Program for Climate, Ecosystem and Fire Applications

Development of Operational Air Quality MM5-MOS Equations for the San Joaquin Valley Air Pollution Control District

Final Report

Hauss J. Reinbold
Timothy J. Brown

Desert Research Institute
Division of Atmospheric Sciences

CEFA Report 10-01

March 2010

Development of Operational Air Quality MM5-MOS Equations for the San Joaquin Valley Air Pollution Control District

Final Report

by

Hauss J. Reinbold and Timothy J. Brown
Program for Climate, Ecosystem and Fire Applications
Desert Research Institute, Reno, Nevada

March 2010

Table of Contents

EXECUTIVE SUMMARY

This report describes the data, analysis and results of developing MOS equations of PM_{10}, $PM_{2.5}$, 1-hour ozone maxima and 8-hour ozone maxima for several stations relevant to the San Joaquin Valley Air Pollution Control District (SJVAPCD). MM5 meteorology output from the California and Nevada Smoke and Air Committee (CANSAC) twice-daily forecast operations provided the predictors for the equations. Validation results based on R^2 goodness-of-fit suggests that there is predictability for most of the stations. Thus, this product can be useful operational forecast guidance. Experimental forecasts for SJVAPCD are performed twice daily for a 3-day forecast period. The product is available online at:
http://cefa.dri.edu/Operational_Products/MOS/txtmosfcsts.php.

1. Introduction and objective

The Desert Research Institute (DRI) program for Climate, Ecosystem and Fire Applications (CEFA) maintains an operational, mesoscale meteorology forecast facility for smoke and fire management under the guidance and support of the California and Nevada Smoke and Air Committee (CANSAC). CANSAC comprises several federal, state and district agencies including the San Joaquin Valley Air Pollution Control District (SJVAPCD). The CANSAC operational facility currently provides high-spatial resolution meteorological information and value-added products for fire weather and smoke management. Three-hourly forecasts out to 72 hours are available twice daily on grids encompassing all of California and Nevada at 36-, 12- and 4-km spatial resolution. Output from these forecasts can be utilized stand-alone or for value-added products, such as the air quality experimental forecasts described in this report.

The objective of this project was to develop model output statistic (MOS) type equations based on historical MM5 output, and utilize the equations operationally to produce forecasts of PM_{10}, $PM_{2.5}$ and ozone. These experimental forecasts can serve as air quality forecast guidance. The project consisted of two primary tasks: 1) Developing MOS equations based upon MM5 archived output; and 2) Implementing the equations operationally to provide guidance forecasts for SJVAPCD. The project deliverables consist of the web based delivery of the operational forecasts, and this report. The project was a collaborative effort between CEFA and meteorologists at SJVAPCD.

2. Data used in the analysis

2.1 Air quality observations

SJVAPCD supplied measurements of PM2.5, PM10, 1-hr ozone and 8-hr ozone as retrieved from the Environmental Protection Agency's Air Quality System (AQS) repository for 29 monitoring sites in the San Joaquin Valley from June 2004 through August 2006. PM_{10} and $PM_{2.5}$ were supplied as both hourly and daily average observations in parts per million (*ppm*), while ozone observations consisted of hourly measurements, and 1-hour and 8-hour daily maximum values in units of units of microgram per cubic meter (μ/m^3).

2.2 MM5 model output

The meteorological model used for the regression analysis was the 4-km resolution CANSAC MM5 from June 2004 through August 2006. The CANSAC MM5 employs the Lambert Conformal map projection centered at 38°N, 121°W, and consists of three nested grids. The outermost grid (36-km horizontal resolution, 97x97x32 grid cells) covers the western U.S., parts of Mexico/Canada and the eastern Pacific. The nested grid (12-km horizontal resolution, 154x154x32 grid cells) encompasses California, Nevada, Oregon, Utah and parts of Idaho, Arizona, Wyoming and Montana. The innermost grid (4-km horizontal resolution, 274x274x32 grid cells) encapsulates the entire California and Nevada boundaries. Figure 1 shows the CANSAC MM5 domains. The specific model elements used in the regression analysis to create the MOS prediction equations were surface and upper air temperature, *u* and *v* wind vectors, perturbation pressure and the water vapor mixing ratio. These variables were originally suggested by meteorologists at SJVAPCD.

