Model Performance Analysis of Annual Meteorology
For Air Quality Modeling Using Mesoscale Model 5 (MM5)

Abstract:

The MM5 model was used to generate meteorological fields over a 14-month time period between December 1, 1999 and February 1, 2001 to provide input to air quality models for particulate matter (PM) modeling. The results of the analyses show that the model was able to capture the overall observed behavior of the atmospheric conditions during the entire 14-month time period, although the model performance varied within some subsets of time periods and at some locations.

Introduction:

Meteorological model setup:

Meteorological conditions leading to elevated ozone and PM levels in the San Joaquin Valley (SJV) over the 14-month period from December 1, 1999 to February 1, 2001 were simulated using the PSU/NCAR Mesoscale Model (MM5) (version 3.6) (Grell et al, 1995). The model is based on non-hydrostatic, fully compressible motions that allow users to study the atmospheric motions at small scales by explicitly treating the effects of convective motions on atmospheric circulations. The MM5 model has been improved over more than two decades by contributions from a broad scientific community (http://www.mmm.ucar.edu/mm5) and is recommended for use in air quality studies by the US EPA (2005).

The MM5 model was set up with three nested grids. First, two coarse grids using (70 x 70) and (133 x 133) grid points with 36 and 12 km horizontal resolutions, respectively, in the (x, y) or (south-north, west-east) directions (Figure 1, domains D01 and D02) were established to provide the large scale initial and boundary conditions (IC/BCs) for the innermost grid. This innermost grid has (94 x 85) grid points with 4 km horizontal resolution to resolve the fine details of atmospheric motions within the SJV domain (Figure 1, domain D03). Domain D01 is centered at 120.5^0W longitude and 37.0^0N latitude. The inner 12 and 4-km domains are placed within their respective coarse parent grids using certain grid point offsets from the lower left corner of their parent grid in (x, y) direction as provided in Table 1.

