6.3 SPATIAL/TEMPORAL CHARACTERISTICS OF AIR QUALITY PREDICTION ERRORS IN THE BLUESKY MODELING FRAMEWORK

K.-S. LII, S. CHEN, S. HUANG
University of California, Riverside

F.M. FUJIOKA*
Forest Fire Laboratory, Pacific Southwest Research Station, USDA Forest Service

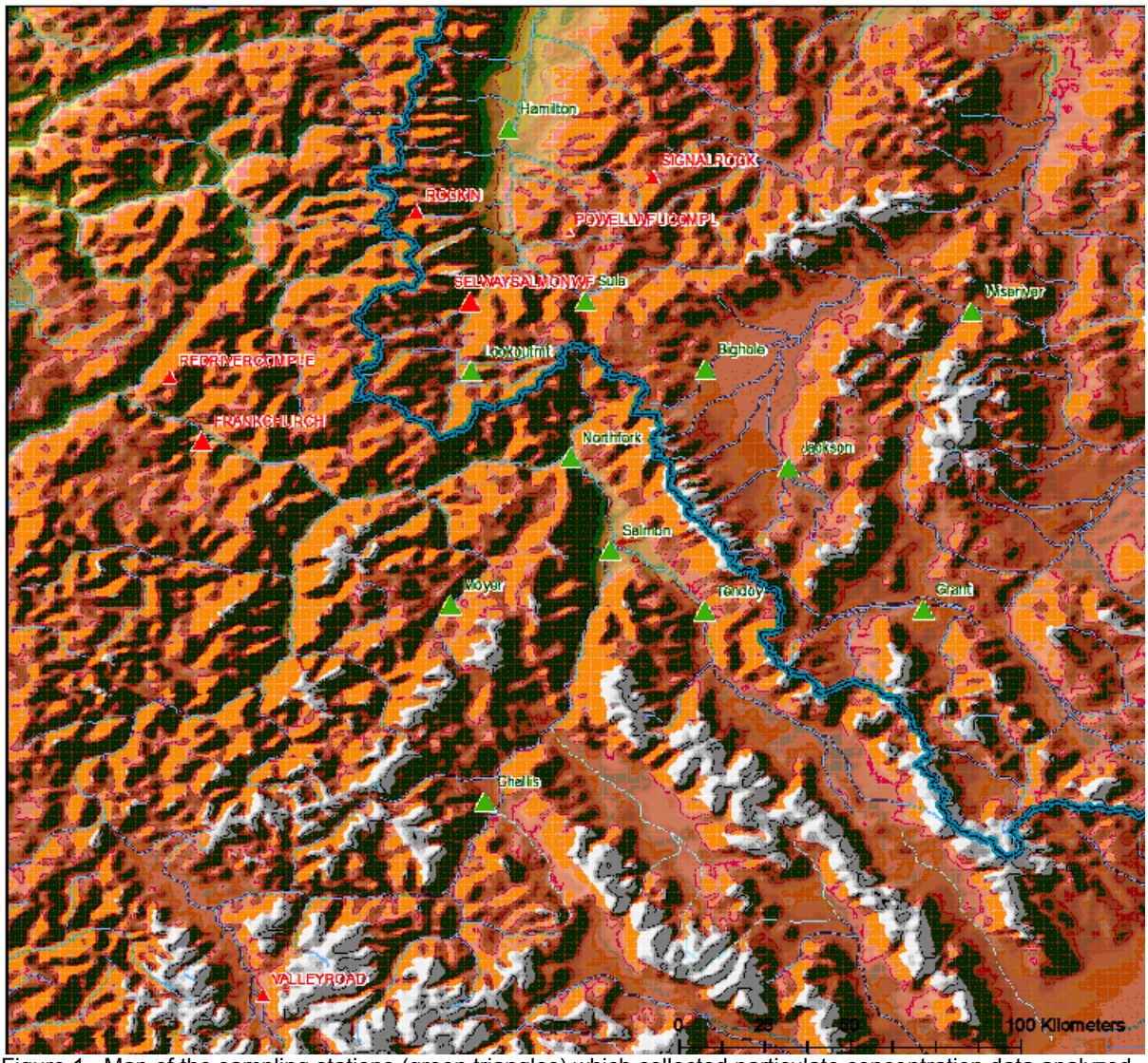

Figure 1. Map of the sampling stations (green triangles) which collected particulate concentration data analyzed in this study. Red triangles locate wildland fires active during the study period.