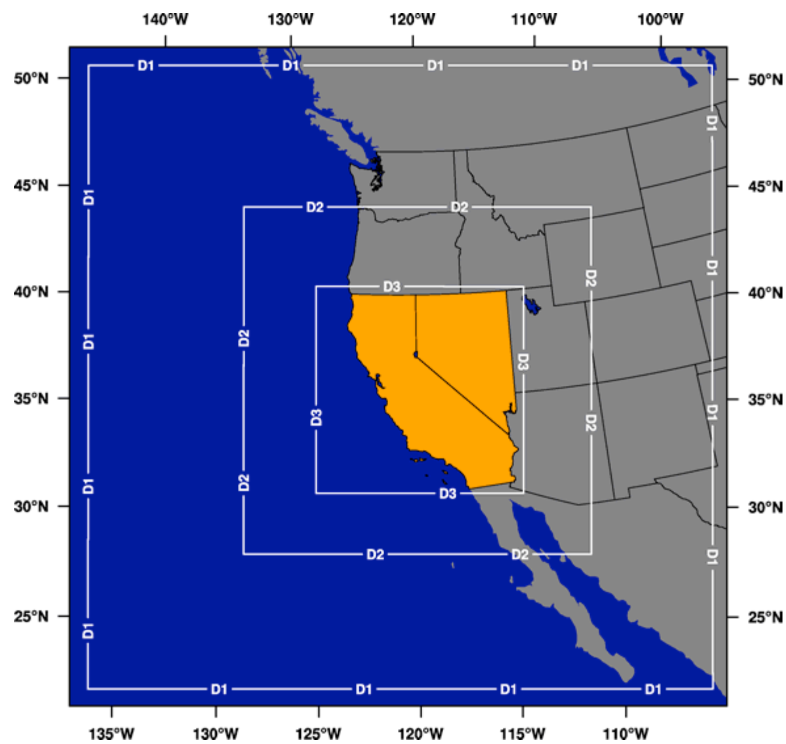

Figure 1. Domains for the CANSAC MM5.

3. Regression analysis

The purpose of the regression analysis was to establish forecast equations for daily averages of PM_{10} and $PM_{2.5}$, and the daily maxima of 1-hour and 8-hour ozone based on the MM5 model forecast elements of surface and upper air temperature, U and V wind vectors, perturbation pressure and the water vapor mixing ratio used as predictors. The first step was to process the AQS observational data into daily averages of PM_{10} and $PM_{2.5}$ and daily maxima of 1-hour and 8-hour ozone. These observations were then paired with model data for the four MM5 grid points closest to each monitoring site. For the regression equation development, the daily air quality values were matched with the MM5 initializations at 00z and 12z separately, so regression equations for each model initialization time could be generated. Regression equations for each air quality measurement at each station for 00z and 12z were then produced using forward stepwise regression analysis from the SPLUS statistical software package.

Statistical validation of each equation was assessed by examining the goodness-of-fit R^2 coefficient; higher R^2 values indicate a better fit. An objective of stepwise regression is maximize the goodness-of-fit measures, but not overfit the equation with too many variables and those having colinearity (highly correlated to each other). Once the equations were developed and tweaked to achieve the maximum R^2 they were evaluated over a few weeks by SJVAPCD meteorologists. Feedback from this assessment allowed for a few more minor equation adjustments. Figures 2 and 3 show R^2 values for the forecast variables of daily average PM_{10} and $PM_{2.5}$, and daily maxima 1-hour ozone and 8-hour ozone for the 12Z forecast time. For most stations the PM_{10} R^2 values are generally high. The R^2 values for PM_{10} generally are larger than for $PM_{2.5}$.