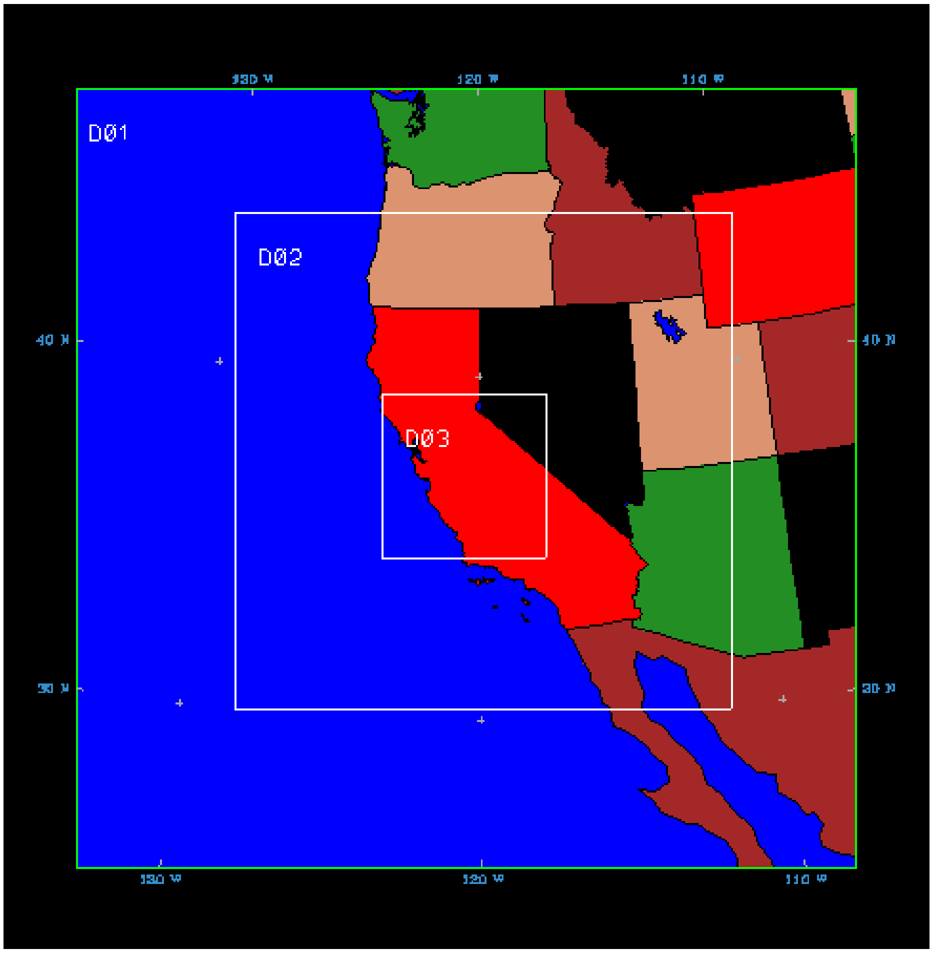


Figure 1: The location of three nested grids designed to study the meteorology and air quality in SJV domain.

	SJV Domain		
Domain name	D01	D02	D03
x-dir resolution	36 km	12 km	4 km
y-dir resolution	36 km	12 km	4 km
# of grids in x-dir	70	133	94
# of grids in y-dir	70	133	85
Lower left corner (N-S)	1	15	43
Lower left corner (W-E)	1	15	54

Table 1: The three nested grids used to study meteorology and air quality in the SJV.

The modeling domain has 30 vertical layers, with a 30-m deep first model layer near the surface, extending to 100 mb at the top of the domain. The thickness of the model layers increase with height. Initially, a higher vertical resolution with more vertical grid levels was used to better resolve atmospheric processes evolving in a stable atmosphere during winter. However, a sensitivity study using a 15 m deep first layer and 50 vertical layers showed that the model results did not improve appreciably, while the total model run time to finish a simulation was nearly doubled. Therefore, the nominal 30 vertical layers with the 30 m deep first model layer were used here.

MM5 has several options to calculate the components of internal and external forces acting on a volume of air including radiation, convection, cloud microphysics, soil fluxes, and boundary layer physics. Many sensitivity studies were conducted using various model options to gain better agreement with observations. Once an optimal configuration was found, it was kept constant for subsequent simulations, although sensitivity studies were still carried out to see if the model results were significantly different from each other. The optimal configuration used here included the Grell (1995) cumulus parameterization scheme (for coarse grids only), the Blackadar boundary layer scheme for calculation of fluxes (Blackadar, 1979, Grell, 1995), the Dudhia simple ice scheme for the treatment of excess moisture (Dudhia, 1989), the Dudhia cloud radiation scheme (Dudhia, 1993) for radiational heating and cooling of the atmosphere, and the Blackadar multi-layer, force-restore method soil model (Blackadar, 1976) for soil physics.

Initial and boundary conditions (IC/BCs) were prepared using the analyses from the National Center for Environmental Prediction (NCEP) archived at NCAR. The 14-month period was first simulated using IC/BCs with the analysis nudging option on the coarse 36 and 12 km grids using two-way nested options. Our modeling experience using various combinations of model options and domain configurations have shown that using the analysis nudging option together with a two-way nested grid option for the first two coarse grids improves the model performance compared to no analysis nudging. Furthermore, the two-way nesting option resolved atmospheric processes within the 12-km grid that the one-way nested option could not. IC/BCs for the 4-km innermost grid were then prepared from the 12 km grid model output, and the.4-km model was run independently from the outer grids.

The year 2000 was chosen for the air quality and meteorology simulations since there were two intensive observational data collection programs carried out that year: 1) Central California Ozone Study (CCOS) between May 1 – September 30, 2000 to study the formation of ozone, and 2) California Regional Particulate Matter Air Quality Study (CRPAQS) between December 1, 1999 and January 31, 2001 to study the formation of PM.

The 14-month period was simulated using one-month time blocks. Each model run was started from the end of a previous numerical simulation using the corresponding history files to create a continuous annual simulation. This method was used since sensitivity simulations showed that the model needed a longer spin-up time of about 12-hr to reach the observed conditions when the model was started without using the history restart files of the previous run. Furthermore, since the analysis nudging option used for the coarse grids forced the model to stay within the boundaries of observed conditions, the model performance did not degrade compared to sensitivity simulations done with shorter time scale such as 1-week long simulations.