1. INTRODUCTION

Since the National Fire Plan funded the Fire Consortia for Advanced Modeling of Meteorology and Smoke (FCAMMS) in 2001, more wildland fire managers around the country have experienced the BlueSky modeling framework of weather and air quality prediction.

Building on the foundation of the BlueSky modeling consortium in the Pacific Northwest (O'Neill et al. 2003), the FCAMMS generate high resolution weather forecasts from a mesoscale model, to use in air quality models that predict smoke trajectories and particulate concentration fields for wildland fires of given size and emission characteristics.

The modeling system invites attention of fire managers who must first protect the public and resources from wildfires, then operate in compliance with federal, state and local smoke

*Corresponding author address: Francis M. Fujioka, USDA Forest Service, 4955 Canyon Crest Dr., Riverside, CA 92507; e-mail: ffujioka@fs.fed.us.

management regulations (USDA Forest Service Fire and Aviation Management Briefing Paper on Wildland Fire and Air Quality, 23 March 2006). Especially when they use prescribed fire, managers must choose environmental conditions under which planned operations will not violate air quality standards. Weather and air quality models in BlueSky were assembled to simulate the complex processes that govern production and dispersion of smoke from wildland fire, hence to be a predictive tool for fire planning.

At the behest of the US Environmental Protection Agency, a BlueSky demonstration project was planned for 2005 in partnership with the federal land management agencies, to determine the feasibility of operating a westwide BlueSky modeling system, and to conduct a limited analysis of the system's performance (BlueSkyRAINS West (BSRW) Demonstration Project, Final Report 2006, available at www.airfire.org/pubs). Fires in national forests in eastern Idaho and western Montana (Figure 1) were the objects of one of two field trials conducted to evaluate the accuracy of the BlueSky simulations. This paper describes an analysis of the error characteristics of the BSRW predictions of the PM2.5 concentration fields, beyond what was covered in the final report. The next sections cover the BlueSky modeling framework, the setup of the BSRW project, and the method of analysis of this study. We conclude with a summary and recommendations.

2. BLUESKY MODELING FRAMEWORK

The BlueSky modeling framework was designed to predict meteorological and particulate concentration fields, and smoke trajectories emanating from active fires (O'Neill et al. 2003). It uses the MM5 mesoscale meteorological model to generate high resolution weather fields (Grell et al. 1994), which in turn drive the air quality models. For the 11 western states covered by the BSRW domain, the Forest Service Rocky Mountain Center ran MM5 twice daily to produce the required weather elements at hourly intervals on a 12 km horizontal grid spacing for a 48-hour period.

Given the location and size of active fires in the BSRW domain, BlueSky can determine the emission rates for particulates and other pollutants not considered in this study. Fire data for the BSRW project came from ICS-209 reports, fuels data from the Fuel Characteristics Classification System, and emissions from application of the CONSUME and Emissions Production Models (BlueSkyRAINS West (BSRW) Demonstration Project, Final Report 2006).

Dispersion is simulated by CALPUFF (Scire et al. 2000), a model that simulates Gaussian dispersion from a point, area, or volume (Figure 2), and estimates plume rise from the fire characteristics determined in the preceding step. Finally, the HYSPLIT model (Draxler and Hess 1998) generates smoke trajectories which track the smoke from each fire, given the meteorological conditions.