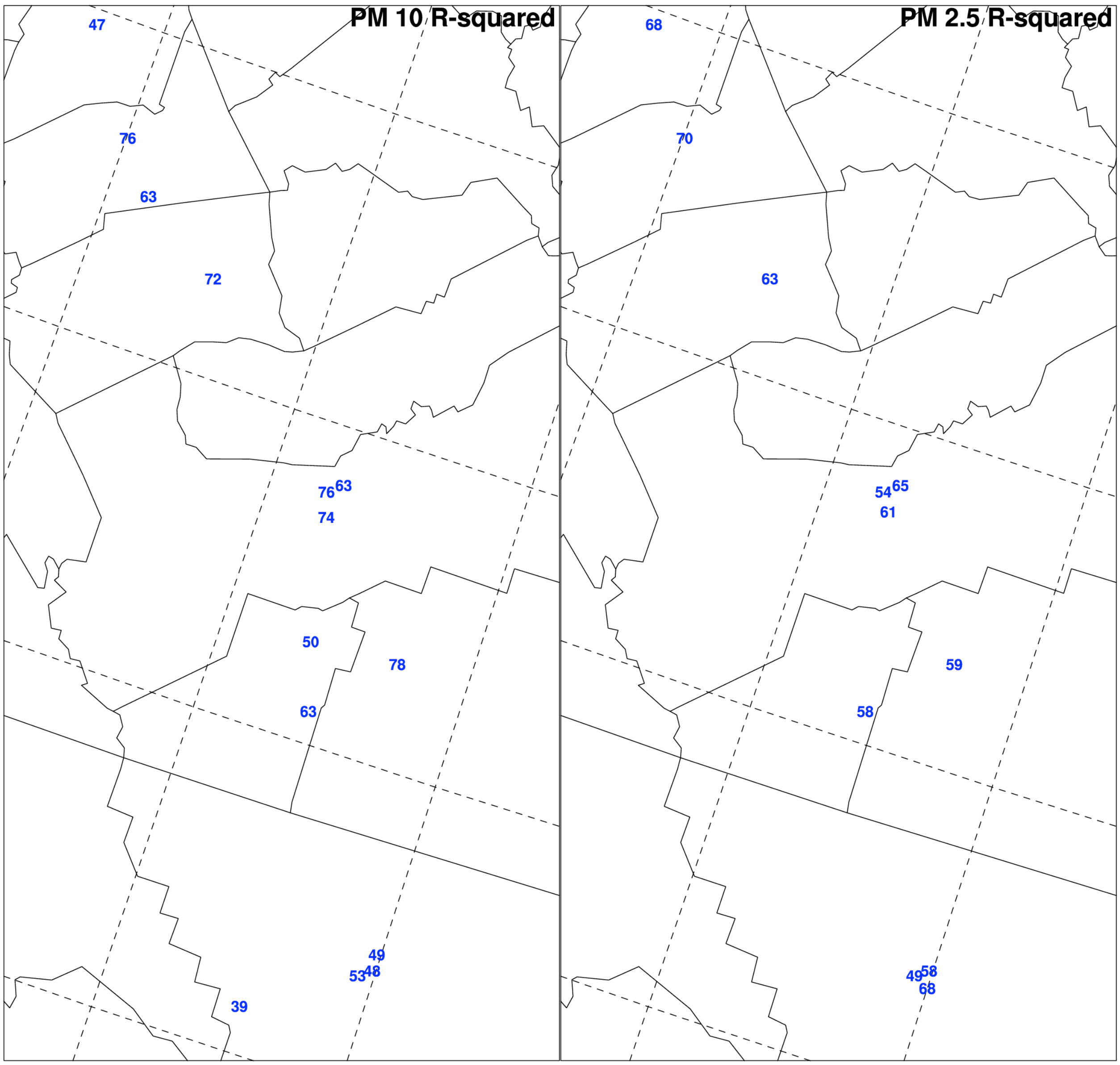

Figure 2. R^2 goodness-of-fit values for each station's MOS equation for PM_{10} and $PM_{2.5}$ for the 12Z forecast.

