

Southeastern Modeling, Analysis and Planning Project

Executive Summary:

WRF Meteorological Modeling Performance Evaluation and Documentation

Prepared for:

**Southeastern States Air Resource Managers, Inc.
526 Forest Pkwy, Ste F
Forest Park, GA 30297-6140**

Prepared by:

**Atmospheric and Environmental Research, Inc.
131 Hartwell Ave
Lexington, MA 02421**

19 December 2011

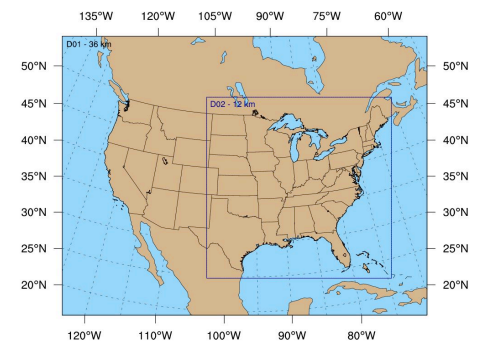
1 Introduction

The Southeastern States' Air Resource Managers, Inc. (SESARM) is required to provide their member states with technical expertise and data in support of each member state's air quality implementation plan. The purpose of this project is to evaluate the quality of a set of weather data provided by SESARM to AER, specifically meteorological fields (temperature, wind, moisture) from a weather model. Determining the suitability of the data for subsequent use in air quality modeling is a driving motivation.

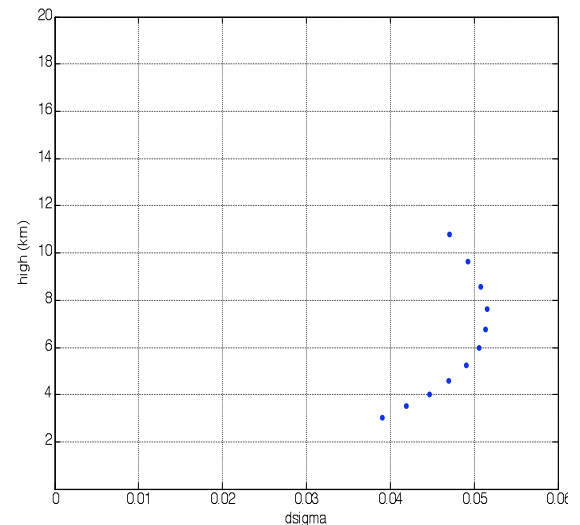
2 Weather Prediction Model

The weather fields to be evaluated in this project were generated by the most recent version (3.1) of a meteorological prediction model called the Weather Research and Forecast model (WRF). The model is open-source software (available from wrf-model.org) managed by the University Center for Atmospheric Research (UCAR) and has a vast user base. A closely related model is currently the principal short-range model used by the US government to predict the weather. The WRF used in this study was selected because it allows the user to choose from multiple configurations to describe how different physical processes (e.g., the movement of heat and moisture) in the atmosphere are represented in the model, and ultimately select the optimal configuration.

The figure at right illustrates the parts of the United States where the weather data were evaluated. Data are available for two areas, domain 1 (36 km) and domain 2 (12 km). The data from domain 2, while covering a smaller horizontal area, benefit from having a more realistic representation of the Earth's surface.



The atmosphere is divided into 35 horizontal layers (34 layers) as shown at right. Each dot signifies one level. The height of the level is shown on the y-axis and the thickness of the atmospheric layers (as a fraction of the overall depth of the atmosphere) is shown on the x-axis. The number of levels is more concentrated below 1 km for increased resolution in the boundary layer. The second level is placed at ~20 m so that the midpoint of the grid cells in the vertical direction is located at 10 m and so corresponds to the standard National Weather Service (NWS) anemometer height without the need for vertical interpolation.



2.1 Evaluation period

Data for the Production Run was evaluated every day for the calendar year 2007. Analysis of Sensitivity Runs was carried out for two 20-day periods, July 28-August 16 ("Summer") and December 5-24 ("Winter") during which air quality standards

were exceeded across large parts of the region. Details of these two different types of runs follow later in this document.

2.2 Observational data and evaluation software

A set of observed weather conditions is required for comparison against the modeled weather data. A full year of observations in “prepBUFR” format was downloaded from the UCAR web server (<http://dss.ucar.edu/datasets/ds337.0/>). This quality-controlled dataset is used by the US government as input to its daily weather model runs. The evaluation software (available from dtcenter.org) is called WRF Model Evaluation Tools (WRF-MET) and was designed by the WRF developers specifically to evaluate their model using the prepBUFR observations.