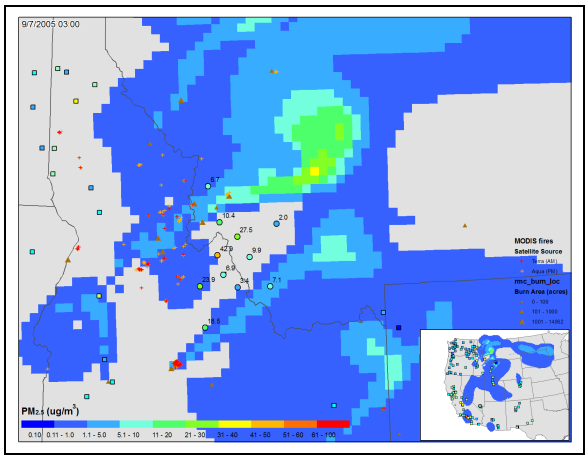

Figure 2. Simulated PM2.5 concentration field for 7 September 2005, 2100 MDT, over Idaho and Montana. Inset shows the entire BSRW domain. Figure provided by Sonoma Technologies.

3. BSRW FIELD TRIAL IN IDAHO/MONTANA

The individual models BlueSky uses have significant histories, but the BlueSky modeling framework was still relatively new in 2005 and had not been thoroughly tested. The BSRW project funded limited field trials for fires in the Gila National Forest of New Mexico, the Salmon-Challis National Forests of Idaho and the Bitteroot National Forest of Montana (Figure 1). The Rocky Mountain Center (RMC) generated 48-hour BlueSky forecasts for the Idaho/Montana fires from 3-16 September 2005, twice daily at 0600 MDT and 1800 MDT.

A team from the Forest Service Pacific Northwest Research Station (PNW) deployed particulate concentration monitors from which a database of hourly PM2.5 concentrations was compiled by a contractor. This study analyzed the errors in the RMC BlueSky PM2.5 forecasts using data from 12 monitoring stations operated in Idaho and Montana over the period 4-15 September 2005 (Table 1). Note that not all stations operated at the same time, and the number of observations varied between stations. PNW also tasked the contractor to analyze the accuracy of the PM2.5 simulations, which was done independently of the simulation error analysis in this paper.

Table 1. Monitoring stations in BSRW project. Times are Mountain Standard.

STATION	LAT (Deg Min N)	LON (Deg Min W)	START Date/Time	END Date/Time	Eq 1 R SQ
Hamilton	46* 16.287'	114* 9.591'	09/04 0100	09/13 1800	0.0044
Salmon	45* 9.779'	113* 53.527'	09/04 2000	09/15 2300	0.0002
Moyer	45* 1.263'	114* 18.723'	09/05 0800	09/11 2000	0.0000
Challis	44* 30.246'	114* 13.303'	09/04 0100	09/14 0400	0.0006
Grant	45* .542'	113* 4.020'	09/05 2300	09/14 0400	0.0507
Jackson	45* 22.974'	113* 25.411'	09/06 0100	09/09 1400	0.0050
North Fork	45* 24.481'	113* 59.797'	09/04 2200	09/15 2300	0.0000
Sula	45* 49.180'	113* 57.414'	09/05 0000	09/13 0700	0.0000
Wise River	45* 47.457'	112* 56.386'	09/06 0100	09/09 2000	0.0273
Big Hole	45* 38.310'	113* 38.557'	09/06 0500	09/14 2300	0.0000
Tendoy	45* .254'	113* 38.662'	09/05 0000	09/12 0400	0.0150
Lookout Mt	45* 38.088'	114* 15.485'	09/07 2000	09/15 1000	0.0203

4. DATA AND METHOD OF ANALYSIS

The principal objective of this study is to describe the spatial and temporal characteristics of the differences between observed and simulated PM2.5 concentrations. Let *o(s,t)* represent the PM2.5 concentration observed at location s and time t, and *f(s,t,k)* the forecasted concentration at (s,t) determined by the k-th BSRW forecast period. Each of s, t, and k can be an index number, where each s corresponds to the spatial coordinates for each observation site, t corresponds to an observation date and time, and k indexes one of four consecutive 12-hour periods issued for every forecast. Each forecast therefore has a maximum of 48 hours of predicted concentration values for each observation site. However, no monitor recorded an observation for every forecast hour, i.e. the observation times comprise a subset of the forecast times. Table 1 lists the station coordinates in terms of latitude and longitude but we used coordinates relative to a Lambert conformal conic projection in the spatial analysis. The concentration data are expressed in μg/m^3.