The R^2 values for ozone tend overall to be higher than for PM (Figure 3). The 1-hour ozone daily maxima generally have higher R^2 values than the 8-hour ozone daily maxima. The R^2 values for both PM and ozone suggest some predictability in the equations for providing air quality forecast guidance.

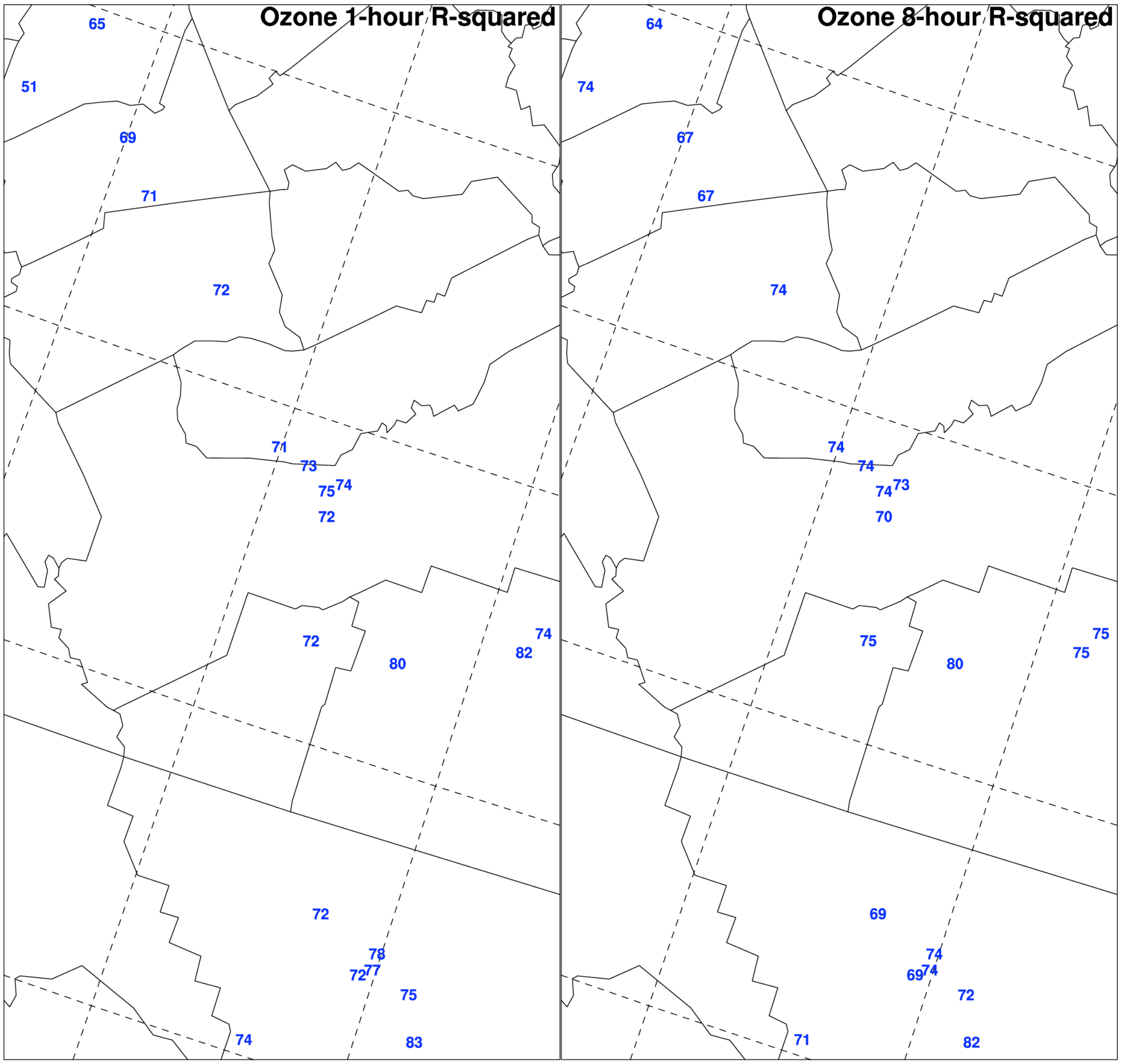

Figure 3. R^2 goodness-of-fit values for each station's MOS equation for ozone 1-hour daily maxima and ozone 8-hour daily maxima for the 12Z forecast.

Figures 4 and 5 show similar plots except for the 00Z forecast time. Though somewhat dependent on the station, in general the PM R^2 values tend to be higher than for 12Z. The same pattern generally applies for the ozone 00Z forecasts. In both cases this suggests that the afternoon forecasts tend to be slightly better than morning.

