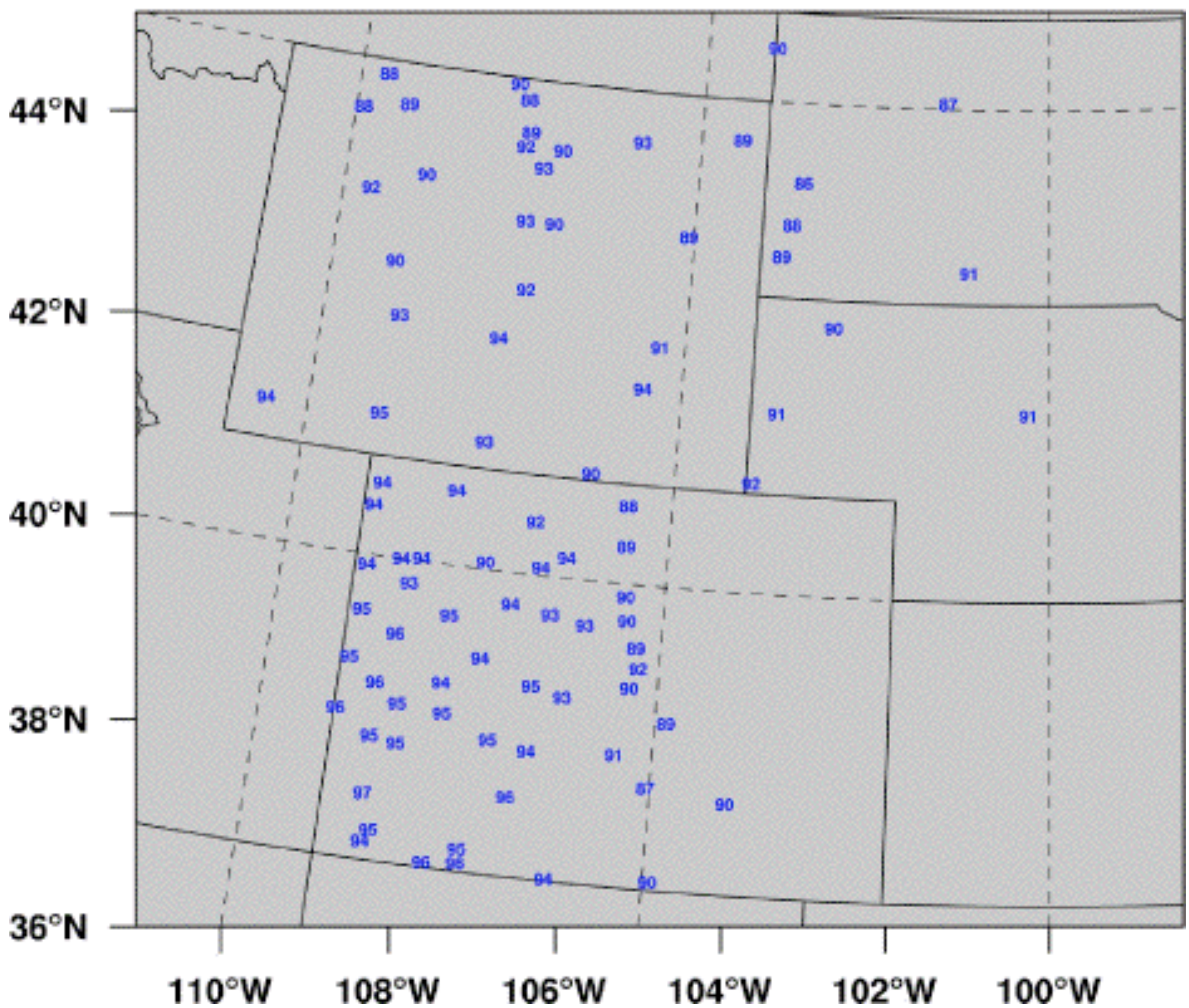

Figure 3. Rocky Mountain Predictive Services area R^2 values for 100-hr fuel moisture