We choose to characterize the relationship between the forecast value and the observed value by a simple regression formula:

$$o(s,t) = af(s,t,k) + b + \varepsilon(s,t) \qquad (1)$$

Assume for now that ε is an independent, identically distributed random variable, therefore the regression constants a and b are solved by ordinary least squares (Draper and Smith 1981). The variable k equals 1 for the initial to 11th forecast hours, k equals 2 for the 12th to 23rd hours, etc. In this setting, the correlation coefficient expresses the strength of the linear association between the observation and the forecast. If the correlation coefficient is high, the regression can be used to calibrate the forecast, so that a mismatch between forecast and observation involving a scaling factor and an offset can be statistically corrected. In fact, the instruments used to measure particulate concentrations in situ are similarly calibrated to match measurements from a Federal Reference Method gravimetric sampler (Trent 2006).

5. RESULTS

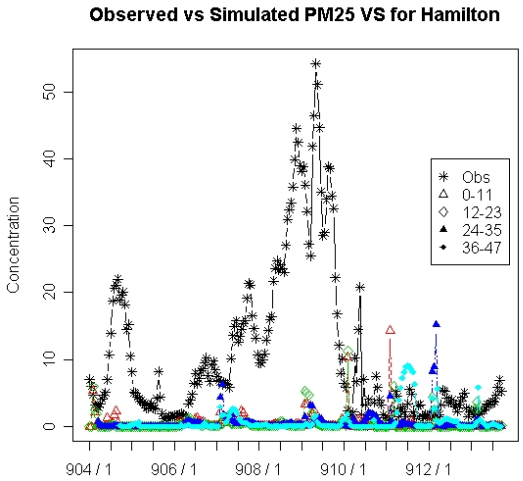

Figure 3. Time series of observed and simulated PM2.5 concentrations for Hamilton.

5.1 *Initial comparison of time trends*

We initially plotted time series pairs of the predicted and observed PM2.5 concentrations at each of the locations in Table 1, which made apparent a tendency of the simulations to underestimate the observed PM2.5 concentrations, particularly for elevated concentration levels. For example, the PM2.5 hourly concentrations (μg/m^3) observed at Hamilton averaged 10.5, whereas the simulated

Hamilton PM2.5 values averaged a mere 0.46. Observed concentrations at Hamilton exceeded 10 for a sustained period between the afternoon of 8 September through all of the following day (Fig. 3). While spikes exceeding 35 were recorded during this time, the simulated values were relatively flat by comparison.

5.2 *Regression analysis*

We know from regression theory that the square of the correlation coefficient is a useful measure of the capacity of the least-squares solution of Equation 1 to explain the variability in the observed concentrations at each location. If o-hat is the predictor derived from Equation 1 and o-bar is the sample mean, r^2 represents the proportion of the observation sum of squares that is accounted for by the regression equation (Draper and Smith 1981):

$$r^2 = \sum_i \left(\hat{o}_i - \bar{o}\right)^2 / \sum_i \left(o_i - \bar{o}\right)^2$$

The last column of Table 1 shows that none of the regressions calculated from the data does well in this regard. The zero entries in the column indicate that the computed r^2 was less than 0.0001.

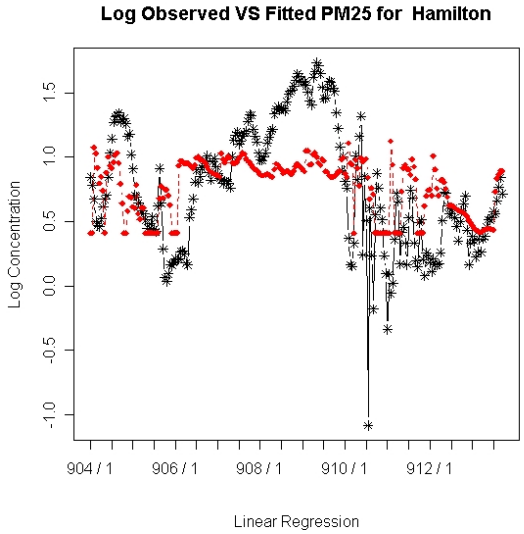

Figure 4. Time series plots of observed log concentrations (black asterisks) and predicted log concentrations (red diamonds) at Hamilton.