2.3 Weather fields and target error limits

The following six weather fields were evaluated by WRF-MET: temperature (degrees Celsius), dew point temperature (degrees Celsius), specific humidity (g/kg), relative humidity (%), wind speed (m/s), and wind direction (degrees). For eventual use in air quality models, errors in the atmospheric model forecasts ideally should be less than the guidelines provided by SESARM in the table. Gross Error, also known as Mean Absolute Error, is a measure of the overall magnitude of the errors and is thus always a positive number. RMSE means Root-Mean-Squared-Error and is similar to the Gross Error, except that it more heavily penalizes very large errors. Mean Bias indicates whether there is a systematic error. For all of these measures, values near zero are desired.

| Quantitative Assessment Metrics | | |
|---------------------------------|-------------|--------------------|
| Wind Speed | RMSE | ≤ 2 m/s |
| | Mean Bias | $\leq \pm 0.5$ m/s |
| Wind Direction | Gross Error | ≤ 30 deg |
| | Mean Bias | $\leq \pm 10$ deg |
| Temperature | Gross Error | ≤ 2 K |
| | Mean Bias | $\leq \pm 0.5$ K |
| Specific Humidity | Gross Error | ≤ 2 g/kg |
| | Mean Bias | $\leq \pm 1$ g/kg |

3 Model fields and evaluation techniques

There were ten different sets of modeled weather fields: the year-long Production Run and nine Sensitivity Runs of 20 days each for the Summer and Winter Evaluation Periods. For each set of model data, AER carried out two different types of evaluations:

1) Subjective: AER closely inspected the details of how the model was run. This included reading the log files and input files that describe the choice of model parameters.



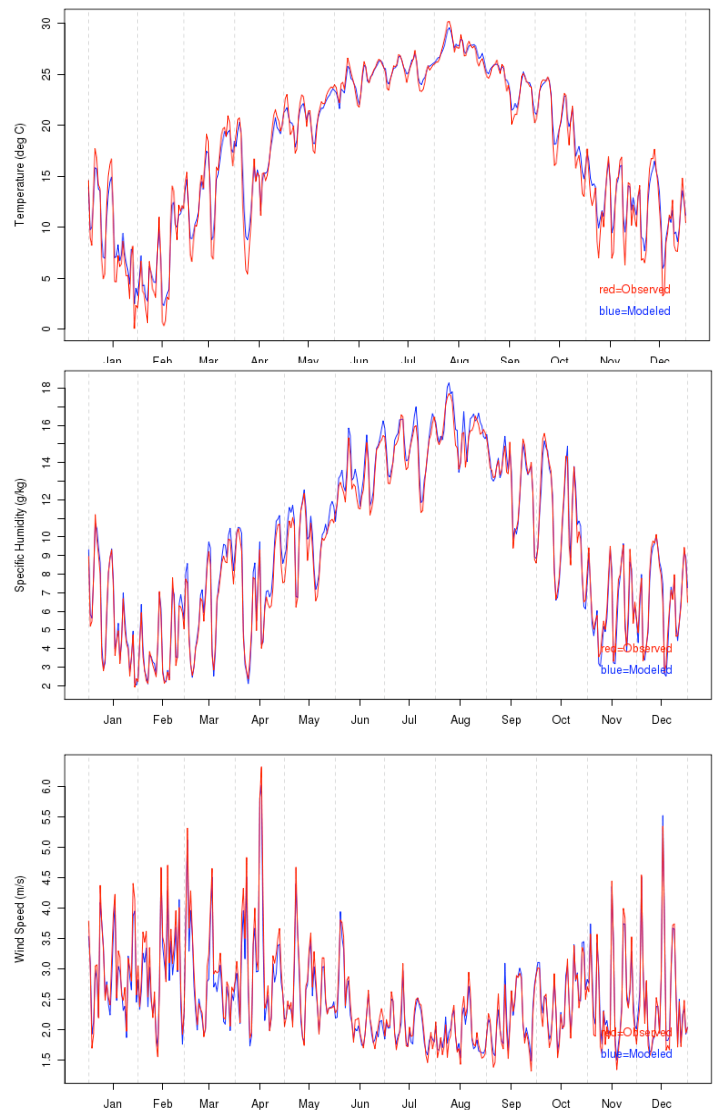
2) Objective: The WRF-MET software was applied to the model data and observations, and error statistics were computed and presented in the form of graphs and tables. WRF-MET carried out the initial steps to properly match in time the model data, which is on a grid similar in concept to the familiar latitude and longitude grid, while the observations mostly are measured at airport locations. These matched pairs of data were then combined, using the freely-available statistical analysis and plotting software called R, over many time periods (e.g., days and months) and then averaged over a selected set of states and Regional Planning Organization regions (RPOs; see figure on previous page): Alabama, Arkansas, District of Columbia, Florida, Georgia, Iowa, Illinois, Indiana, Kentucky, Louisiana, Maryland, Michigan, Mississippi, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, Wisconsin, West Virginia, MANEVU (ME, MA, VT, NH, NY, PA, NJ, MD, DE), CENRAPN (MN, IA, NE, KS, MO), CENRAPN (TX, LA, OK, AR), SEMAP [VISTAS] (VA, WV, KY, TN, NC, SC, GA, AL, MS, FL), MRPO (MI, WI, IN, IL, OH) and FULL (entire model domain).