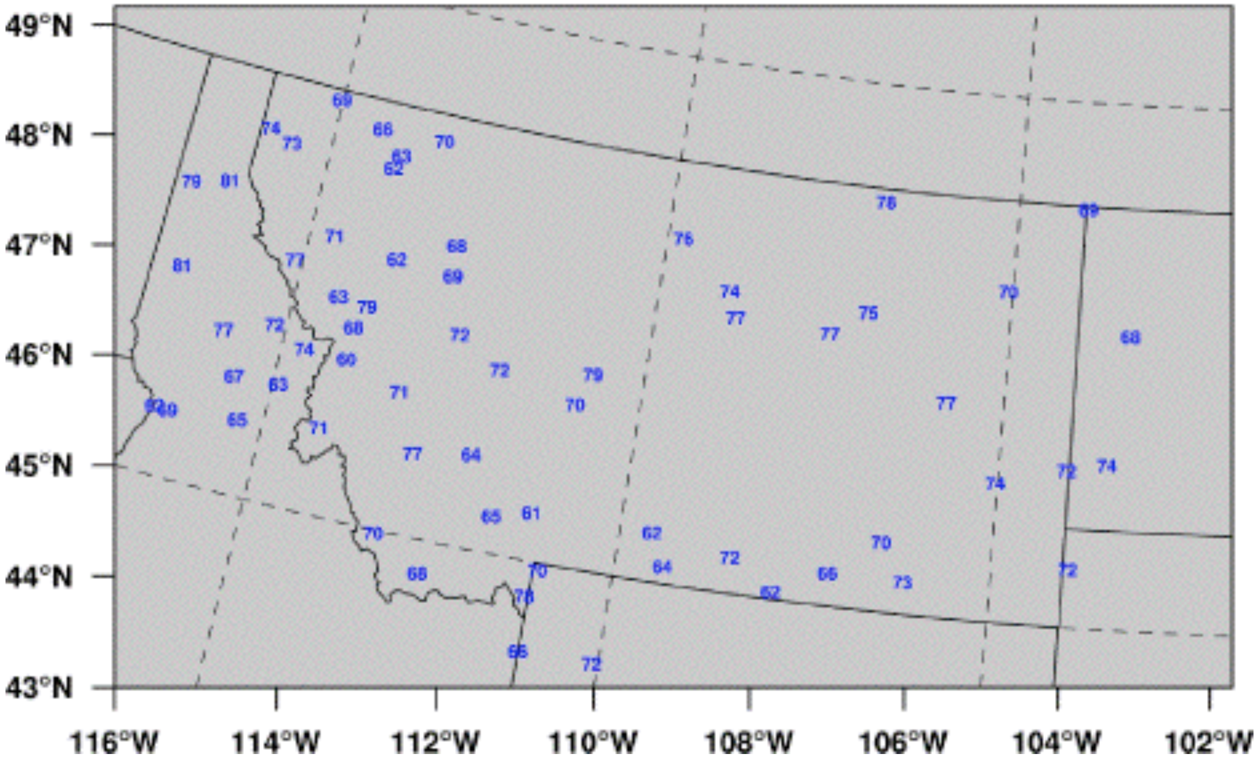

Figure 4. Northern Rockies Predictive Services area R^2 values for minimum RH

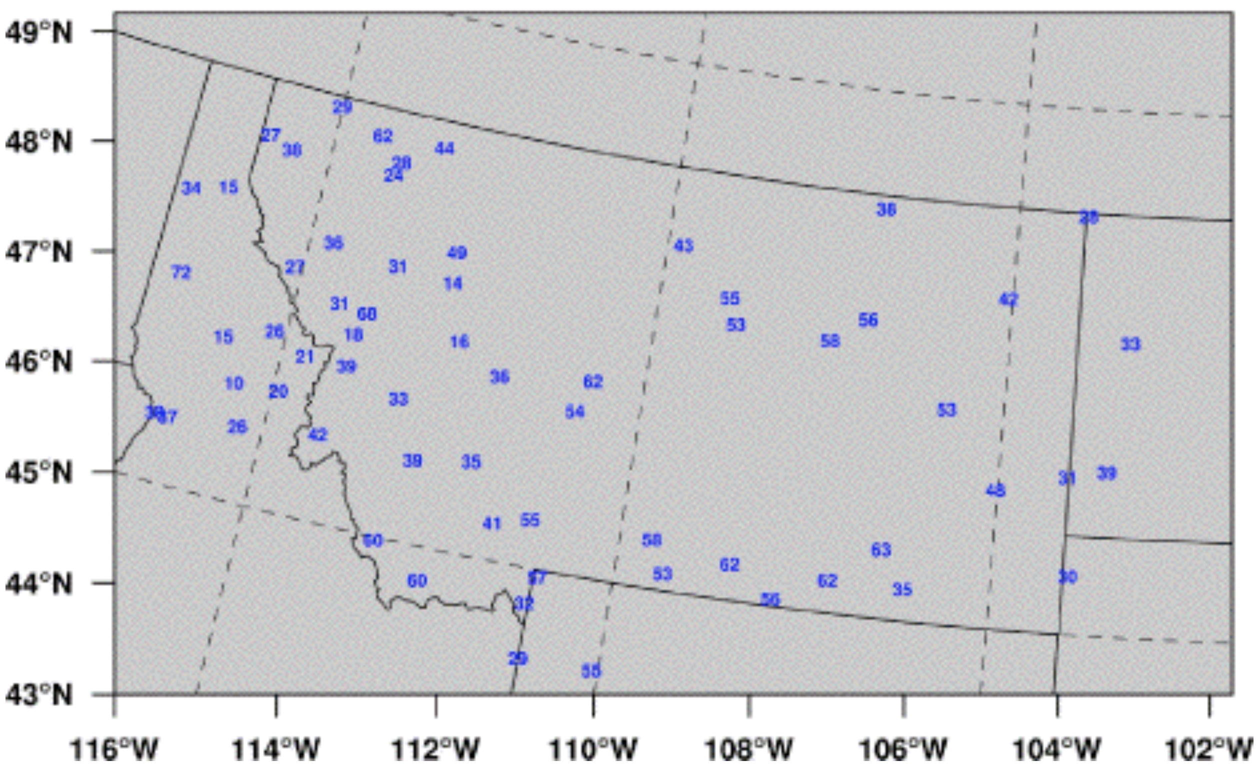