For Hamilton, the simulated series for the first 12 hours of each forecast peaked twice in roughly the same period of high observed concentrations, but at an order of magnitude less than the latter. This suggests that a log transformation on both the observed and simulated concentration values might improve the correlation between them. The non-negativity constraint on the log function requires the addition of a small positive value to the concentration data (notated by tilde), which we set arbitrarily to 10^{-5}. We would then modify the regression as in Equation 2:

$$\log \tilde{o}(s,t) = a \log \tilde{f}(s,t,k) + b + \varepsilon(s,t) \qquad (2)$$

The transformation did improve the correlation, and the r^2 score jumped from 0.0044 under Equation 1 to 0.2034 under Equation 2. This would be a highly significant result under the standard assumptions made for the distribution of the residuals, but it would be difficult to declare that the regression fit the observations well, with an r^2 that shows only 20% of the variance explained. The time series of the observed log concentrations and the corresponding regression estimated log concentrations in Figure 4 indicate that the observed values vary with greater amplitude than the predicted values.

Another alternative we pursued to match the observed trend with a statistical model was nonparametric regression, which does not restrict the relationship between dependent and independent variables in the regression to a linear form, and incorporates local fitting (Cleveland et al. 1992). We used a nonparametric model of degree 1 nevertheless, with gaussian smoothing, which yielded an r^2 of 0.2380, adjusted for the equivalent degrees of freedom in the nonparametric fit. We repeated the nonparametric fit after performing a log transform on the data, as we had in the simple regression case. The transform increased the adjusted r^2, but only to 0.2559.

5.3 *Comparison by forecast lead times*

So far, we have only analyzed the performance of the first 12 hours of the forecasts, k=1 in Equation 1 (lead times of 0 to 11 hours). It is apparent that the forecasts tend to underestimate the observations by an order of magnitude, which can be ameliorated somewhat by a log transformation of the data. We have not yet applied regression analysis to the other forecast periods, but we have begun to explore the distributional properties of the log difference between forecasts and observations for all k. We make use of the relationship:

$$\log\left(\frac{\tilde{o}(s,t)}{\tilde{f}(s,t,k)}\right) = \log \tilde{o}(s,t) - \log \tilde{f}(s,t,k)$$