4 Results

4.1 Subjective

In our opinion, the WRF model configuration selected for the production run is acceptable for generation of meteorological fields to be input to air quality models. A reasonable mix of physics options and model configuration parameters were selected using an up-to-date release of the WRF model. Sensitivity run number three (“px_acm2_morr_rrtmg”; see table in Section 4.3) was selected by SESARM as the configuration for the production run, and used the following settings: Morrison explicit moisture, RRTMG radiation, PX land surface and surface layer physics, Kain-Fritsch cumulus parameterization, and ACM2 planetary boundary layer (PBL) scheme.

The model configurations for the nine sensitivity runs were also set up correctly and spanned a sufficiently broad set of physics choices such that the final choice of production run configuration could be satisfactorily determined. Sufficient documentation was provided by SESARM to be confident that the WRF model was run in the manner designed and intended.



4.2 Objective statistical analysis of full-year production run

The three time-series charts on the previous page serve as a demonstration of the overall good agreement between the modeled and observed temperature (top), specific humidity [moisture] (middle) and wind speed (bottom) fields for the SEMAP states. The observed weather fields are in red and the modeled fields are in blue.

A summary of domain 2 annual statistics for the FULL and SEMAP regions are presented in the tables below. Statistics are compared to those target values found in the Metric table in Section 2.3. Error statistics that exceed the target errors are highlighted in red; otherwise the entries are colored green. The reader should be reminded that the values contained in Section 2.3 are target values from the historical record and the color scheme should not be construed simply as a pass or fail designation. Note that rounding may introduce small disagreements between values in different columns.

4.2.1 FULL

| Field | Number of Obs | Model Mean | Obs Mean | Mean Bias | Mean Absolute Error | Root-Mean-Squared Error | Correlation Coefficient | Multiplicative Bias |
|-----------------------------------|---------------|------------|----------|-----------|---------------------|-------------------------|-------------------------|---------------------|
| Temperature (degrees C) | 18082480 | 12.7 | 12.6 | 0.2 | 1.9 | 2.5 | 0.92 | 1.001 |
| Dew Point Temperature (degrees C) | 16272421 | 7.4 | 6.8 | 0.7 | 2.2 | 3.0 | 0.89 | 1.002 |
| Relative Humidity (%) | 16272365 | 73.2 | 69.8 | 3.4 | 11.4 | 14.6 | 0.59 | 1.049 |
| Specific Humidity (g/kg) | 17015610 | 8.0 | 7.9 | 0.2 | 1.0 | 1.3 | 0.90 | 1.028 |
| Wind Speed (m/s) | 17648777 | 3.4 | 3.5 | -0.2 | 1.4 | 1.8 | 0.68 | 0.994 |
| Wind Direction (degrees) | 17648777 | | | -4.2 | 34.6 | | | |

4.2.1.1 Temperature

Annual MAEs for temperature lie within the target value and daily values do not exhibit any unexpected characteristics. Daily MAE values are largest, often upwards of 2.5 degrees Celsius, for a number of days during the cold months of November through April, with summertime values at times below 1.5 degree Celsius. Some of the larger daily errors are clustered together in time, often in five-day segments, which strongly indicates that individual model runs suffered from poor initialization. Annual values of mean bias are small. Daily values

during the summer almost exclusively are within 1 degree Celsius, while the most extreme values occurred during the three other seasons. The spring and early summer from March through July produced the vast majority of days with a negative temperature bias. Days with the largest positive bias occurred mostly from October through April. An especially poor period with large (positive) forecast errors occurred in early April.

4.2.1.2 Dew point temperature

The annual mean bias of 0.7 degrees indicates that, on average, modeled dew point temperatures were higher than observed. Significantly contributing to this domain-wide bias is the widespread over-prediction of dew point temperatures in the northern parts of the modeling domain during the colder months. This problem persists from December through April. The MAEs are notably larger during the winter months, again predominantly in the northern part of the modeling domain. Problems with snow cover in the model may be a contributing factor.

4.2.1.3 Relative humidity

Annual mean bias values also indicate modeled fields with slightly higher levels of relative humidity than observed. The occurrence of days with over-predicted relative humidity is concentrated strongly in December and from January through May and, again, in the northern parts of the modeling domain. March, by far, has the most days with the largest mean bias values. The only months with daily negative mean bias values of magnitude greater than 5% are October and November. The overall annual MAE of 11.4% is generally representative of individual daily MAE values. As might be expected, there is a minimum during the summer months, while maximum winter gross errors exceed 15% on a handful of days.

4.2.1.4 Specific humidity

A small positive mean bias, well within the target range, is seen in the annual specific humidity field, which agrees with the characteristics of the dew point temperature and relative humidity. A positive bias in the Northeastern states is present during March and April. The magnitude, persistence and large-scale nature of this bias, especially during the cold season when absolute values of the atmospheric moisture are at a minimum, again suggests a link to how the model handles snowcover. MAE values overall are very close to the target value and are largest south of 40 °N latitude during the months of August and September. The lack of a widespread and persistent mean bias signature at this time, indicative of errors all of the same sign, suggests that the larger MAE values may be largely due to the fact that atmospheric moisture content is close to its maximum and thus subject to the largest error.