Figure 5. Northern Rockies Predictive Services area R^2 values for maximum RH

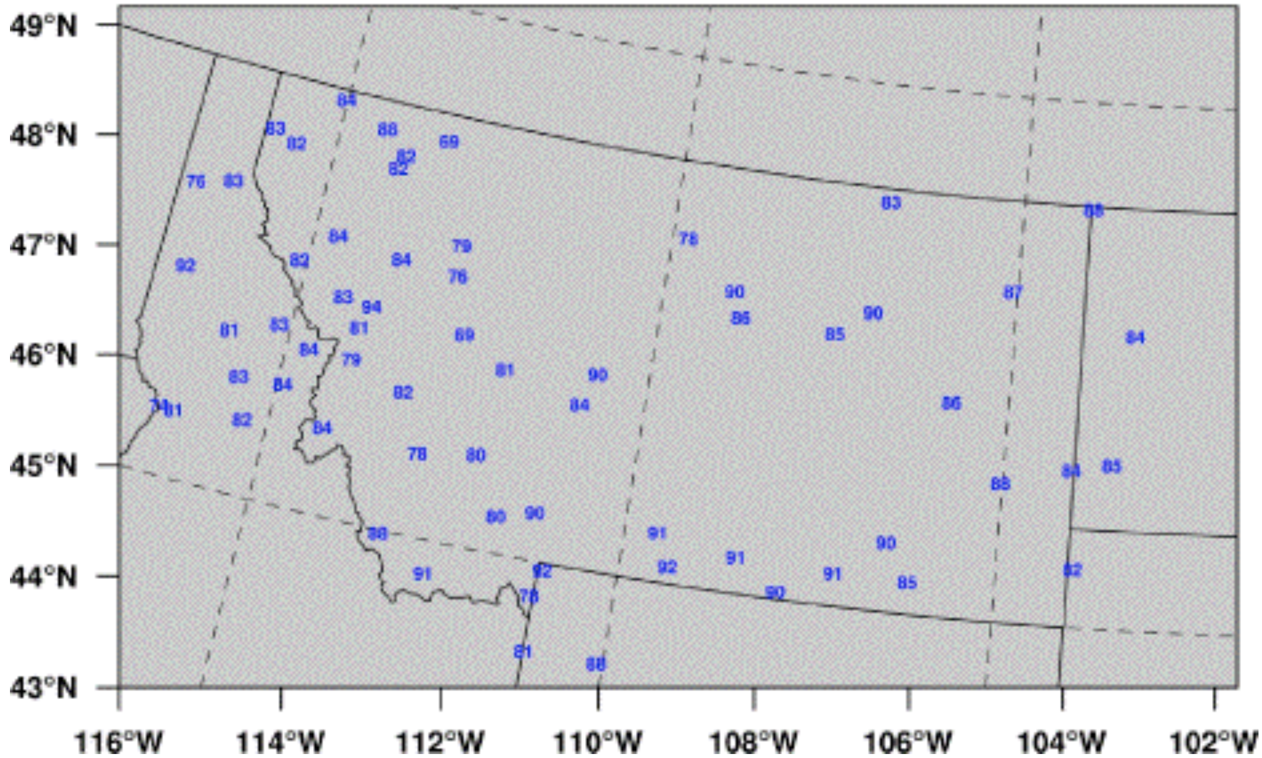

Figure 6. Northern Rockies Predictive Services area R^2 values for minimum temperature