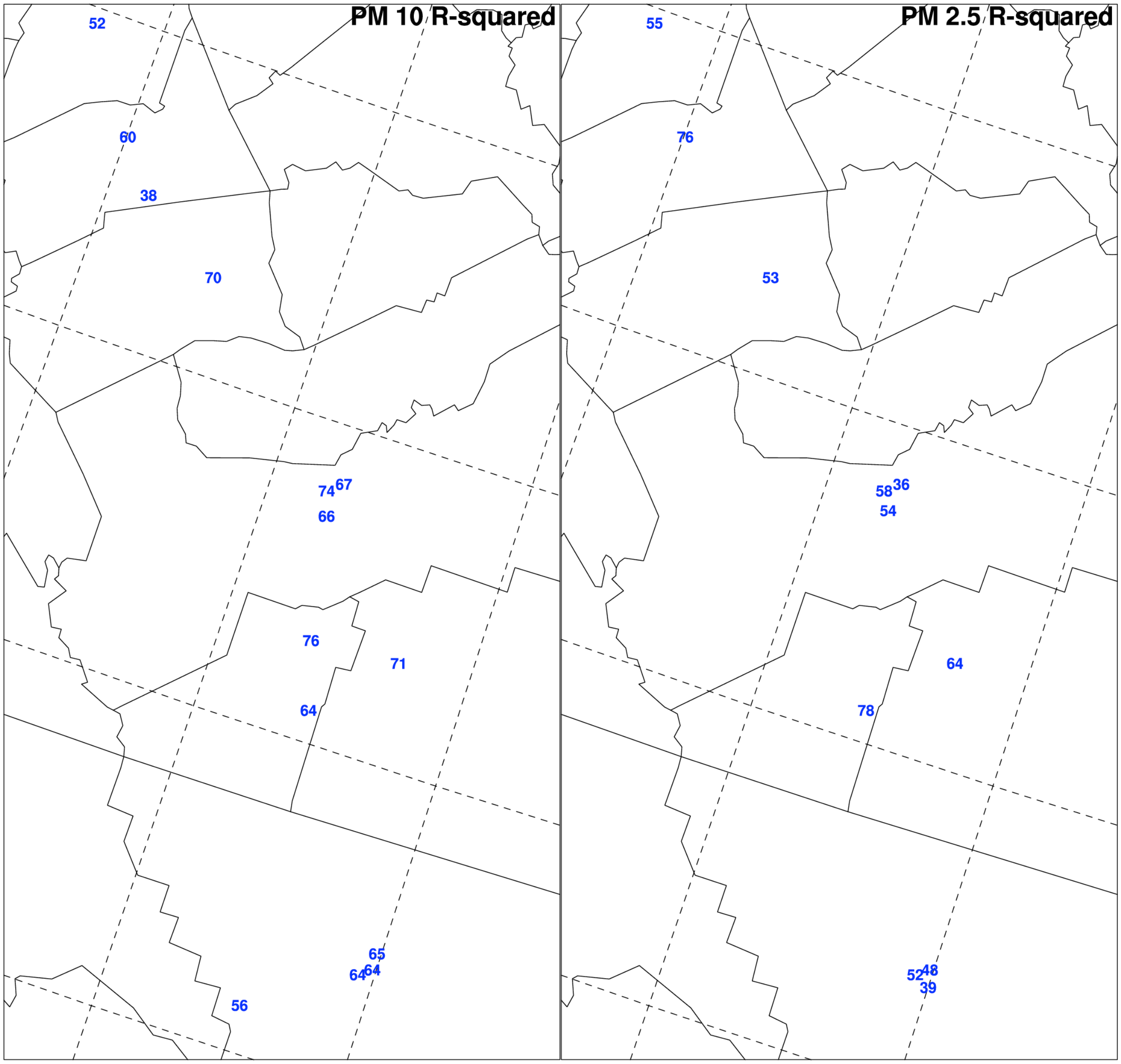

Figure 4. R^2 goodness-of-fit values for each station's MOS equation for PM_{10} and $PM_{2.5}$ for the 00Z forecast.