Performance Evaluation of 4-km SJV Simulation:

Observational data from 13 surface meteorological stations located within the SJV were analyzed and compared against model estimates to study the temporal and three-dimensional spatial evolution of atmospheric motions as well as to evaluate the model performance. The choice of stations was based on data availability and location within the domain. Figure 2 and Table 2 show the location of the monitoring stations located within the SJV. Since there are a large number of figures analyzed, only selected figures are shown here to demonstrate the model performance. The figures generated for all 13 stations can be found in Attachments A, B, and C.

There are various analysis methods used in the scientific community to study the results and evaluate the performance of a meteorological model. The most commonly used methods are summarized in Attachment D. Three methods are used here:

(1) Direct comparisons of all hourly values and statistical analyses using mean, standard deviation, correlation coefficient, root mean square error (RMSE), mean bias, and index of agreement of temperature, relative humidity, u and v-component of the wind over the entire year and 3-month seasonal periods (January-March, April-June, July-September, and October-December, 2000) at Arvin, Bakersfield, Fresno, Hanford, Madera, Meadows Field Airport in Bakersfield, Merced, Modesto, Parlier, San Luis Obispo, Santa Rita, Visalia, and Yosemite (Attachment A);

(2) Histograms for the frequency of occurrence of temperature, wind speed, and relative humidity over a 24-hour time period for a one-month period every other month (January, March, May, July, September, November) 2000 and January 2001) at sites in the vicinity of Fresno and Bakersfield (Attachment B); and

(3) Hourly comparisons of wind speed, direction, and temperature over two five-day periods within each month (days 1 through 6, and 11 through 16 for 14 months) at Arvin, Bakersfield, Fresno, Merced, Modesto, Parlier, and Visalia (Attachment C).

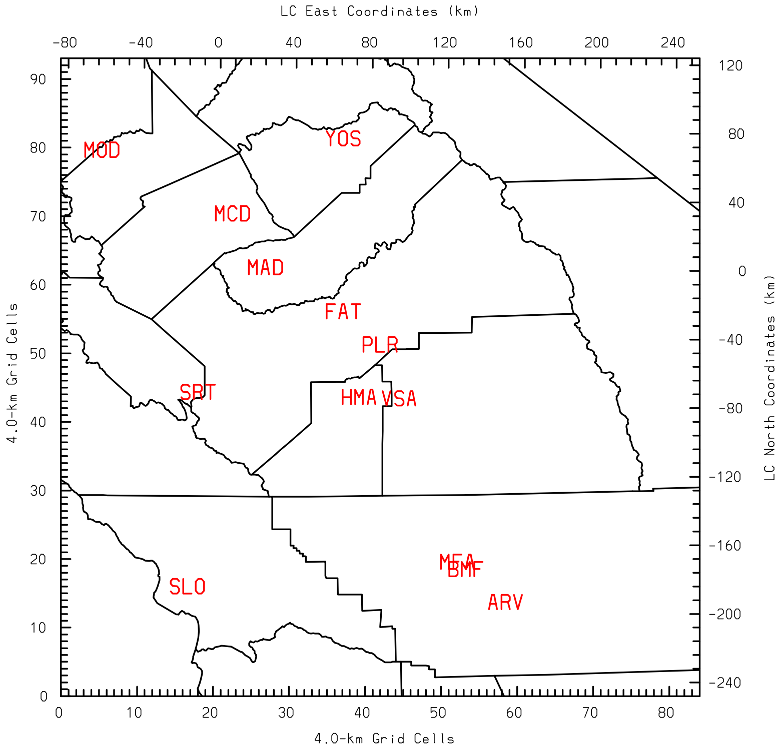


Figure 2: The location of observation stations used in the model performance analysis of MM5.