4.2.1.5 Wind speed

The annual mean bias lies within the target range, and both it and the MAE are quite small and show no significant temporal variability during the course of the year. MAE errors overall are slightly higher in the northern parts of the modeling domain in the winter and spring when overall wind speeds are highest and the passage of synoptic weather systems increases the variability of the wind field. The annual RMSE is slightly less than the target of 2 m/s.

4.2.1.6 Wind direction

(Note that, by convention, wind direction indicates the direction from which the wind is blowing. Here we denote, for instance, zero degrees as a North wind, 90 degrees as an East wind and -90 degrees as a West wind.) While the mean bias is well inside the target range, the MAE is slightly larger than the target value of 30 degrees. Thus, while there are substantial deviations from the observed wind, they do not exhibit a substantial bias in direction. The wind direction field is strongly affected by local topography, and, while there are no expansive and persistent regions of significant mean bias, individual stations for any one day can exhibit large biases. These individual episodes of large bias are likely tied to poor modeling of synoptic features and can be sufficiently persistent to affect monthly averages, especially in mountainous or otherwise elevated terrain. Indeed, such areas suffer from the largest and most persistent MAE values. This is most noticeable in the Appalachian mountains throughout the year. During the summer months with lighter winds, individual sites throughout the modeling domain can suffer at times from large MAE values.

4.2.2 SEMAP

| Field | Number of Obs | Model Mean | Obs Mean | Mean Bias | Mean Absolute Error | Root-Mean-Squared Error | Correlation Coefficient | Multiplicative Bias |
|-----------------------|---------------|------------|----------|-----------|---------------------|-------------------------|-------------------------|---------------------|
| Temperature | 3250685 | 17.9 | 17.6 | 0.3 | 1.8 | 2.3 | 0.81 | 1.001 |
| Dew Point Temperature | 3104743 | 10.9 | 10.6 | 0.4 | 2.2 | 2.8 | 0.78 | 1.001 |
| Relative Humidity | 3104716 | 68.9 | 67.5 | 1.5 | 11.2 | 14.2 | 0.47 | 1.021 |
| Specific Humidity | 3107904 | 9.6 | 9.5 | 0.2 | 1.2 | 1.5 | 0.76 | 1.019 |
| Wind Speed | 3139180 | 2.5 | 2.6 | -0.0 | 1.3 | 1.6 | 0.52 | 1.115 |
| Wind Direction | 3139180 | | | -1.5 | 42.8 | | | |

4.2.2.1 Temperature

Annual MAEs for temperature lie within the target value and daily values do not exhibit any unexpected characteristics. Daily MAE values are largest, often upwards of 2.5 degrees Celsius, for a number of days during the cold months of October through April, with summertime values mostly below 2 degrees Celsius. Some of the larger daily errors are clustered together in time, often in five-day segments, strongly indicating that individual model

runs suffered from poor initialization. As for FULL, an especially bad forecast segment occurred in early April. Annual values of mean bias are small and, again, do not indicate a problem. Daily values during the summer almost exclusively are within 1 degree Celsius, while the most extreme values occurred during the three other seasons. During these cooler seasons, individual days exhibit the largest mean biases, sometimes in excess of 2.5 degrees Celsius.

4.2.2.2 Dew point temperature

The annual mean bias of 0.4 degrees indicates that, on average, modeled dew point temperatures were higher than observed. Of note are the episodes of poorly predicted dew point temperatures centered over North Carolina during November. Detailed analysis is beyond the scope of this project, but during this month there were a number of consecutive days that exhibited large negative mean bias values. The largest MAEs are found in the mountainous or otherwise elevated regions.

4.2.2.3 Relative humidity

Annual mean bias values also indicate modeled fields with slightly higher levels of relative humidity than observed. The occurrence of days with under-predicted relative humidity is concentrated strongly in November, as noted above for dew point temperature. During the rest of the year, there is a predominance of positive daily mean bias of up to 10-15%. The annual MAE of 11.2% is generally representative of individual daily MAE values. As might be expected, there is a minimum during the summer months, while maximum winter gross errors exceed 15% on a handful of days. These are focused somewhat during the month of November.

4.2.2.4 Specific humidity

A small positive mean bias, well within the target range, is seen in the annual specific humidity field, which agrees with the characteristics of the dew point temperature and relative humidity. A positive bias, with a corresponding increase in MAE, in some individual locations that are mostly located in mountainous terrain or plateaus, develops during the summer, peaks in August, then diminishes during the fall. Mean bias values in the fall and early winter are then near zero or slightly negative. It is hypothesized that the positive bias that peaks in August may reflect inadequate model representation of topography, with model locations being positioned at lower elevations than in reality and subjected to higher values of PBL moisture in an otherwise horizontally-uniform moisture environment.