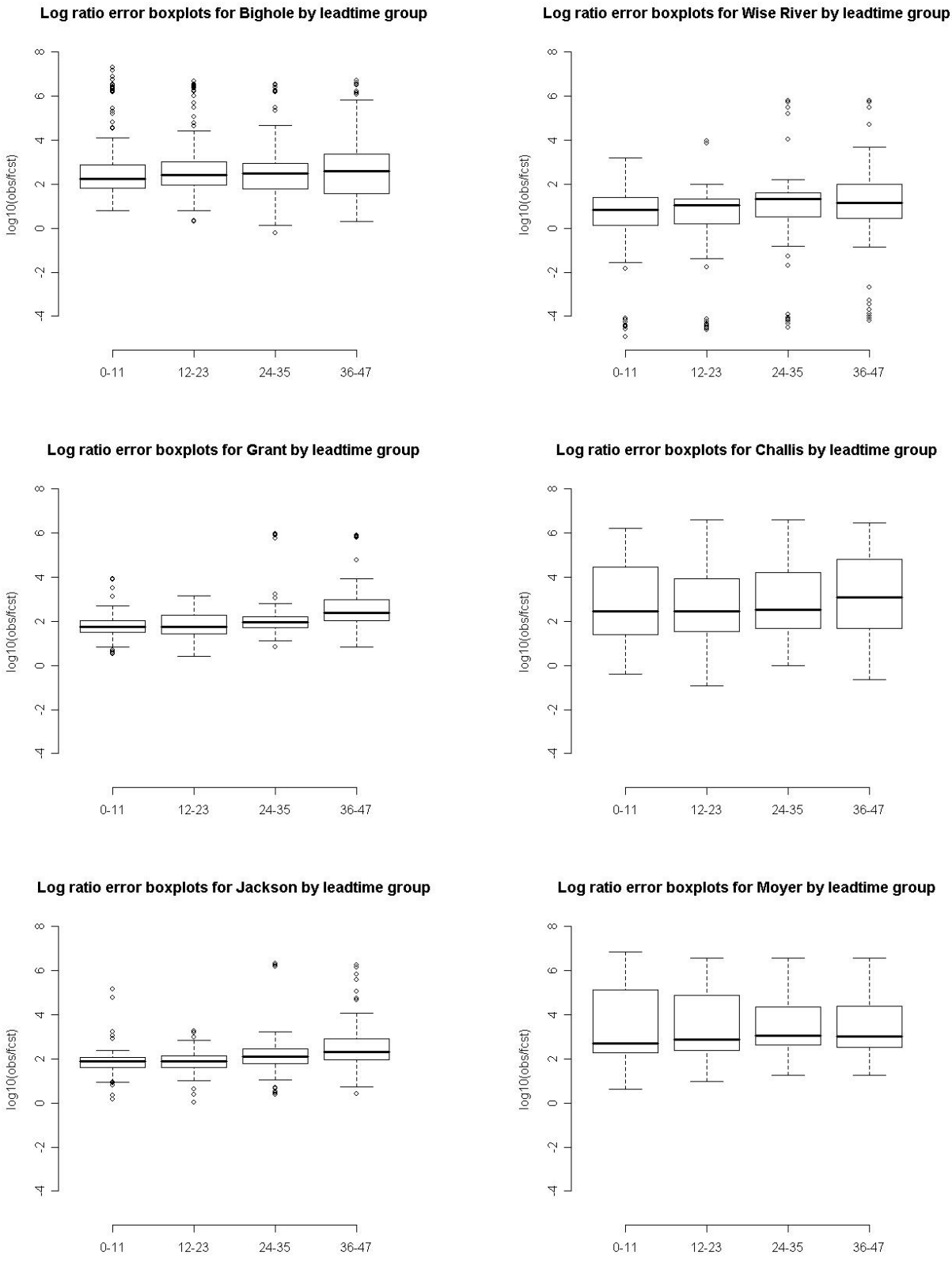

Figure 5. Boxplots of the log10 transform of the observed to forecast concentration ratio for the first six stations in Table 1.