Site name	Site location	Site county	Lat(degree)	Lon (degree)
ARV	Arvin-Edison	Kern	35.21	-118.78
BMF	Bakersfield Meadows Field	Kern	35.39	-119.01
FAT	Fresno Air Terminal	Fresno	36.78	-119.72
PLR	Parlier	Fresno	36.6	-119.5
MCD	Merced	Merced	37.31	-120.39
MOD	Modesto #3	Stanislaus	37.65	-121.19
MFA	Meadow Field Airport	Kern	35.43	-119.06
VSA	Visailia Airport	Tulare	36.31	-119.39
YOS	Yosemite National Park	Mariposa	37.71	-119.71
SRT	Santa Rita	San Benito	36.35	-120.6
HMA	Hanford Municipal Airport	Kings	36.32	-119.63
MAD	Madera	Madera	37.02	-120.19
SLO	San Luis Obispo	San Luis Obispo	35.306	-120.66

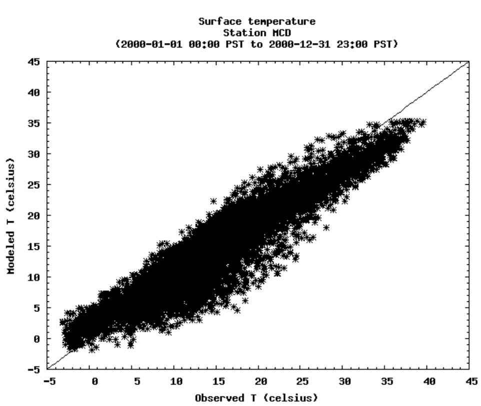
Table 2:The observation stations used in the model performance analysis of MM5.

The results of method 1:

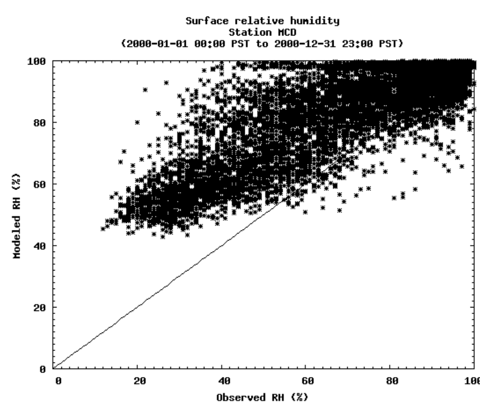
Figures 3a-d compare observed air temperature, relative humidity, u-component, and v-component of the wind near the surface against model estimates at Merced (MCD), for the entire year 2000. Similarly, Tables 3a-d show the statistical calculations performed for the same variables at the same station for the same time period. Merced was chosen as a representative site; figures and statistical comparisons for other sites can be found in Attachment A.

Both figures and tables indicate relatively high correlation between model results and observations for temperature (0.89) and relative humidity (0.71). However, the correlation decreases for the u-component of the wind (0.53) and the v-component (0.27). Moreover, although the correlation for relative humidity is higher than those for the u and v-components of the wind, the model over-predicts the relative humidity most of the time, especially at higher relative humidities. Similar degrees of agreement between model results and observations are also seen at other stations (see Attachment A) for all meteorological variables used. The smaller correlation for the wind components is partly due to the fact that analysis or observational nudging was not used for the 4-km grid.

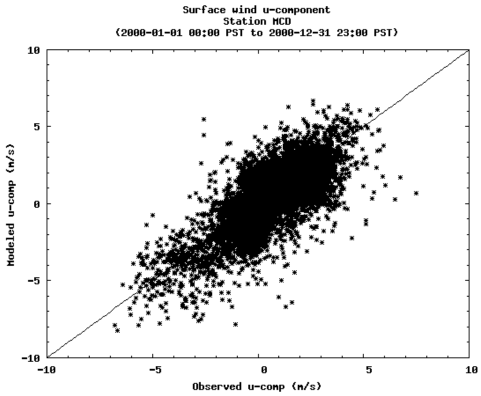
The index of agreement between model results and observations for each model variable do not show large seasonal variations. For example, relative humidity at Merced is 0.79 for the entire year while it is 0.82, 0.78, 0.69, and 0.75 in winter, spring, summer, and fall, respectively. The correlation between model results and observations is the highest for temperature for the entire year and different seasons, but lower for the u and v-component of the wind for each of these time periods. For example, while the correlation is relatively high for the u-component (0.79) for the entire year, it is below 0.4 in other seasons, and it is less than 0.29 for the v-component for all seasons.