4.2.2.5 Wind speed

The annual mean bias and MAE both lie within the target range and show no important temporal variability during the course of the year. The annual RMSE is acceptably less than the target of 2 m/s.

4.2.2.6 Wind direction

The mean bias for SEMAP lies within the target range, while, as for FULL, the MAE is slightly larger than the target of 30 degrees. As for FULL, deviations of the wind from observed while large do not exhibit a substantial bias in direction. This meteorological field is strongly affected by local topography, and, while there are no expansive and persistent regions of significant mean bias (since the signs of the individual errors often cancel), regions of mountainous topography do suffer from the largest and most persistent MAE values. Along the spine of the Appalachian mountains, it is commonplace throughout the year to exceed the target MAE value. During the summer months with lighter winds, individual sites throughout the modeling domain can suffer at times from large MAE values.

4.3 Analysis of sensitivity runs

A similar subjective and objective analysis was carried out for the sensitivity runs. The table below lists the model configurations for each of the nine runs. The reader is referred to the Final Report for more details. Note that the production run, though it used the model configuration for sensitivity run number three, initialized its run segments at different times (among other potential differences) and so a separate set of production run statistics is included in the error tables below. Also note that run seven is unavailable for the Summer period. The magnitude of the errors has been color-coded as in Section 4.2.

| CENTER | RUN NUMBER/ NAME | MOIST PHYSICS | LONG WAVE | SHORT WAVE | LAND SURFACE | SOIL LAYERS | PBL | SURFACE LAYER | CUMULUS |
|--------|----------------------------------|------------------|--------------|---------------|-----------------|----------------|-------------------|------------------|---------|
| IA | 1 - px-acm2_wsm5 | WSM5(4) | RRTM(1) | Dudhia(1) | PX(7) | 2 | ACM2(7) | PX(7) | KF (1) |
| IA | 2 - px-acm2_wsm6 | WSM6(6) | RRTM(1) | Dudhia(1) | PX(7) | 2 | ACM2(7) | PX(7) | KF (1) |
| IA | 3 - px_acm2_morr_ rrtmg | Morrison (10) | RRTMG(4) | RRTMG(4) | PX(7) | 2 | ACM2(7) | PX(7) | KF (1) |
| IA | 4 - px-acm2_morr _rrtmg_ipxwf | Morrison (10) | RRTMG(4) | RRTMG(4) | PX(7) IPXWRF | 2 | ACM2(7) | PX(7) | KF (1) |
| NY/UMD | 5 - myj.wsm5 | WSM5(4) | RRTM(1) | Dudhia(1) | Noah(2) | 4 | MYJ(2) | MO(2) | KF (1) |
| NY/UMD | 6 - myj.wsm6 | WSM6(6) | RRTM(1) | Dudhia(1) | Noah(2) | 4 | MYJ(2) | MO(2) | KF (1) |
| NY/UMD | 7 - ysu.wsm5 | WSM5(4) | RRTM(1) | Dudhia(1) | Noah(2) | 4 | YSU(1) | MO(1) | KF (1) |
| NY/UMD | 8 - ysu.wsm6 | WSM6(6) | RRTM(1) | Dudhia(1) | Noah(2) | 4 | YSU(1) | MO(1) | KF (1) |
| NY/UMD | 9 - OTC | WSM6(6) | RRTM(1) | Dudhia(1) | PX(7) | 2 | Blackadar (11) | PX(7) | KF (1) |

The following tables summarize the domain 2 Summer and Winter statistics at the surface for each sensitivity run for SEMAP. Noteworthy aspects of the statistics are briefly discussed. Note that rounding may introduce small disagreements between values in different columns.

4.3.1 Temperature

Summer Temperature Celsius

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 176141 | 27.5 | 27.6 | -0.1 | 1.5 | 2.0 | 0.68 | 1.000 |
| 1-px-acm2_wsm5 | 176141 | 27.4 | 27.6 | -0.2 | 1.5 | 2.0 | 0.69 | 0.999 |
| 2-px-acm2_wsm6 | 176141 | 27.4 | 27.6 | -0.2 | 1.5 | 2.0 | 0.69 | 0.999 |
| 3-px_acm2_morr_rrtmg | 176141 | 27.5 | 27.6 | -0.1 | 1.5 | 1.9 | 0.69 | 1.000 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 176141 | 27.8 | 27.6 | 0.2 | 1.6 | 2.1 | 0.67 | 1.001 |
| 5-myj.wsm5 | 176141 | 27.5 | 27.6 | -0.1 | 1.5 | 2.0 | 0.71 | 1.000 |
| 6-myj.wsm6 | 176141 | 27.5 | 27.6 | -0.1 | 1.5 | 2.0 | 0.71 | 1.000 |
| 7-ysu.wsm5 | 0 | NA | NA | NA | NA | NA | NA | NA |
| 8-ysu.wsm6 | 176141 | 27.6 | 27.6 | 0.0 | 1.6 | 2.1 | 0.70 | 1.000 |
| 9-OTC | 176141 | 27.2 | 27.6 | -0.4 | 1.5 | 1.9 | 0.71 | 0.999 |