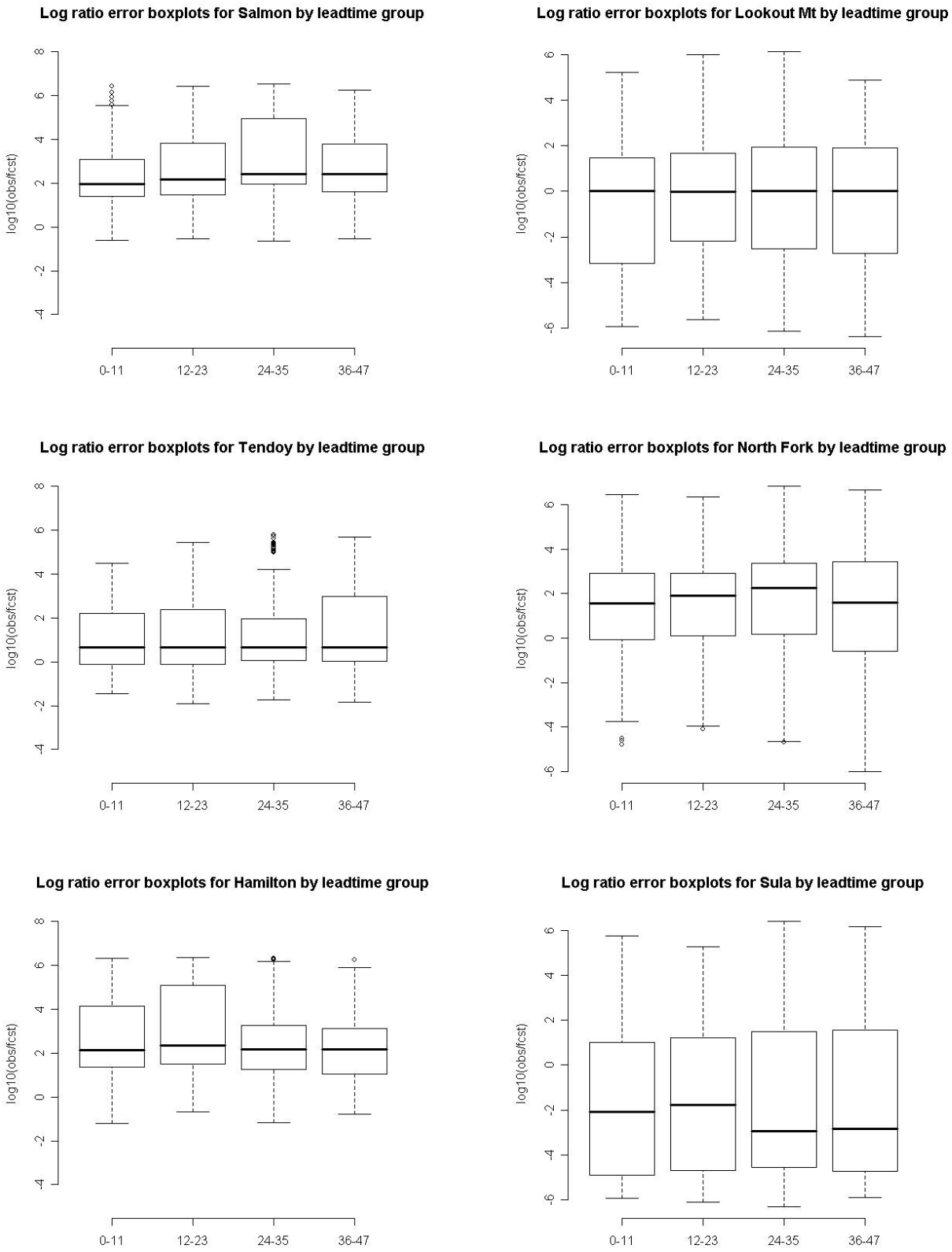

Figure 6. Boxplots of the log10 transform of the observed to forecast concentration ratio for the second six stations in Table 1.

Note in the boxplots that the log transforms are base 10, so the differences between observed and forecast concentrations were often substantial. Most of the time, the forecasts underestimated the observed concentrations. Sula was the exception. Bighole, Jackson, and Wise River had the most outliers. Shorter lead times did not always result in greater accuracy or precision. Hamilton provided the best counterexample to this rule. There, the interquartile range of the log ratio error (Figure 6) shrank noticeably for the 24-35 and 36-47 hour lead time periods. Otherwise, the distribution of the log ratio error overall was fairly regular between lead time groups at each station.