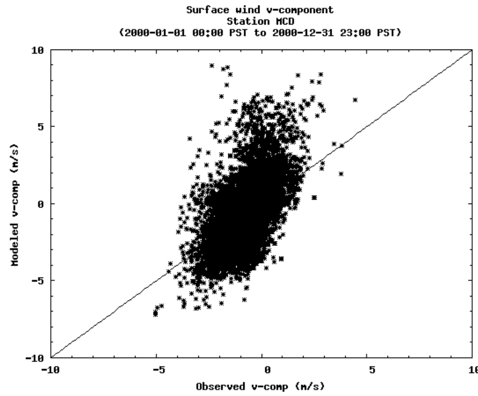
(a)



(b)



(c)



(d)

Figures 3a-d: Comparisons of model results against observations for a) temperature, b) relative humidity, c) u-component, and d) v-component of the wind near the surface at Merced (MCD) for the entire year.

Performance stats.			Obs.	Mod.
R-squared value	0.89	Average value	15.664	14.784
Mean Abs. error	2.34	Standard deviation	8.605	7.889
Mean bias	-0.88	Maximum value	39.56	35.3
RMSE	3	Minimum value	-3.17	-1.83
Agreement index	0.97	Data points	8751	8751

(a)

Performance stats.			Obs.	Mod.
R-squared value	0.71	Average value	69.689	82.937
Mean Abs. error	14.46	Standard deviation	23.769	15.336
Mean bias	13.25	Maximum value	100	100
RMSE	19.01	Minimum value	12	42.76
Agreement index	0.79	Data points	8751	8751

(b)

Performance stats.			Obs.	Mod.
R-squared value	0.53	Average value	0.697	0.453
Mean Abs. error	1.06	Standard deviation	1.7	1.926
Mean bias	-0.24	Maximum value	7.51	6.69
RMSE	1.38	Minimum value	-6.79	-8.24
Agreement index	0.84	Data points	8752	8752

(c)

Performance stats.			Obs.	Mod.
R-squared value	0.27	Average value	-0.519	-0.099
Mean Abs. error	1.27	Standard deviation	0.926	1.962
Mean bias	0.42	Maximum value	4.43	8.99
RMSE	1.73	Minimum value	-5.04	-7.21
Agreement index	0.6	Data points	8752	8752

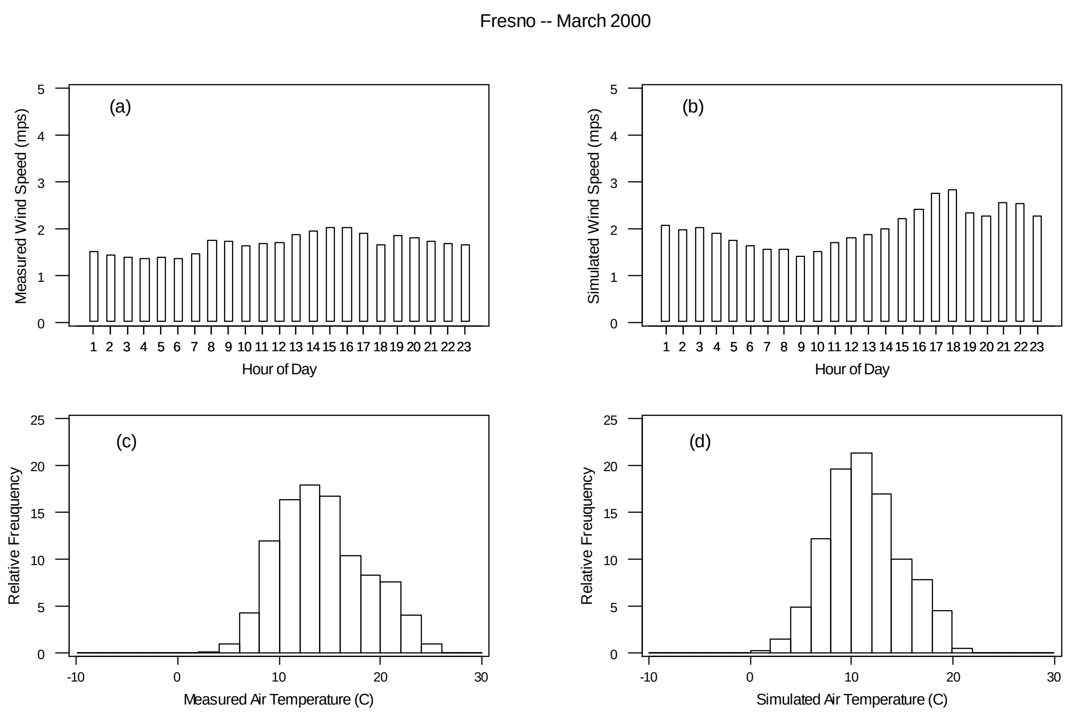
(d)