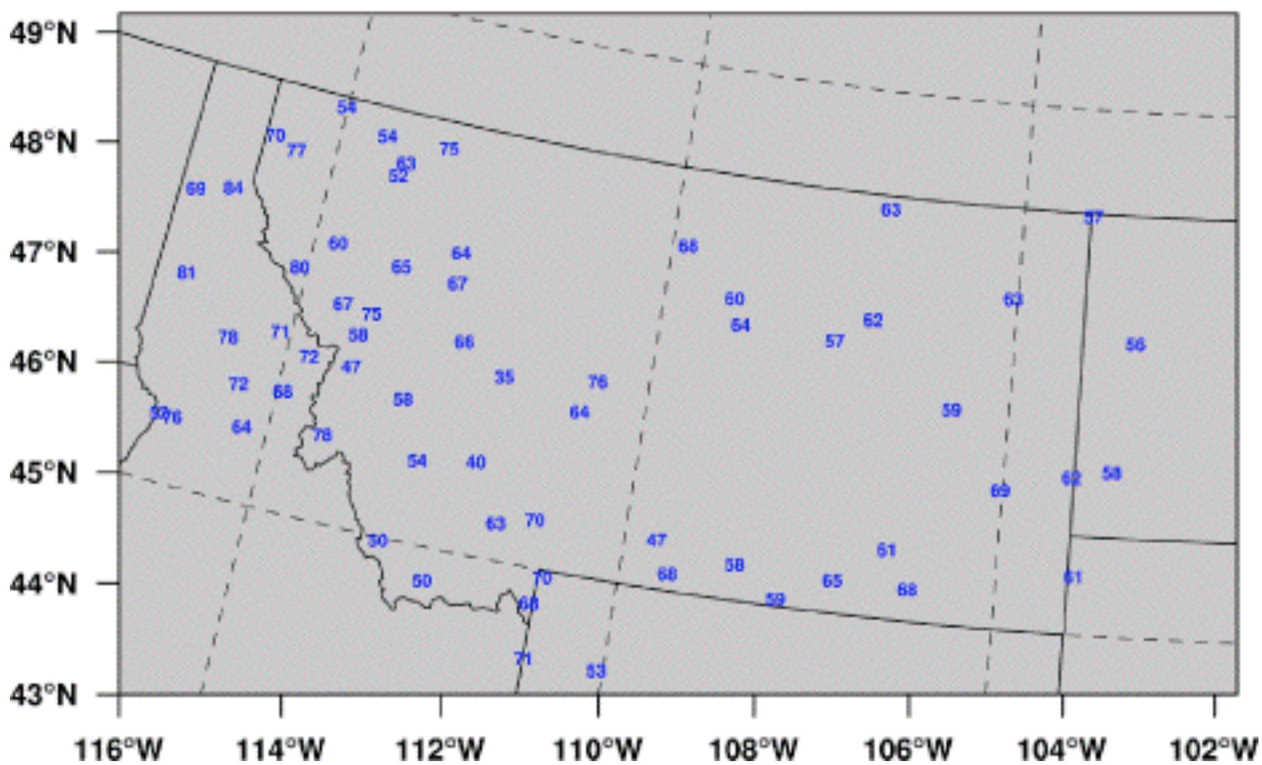

Figure 7. Northern Rockies Predictive Services area R^2 values for IC

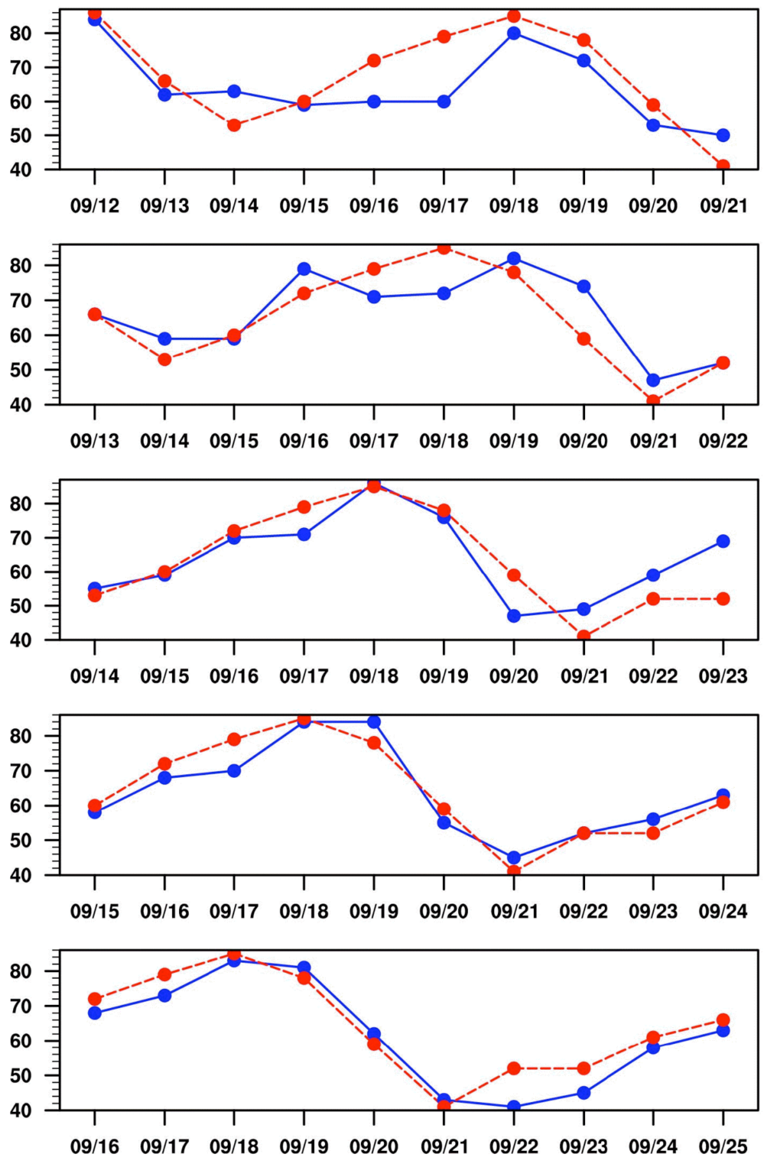

Figure 8. Example comparison of MOS forecast (blue) and RAWS observations (red) for a station in the Rocky Mountains.