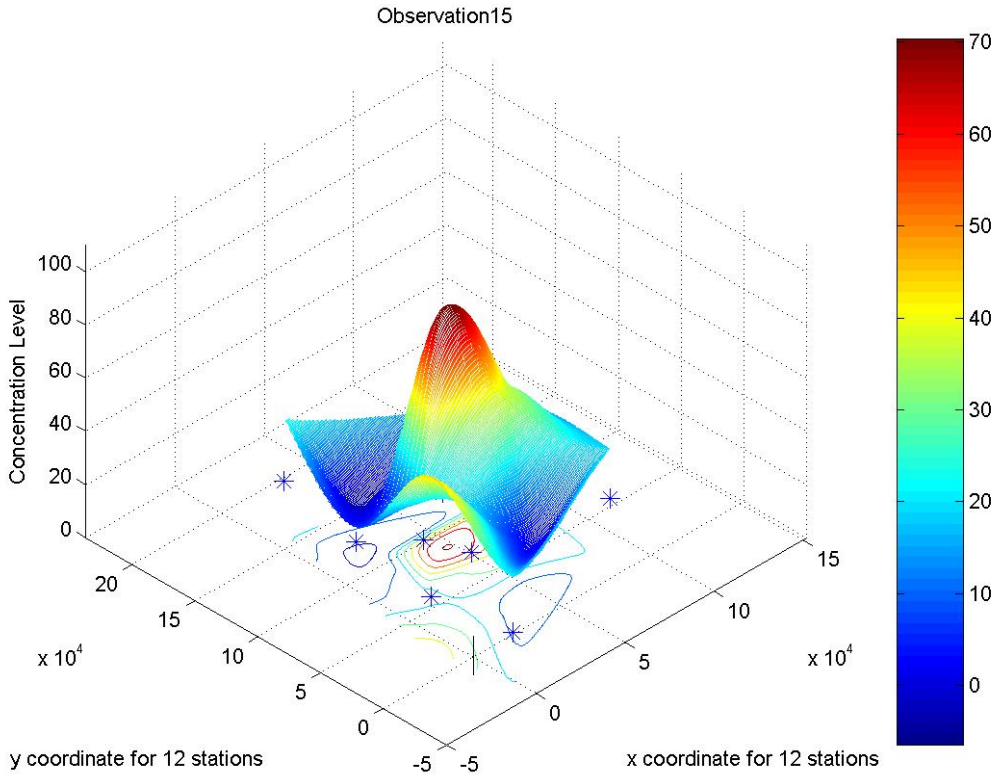

Figure 7. PM2.5 concentration field estimated from station data by application of a geostatistical model. Asterisks in the x-y plane mark the positions of stations not hidden by the surface. The y-axis aligns northerly in Lambert conformal conic coordinates, and the x-axis increases eastward.

6. SUMMARY AND RECOMMENDATIONS

The 2005 BSRW Demonstration Project provided limited data for a case study of the accuracy of PM2.5 dispersion simulations from fires near the Idaho-Montana border. The analysis showed that the simulations grossly underestimated the observed PM2.5 concentrations. An attempt to calibrate the simulations using regression techniques to statistically adjust the simulations achieved marginal success.

Because the simulated values were generally an order of magnitude smaller than the observed concentrations, a log transform was applied to the data, which improved the regression fit, but not substantially. The design of the BSRW Project provided limited opportunities to determine weaknesses in the models. The final report from the Project cited deficiencies in simulating the boundary layer collapse at the end of the day and the fire emissions (BlueSkyRAINS West (BSRW) Demonstration Project, Final Report 2006). Clearly, the simulations require improvement.

What is also needed is an analysis of the simulations *between* stations. Studies to do this are ongoing. We have initiated an attempt to create a spatially and temporally continuous picture of the PM2.5 concentration field using geostatistical methods with the concentrations observed over the sampling network. A first step in universal kriging (Cressie 1991) is to determine the spatial trend surface (Figure 7).

The same methods can and should be used to determine the spatial/temporal variability of the simulation errors that have been so far analyzed only pointwise at the sampling stations.

REFERENCES

Cleveland, W.S., E. Grosse and W.M. Shyu, 1992: Local regression models. *Statistical Models in S*, J.M. Chambers and T.J. Hastie (eds.), Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 309-376.

Cressie, N.A.C., 1991: *Statistics for Spatial Data*. John Wiley, New York, 900p.

Draper, N. and H. Smith, 1981: *Applied Regression Analysis, 2nd edition*. John Wiley and Sons, New York, 1-55.

Draxler, R.R. and G.D. Hess, 1998: An overview of the HYSPLIT_4 modelling system for trajectories, dispersion, and deposition. *Austral. Met. Mag.*, 47:295-308.

Grell, G., A., J. Dudhia, D. R. Stauffer, 1994: A description of the fifth generation Pennsylvania State Unversity/NCAR mesoscale model. NCAR/TN-398+IA, National Center for Atmospheric Research, Boulder, Co, 116p.

O'Neill, S.M., S.A. Ferguson, J. Peterson, and R. Wilson, 2003. The BlueSky smoke modeling framework. Proc., 5th Symp. on Fire and Forest Meteorol., 16-20 Nov. 2003, Orlando, FL. 4 p.

Scire, J.S., D.G. Strimaitis, and R.J. Yamartino, 2000: A user's guide for the CALPUFF dispersion model (version 5). Tech. Rep., Earth Tech Inc., Concord, MA. 521 p.

Trent, A., 2006: Smoke particulate monitors: 2006 update. Tech. Rep. 0625-2842-MTDC. USDA Forest Service, Missoula Technology and Development Center, 14p.