Tables 3a-d: Statistical comparisons for a) temperature, b) relative humidity, c) u-component, and d) v-component of the wind near the surface at Merced (MCD) for the entire year.

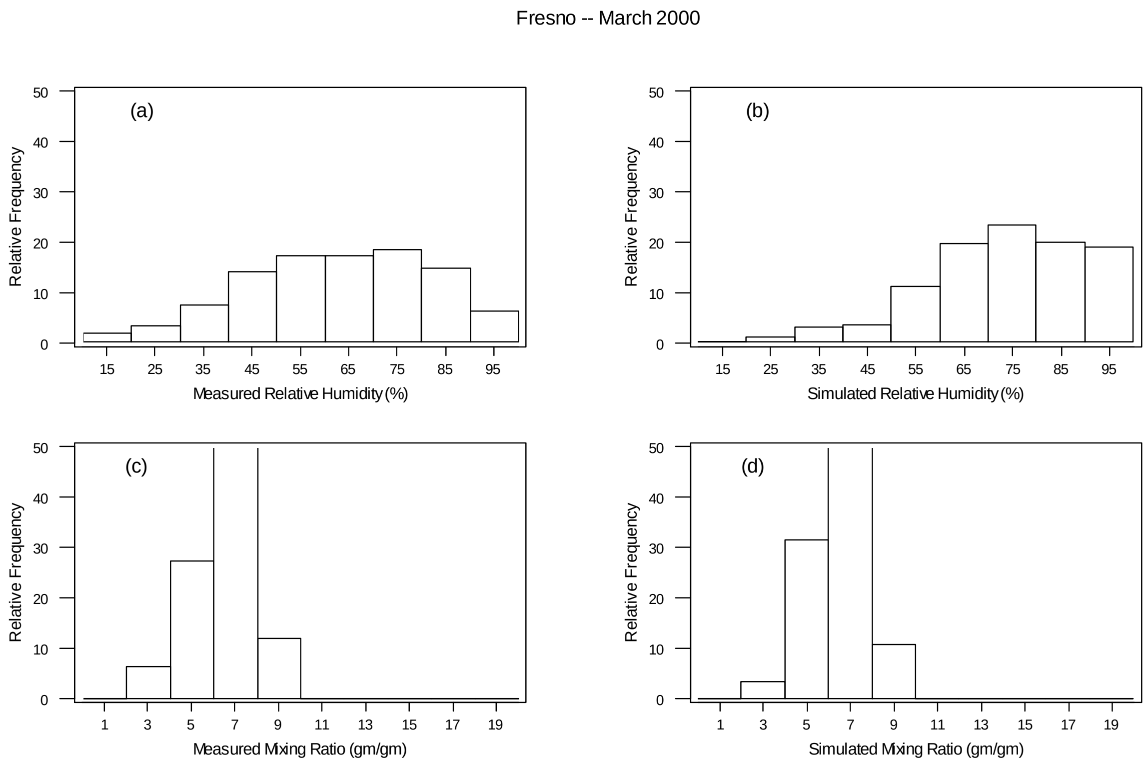
The results of method 2:

In these comparisons, four representative monitoring sites in the vicinities of Fresno and Bakersfield were used and grouped for the analyses performed (Table 1, Attachment B). Only a sample set of figures for Fresno for March 2000 are shown here as samples of the results, while all results from Fresno can be found in Attachment B in greater detail. The results for Bakersfield are similar to those of Fresno and are not shown.