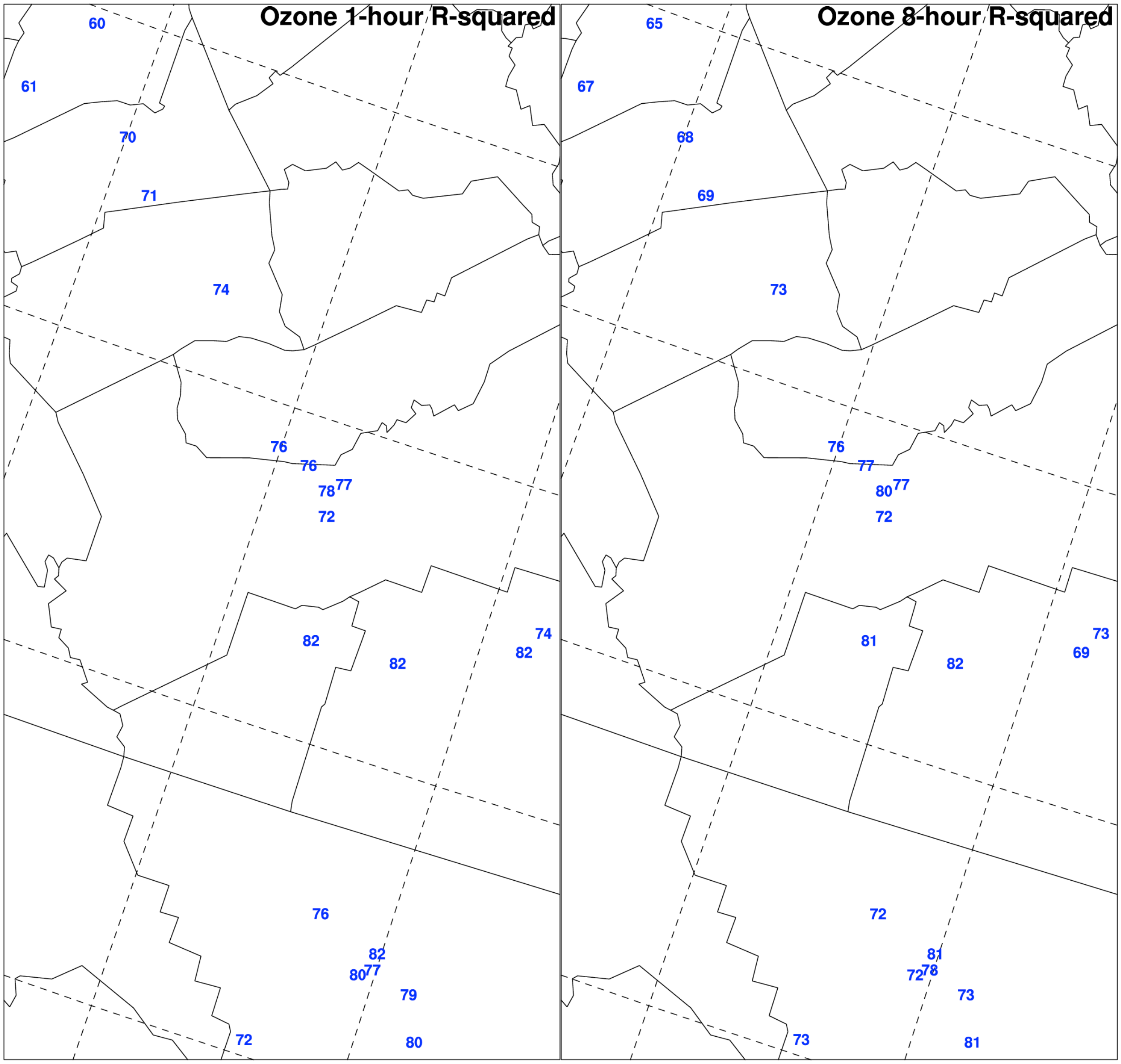

Figure 5. R^2 goodness-of-fit values for each station's MOS equation for ozone 1-hour daily maxima and ozone 8-hour daily maxima for the 00Z forecast.

Table 1 shows the predictors and associated coefficients for an example MOS station regression equation, where *T* is temperature, *U* is the *u* wind vector component, *PP* is perturbation pressure, *.1-4* one of the four model grid point corners closest to the station and *v30* the model sigma level. The resulting equation is:

PM2.5 = 10.8907043811116*T.2.v30 + 0.024970397530653*PP.4.v31 - 19.9993173151886*U.2.v12 - 0.0197082956561172*PP.2.v30 + 18.1993986369606*U.3.v12 + 2.98676339567287*U.1.v11 - 1.48998402218158*U.4.v10 + 0.998564485959724*T.1.v16 - 11.9403735152259*T.3.v30 + 35.1729870380697

Table 1. Example set of predictor variables and associated coefficient for the $PM_{2.5}$ MOS regression equation.

Predictor variable	Coefficient
(Intercept)	35.1729870380697
PP.4.v31	0.0249703975306531
U.2.v12	-19.9993173151886
PP.2.v30	-0.0197082956561172
U.3.v12	18.1993986369606
U.1.v11	2.98676339567287
U.4.v10	-1.48998402218158
T.1.v16	0.998564485959724
T.3.v30	-11.9403735152259
T.2.v30	10.8907043811116

4. Deliverable

The 00Z and 12Z 3-day forecast output is available from: http://cefa.dri.edu/Operational_Products/MOS/txtmosfcsts.php. The table below shows example forecast output for three air quality stations denoted by their AQS station ID. PM forecasts are in units of *ppm* and ozone forecasts in μ/m^3.

```
060190007
Fcst Dy          03/11   03/12   03/13
PM 2.5               0       0       0
PM 10                4       5       6
Ozone 1-hour        24      37      26
Ozone 8-hour        21      30      26

060190008
Fcst Dy          03/11   03/12   03/13
PM 2.5              25      24       9
PM 10                1       5       4
Ozone 1-hour        25      36      33
Ozone 8-hour        24      27      28

060190242
Fcst Dy          03/11   03/12   03/13
PM 2.5              22      20      20
PM 10                9       7      19
Ozone 1-hour        28      39      34
Ozone 8-hour        21      30      30
```


Acknowledgements

We very much thank Evan Shipp, formerly at SJVAPCD, for project support and guidance, providing the air quality observations and evaluating the MOS equations. The San Joaquin Valley Air Pollution Control District funded this project.