Winter Temperature Celsius

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 199175 | 11.4 | 11.0 | 0.4 | 2.1 | 2.7 | 0.89 | 1.001 |
| 1-px-acm2_wsm5 | 199175 | 11.0 | 11.0 | 0.0 | 2.1 | 2.7 | 0.89 | 1.000 |
| 2-px-acm2_wsm6 | 199175 | 11.0 | 11.0 | 0.0 | 2.1 | 2.7 | 0.89 | 1.000 |
| 3-px_acm2_morr_rrtmg | 199175 | 11.3 | 11.0 | 0.3 | 2.1 | 2.6 | 0.89 | 1.001 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 199175 | 11.5 | 11.0 | 0.5 | 2.2 | 2.8 | 0.89 | 1.002 |
| 5-myj.wsm5 | 199175 | 10.6 | 11.0 | -0.4 | 1.8 | 2.3 | 0.90 | 0.999 |
| 6-myj.wsm6 | 199175 | 10.6 | 11.0 | -0.4 | 1.8 | 2.4 | 0.90 | 0.999 |
| 7-ysu.wsm5 | 199175 | 10.5 | 11.0 | -0.5 | 1.8 | 2.3 | 0.91 | 0.998 |
| 8-ysu.wsm6 | 199175 | 10.5 | 11.0 | -0.5 | 1.8 | 2.3 | 0.91 | 0.998 |
| 9-OTC | 199175 | 11.2 | 11.0 | 0.2 | 1.9 | 2.4 | 0.91 | 1.001 |

Model performance was generally very good for all sensitivity runs for temperature. Summer correlation values were notably lower than for winter. It is anticipated that correlation values will be lower during periods of more uniform temperature, such as during the summer, with short-term periods of cooling driven largely by less predictable diurnal convection. Mean bias values fell within target values, as did MAE values, except for the winter period for runs that utilized the px-acm2 parameterizations. These runs had gross errors of slightly larger than the target of 2 degrees Celsius.

4.3.2 Relative humidity

Summer RH %

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 167793 | 72.4 | 69.3 | 3.1 | 9.7 | 12.3 | 0.43 | 1.049 |
| 1-px-acm2_wsm5 | 167793 | 72.4 | 69.3 | 3.1 | 10.0 | 12.6 | 0.45 | 1.048 |
| 2-px-acm2_wsm6 | 167793 | 72.4 | 69.3 | 3.1 | 10.0 | 12.6 | 0.45 | 1.048 |
| 3-px_acm2_morr_rrtmg | 167793 | 72.4 | 69.3 | 3.1 | 10.0 | 12.6 | 0.44 | 1.048 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 167793 | 73.4 | 69.3 | 4.1 | 10.3 | 12.9 | 0.38 | 1.070 |
| 5-myj.wsm5 | 167793 | 70.2 | 69.3 | 0.9 | 9.5 | 12.1 | 0.48 | 1.015 |
| 6-myj.wsm6 | 167793 | 70.1 | 69.3 | 0.8 | 9.5 | 12.1 | 0.48 | 1.015 |
| 7-ysu.wsm5 | 0 | NA | NA | NA | NA | NA | NA | NA |
| 8-ysu.wsm6 | 167793 | 64.9 | 69.3 | -4.4 | 9.4 | 12.2 | 0.51 | 0.935 |
| 9-OTC | 167793 | 73.6 | 69.3 | 4.3 | 10.5 | 13.3 | 0.45 | 1.069 |

Winter RH %

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 188866 | 74.6 | 75.5 | -0.9 | 12.1 | 15.2 | 0.46 | 0.980 |
| 1-px-acm2_wsm5 | 188866 | 77.2 | 75.5 | 1.7 | 12.4 | 15.5 | 0.44 | 1.019 |
| 2-px-acm2_wsm6 | 188866 | 77.1 | 75.5 | 1.7 | 12.4 | 15.5 | 0.44 | 1.019 |
| 3-px_acm2_morr_rrtmg | 188866 | 76.0 | 75.5 | 0.5 | 11.8 | 14.9 | 0.46 | 1.000 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 188866 | 77.3 | 75.5 | 1.9 | 12.0 | 15.1 | 0.41 | 1.019 |
| 5-myj.wsm5 | 188866 | 79.3 | 75.5 | 3.8 | 9.6 | 12.2 | 0.49 | 1.056 |
| 6-myj.wsm6 | 188866 | 79.5 | 75.5 | 4.0 | 9.8 | 12.4 | 0.48 | 1.060 |
| 7-ysu.wsm5 | 188866 | 76 | 75.5 | 0.5 | 8.9 | 11.5 | 0.57 | 1.010 |
| 9-ysu.wsm6 | 188866 | 76.2 | 75.5 | 0.7 | 9.0 | 11.6 | 0.56 | 1.013 |
| 9-OTC | 188866 | 76.2 | 75.5 | 0.8 | 11.6 | 14.6 | 0.50 | 1.005 |