Figures 4a-d and 5a-d show the comparisons of measured and simulated hourly averaged values of air temperature, wind speed, relative humidity, and mixing ratio, respectively, near the surface for March 2000 using histograms. The frequency distribution of air temperature, relative humidity, and mixing ratio for observations and model results show reasonably good agreement. There are only small differences in observed and simulated diurnal wind speed patterns.



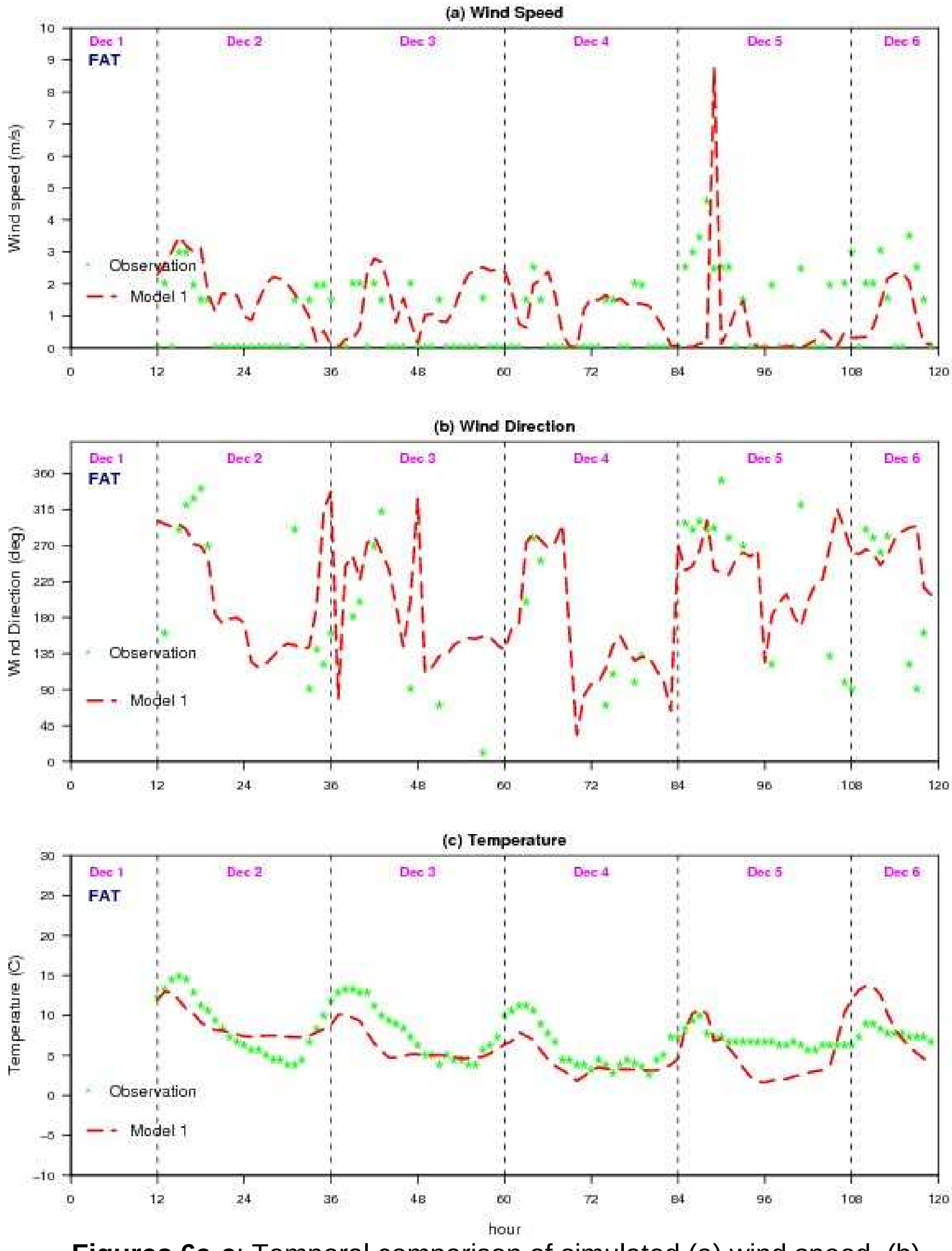
Figures 4a-d: Hourly averaged wind speed (a) measured and (b) simulated, and the frequency of occurrence of air temperatures (c) measured and (d) simulated for sites in the vicinity of Fresno during March, 2000.



Figures 5a-d: Frequency of occurrence of (a) measured and (b) simulated relative humidity, and of (c) measured and (d) simulated water mixing ratio for sites in the vicinity of Fresno during March, 2000.

<u>The results of method 3:</u>

A direct comparison of the model results against observations was also carried out by plotting observed and simulated values for each station over two 5-day periods that were randomly selected within each month (days 1-6 and 11-16). Only a sample figure is given here while all other figures can be found in Attachment C. Temporal comparisons of wind speed, wind direction, and temperature for the Fresno station for December 1-6, 1999 period (Figures 6a-c) show that the observed conditions appear to be captured by the model well. Figures 3-5 a-c and 3-6 a-c in Attachment C show that the model can capture the diurnal evolution of observed wind speed, wind direction and temperature variations reasonably well. While the simulated wind speed and direction are in good agreement with the observations, the simulated temperature also shows variations in capturing the observed maxima and minima. For example, Figures 3-8 a-c indicate that the model overestimates somewhat the observed daily maxima and minima of temperature during March 1-6, 2000 period at Fresno, and the estimates for the evolution of the wind speed and direction differ somewhat. However, the examination of all figures given in Attachment C indicates that the model does capture the overall evolution. Furthermore, the model's ability in recreating observed conditions also shows variations from one station to another (Attachment C).



Figures 6a-c: Temporal comparison of simulated (a) wind speed, (b) wind direction, and (c) temperature near the surface against observations at Fresno Station during December 1-6, 1999 period.

Summary and conclusions:

The model performance analyses described above indicate that MM5 is able to reproduce the overall statistical characteristics of observed meteorological conditions. The model can recreate the mean, standard deviation, and minimum and maximum values of wind speed and direction over the entire year and each season. However, the model estimates of maxima and minima for near surface temperature can show variations in time and location. The correlation coefficient of temperature is high for most stations, although the model underestimates the maximum surface temperature. The model generally captures the observed conditions that show large variations in maxima and minima from one day to the next over some 5-day periods well.

It should be noted that the simulated time period is 14 months, and no analysis or observational nudging was performed in the 4-km domain. It has been shown by many numerical modelers that observational nudging improves the overall model performance. However, the use of observational nudging can result in the distortion of overall wind flow patterns, or can induce instabilities. Observational nudging was not used in this study, however the model was able to capture expected and observed conditions well.

References:

Blackadar, A. K., 1976: Modeling the Nocturnal Boundary Layer. Preprints of Third Symposium on Atmospheric Turbulence and Air Quality, Raleigh, NC. 19-22, October 1976, Amer. Meteor. Soc., Boston, 46-49.

Blackadar, A. K., 1979: Advances in Environmental Science and Engineering, 1. No. 1. Pfaflin and Ziegler, Eds. Gordon and Breach Publishers, 50-85.

Dudhia, J., 1989: Numerical Study of Convection Observed During the Winter Monsoon Experiment Using a Two-Dimensional Model. J. Atmos. Sci., 46, 3077-3107.

Dudhia, J., 1993: Radiation Studies with a High-Resolution Mesoscale Model. Proceedings of the Third Atmospheric Radiation Measurement (ARM) Science Team Meeting. March 1-4, 1993, Norman, OK. pp. 363-366.

Grell, G. A., J. Dudhia and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). *NCAR Technical Note*, NCAR/TN-398+STR.

US EPA, 2005: Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hour Ozone NAAQS. EPA-454/R-05-002. OAQPS/EMAD/AQMG, RTP, NC, October 2005.