Modeled relative humidity forecasts typically exhibited a positive bias of up to 4%, as indicated also by the multiplicative bias scores. It should be kept in mind that relative humidity is dependent not only on the absolute amount of atmospheric moisture, but also the temperature. See the specific humidity tables below for evaluation solely of the amount of moisture. Use of the MYJ planetary boundary layer scheme improved the modeled RH fields in the summer, with the best scores across all statistics. In the winter, on the other hand, this scheme's large positive mean bias values, but relatively small MAE errors, indicate that the sign of the forecast errors did not cancel and RH was generally overpredicted. The YSU PBL scheme had the best error statistics in the winter.

4.3.3 Specific humidity

Summer SPFH g/kg

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 167951 | 16.1 | 15.9 | 0.3 | 1.7 | 2.2 | 0.56 | 1.017 |
| 1-px-acm2_wsm5 | 167951 | 15.9 | 15.9 | 0.1 | 1.8 | 2.2 | 0.59 | 1.005 |
| 2-px-acm2_wsm6 | 167951 | 15.9 | 15.9 | 0.1 | 1.8 | 2.2 | 0.59 | 1.005 |
| 3-px_acm2_morr_rrtmg | 167951 | 16.1 | 15.9 | 0.2 | 1.8 | 2.3 | 0.57 | 1.013 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 167951 | 16.8 | 15.9 | 0.9 | 2.0 | 2.5 | 0.50 | 1.057 |
| 5-myj.wsm5 | 167951 | 15.5 | 15.9 | -0.3 | 1.8 | 2.3 | 0.61 | 0.980 |
| 6-myj.wsm6 | 167951 | 15.5 | 15.9 | -0.3 | 1.8 | 2.3 | 0.61 | 0.979 |
| 7-ysu.wsm5 | 0 | NA | NA | NA | NA | NA | NA | NA |
| 8-ysu.wsm6 | 167951 | 14.4 | 15.9 | -1.4 | 2.0 | 2.5 | 0.62 | 0.909 |
| 9-OTC | 167951 | 16.1 | 15.9 | 0.3 | 1.9 | 2.4 | 0.57 | 1.016 |

Winter SPFH g/kg

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 189058 | 6.7 | 6.8 | -0.1 | 1.0 | 1.2 | 0.83 | 0.983 |
| 1-px-acm2_wsm5 | 189058 | 6.8 | 6.8 | 0.0 | 0.9 | 1.2 | 0.84 | 0.997 |
| 2-px-acm2_wsm6 | 189058 | 6.7 | 6.8 | 0.0 | 0.9 | 1.2 | 0.84 | 0.997 |
| 3-px_acm2_morr_rrtmg | 189058 | 6.8 | 6.8 | 0.0 | 0.9 | 1.2 | 0.84 | 0.999 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 189058 | 7.0 | 6.8 | 0.2 | 0.9 | 1.2 | 0.83 | 1.027 |
| 5-myj.wsm5 | 189058 | 6.9 | 6.8 | 0.1 | 0.8 | 1.1 | 0.85 | 1.023 |
| 6-myj.wsm6 | 189058 | 6.9 | 6.8 | 0.1 | 0.8 | 1.1 | 0.85 | 1.023 |
| 7-ysu.wsm5 | 189058 | 6.6 | 6.8 | -0.2 | 0.8 | 1.1 | 0.85 | 0.970 |
| 8-ysu.wsm6 | 189058 | 6.6 | 6.8 | -0.2 | 0.8 | 1.1 | 0.85 | 0.970 |
| 9-OTC | 189058 | 6.9 | 6.8 | 0.1 | 0.9 | 1.2 | 0.84 | 1.005 |

All sensitivity runs met the target values for mean bias and mean absolute error (with one exception). Use of the YSU PBL scheme appears to result in the lowest amount of atmospheric moisture, perhaps related to deficiencies in this scheme's amount of vertical mixing. As expected, error statistics for the winter are smaller, due largely to the significantly smaller amount of atmospheric moisture during this time period. It is suggested, therefore, that statistics for the summer period represent a more rigorous evaluation. Of note is the degradation of the model performance (e.g., increased positive mean bias) through introduction of the IPX-WRF (runs number 4). The IPXWRF software is used to initialize soil moisture in the PX land surface model.

4.3.4 Wind speed

Summer WSPD m/s

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 168497 | 1.9 | 1.8 | 0.1 | 1.3 | 1.6 | 0.35 | 1.319 |
| 1-px-acm2_wsm5 | 168497 | 2.2 | 1.8 | 0.3 | 1.2 | 1.5 | 0.40 | 1.462 |
| 2-px-acm2_wsm6 | 168497 | 2.2 | 1.8 | 0.3 | 1.2 | 1.5 | 0.40 | 1.462 |
| 3-px_acm2_morr_rrtmg | 168497 | 2.2 | 1.8 | 0.3 | 1.2 | 1.5 | 0.40 | 1.468 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 168497 | 2.2 | 1.8 | 0.3 | 1.2 | 1.6 | 0.41 | 1.475 |
| 5-myj.wsm5 | 168497 | 2.3 | 1.8 | 0.5 | 1.2 | 1.5 | 0.42 | 1.496 |
| 6-myj.wsm6 | 168497 | 2.3 | 1.8 | 0.5 | 1.2 | 1.5 | 0.42 | 1.496 |
| 7-ysu.wsm5 | 0 | NA | NA | NA | NA | NA | NA | NA |
| 8-ysu.wsm6 | 168497 | 2.1 | 1.8 | 0.2 | 1.2 | 1.5 | 0.41 | 1.395 |
| 9-OTC | 168497 | 1.6 | 1.8 | -0.3 | 1.2 | 1.6 | 0.40 | 1.015 |

Winter WSPD m/s

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error | Root Mean Square Deviation | Correlation Coefficient | Multiplicative Bias |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|----------------------------|-------------------------|---------------------|
| production | 194801 | 2.6 | 2.5 | 0.0 | 1.2 | 1.5 | 0.57 | 1.085 |
| 1-px-acm2_wsm5 | 194801 | 2.8 | 2.5 | 0.3 | 1.2 | 1.5 | 0.62 | 1.194 |
| 2-px-acm2_wsm6 | 194801 | 2.8 | 2.5 | 0.3 | 1.2 | 1.5 | 0.62 | 1.195 |
| 3-px_acm2_morr_rrtmg | 194801 | 2.9 | 2.5 | 0.3 | 1.2 | 1.5 | 0.61 | 1.207 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 194801 | 2.9 | 2.5 | 0.3 | 1.2 | 1.5 | 0.61 | 1.209 |
| 5-myj.wsm5 | 194801 | 3.0 | 2.5 | 0.5 | 1.2 | 1.5 | 0.62 | 1.257 |
| 6-myj.wsm6 | 194801 | 3.0 | 2.5 | 0.5 | 1.2 | 1.5 | 0.61 | 1.255 |
| 7-ysu.wsm5 | 194801 | 2.9 | 2.5 | 0.4 | 1.3 | 1.6 | 0.60 | 1.241 |
| 8-ysu.wsm6 | 194801 | 2.9 | 2.5 | 0.4 | 1.3 | 1.6 | 0.60 | 1.241 |
| 9-OTC | 194801 | 2.5 | 2.5 | 0.0 | 1.2 | 1.5 | 0.60 | 1.034 |

All sensitivity runs for each time period met the target mean bias and RMSE ranges for wind speed. MAE and RMSE differences amongst the runs were unimportant. Overall mean bias errors were small and positive, likely influenced by the models' well-known inability to reproduce coincident observations of calm air. Interestingly, the production run had the smallest overall mean bias, especially when compared against the sensitivity run that used same model configuration. As mentioned above, one factor contributing to this difference in error statistics is the different start times for each 5.5-day run segment. Overall, the correlation coefficient scores were significantly better during the winter, likely a result of the predominance of more robust synoptic-scale flow.

4.3.5 Wind direction

Summer WDIR degrees

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|
| production | 168497 | -128.3 | -132.4 | 4.1 | 54.7 |
| 1-px-acm2_wsm5 | 168497 | -132.1 | -132.4 | 0.3 | 52.6 |
| 2-px-acm2_wsm6 | 168497 | -131.8 | -132.4 | 0.6 | 52.6 |
| 3-px_acm2_morr_rrtmg | 168497 | -133.6 | -132.4 | -1.2 | 52.5 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 168497 | -137.7 | -132.4 | -5.3 | 52.3 |
| 5-myj.wsm5 | 168497 | -132.1 | -132.4 | 0.3 | 51.7 |
| 6-myj.wsm6 | 168497 | -132.2 | -132.4 | 0.2 | 51.7 |
| 7-ysu.wsm5 | 0 | NA | NA | NA | NA |
| 8-ysu.wsm6 | 168497 | -133.6 | -132.4 | -1.2 | 52.0 |
| 9-OTC | 168497 | -130.7 | -132.4 | 1.7 | 52.1 |

Winter WDIR degrees

| Run | Number of Obs | Model Mean | Observed Mean | Mean Bias | Mean Absolute Error |
|-----------------------------|---------------|------------|---------------|-----------|---------------------|
| production | 194801 | -29.0 | -48.0 | 19.0 | 40.9 |
| 1-px-acm2_wsm5 | 194801 | -13.0 | -48.0 | 35.0 | 38.2 |
| 2-px-acm2_wsm6 | 194801 | -14.1 | -48.0 | 33.9 | 38.3 |
| 3-px_acm2_morr_rrtmg | 194801 | -26.5 | -48.0 | 21.5 | 38.4 |
| 4-px-acm2_morr_rrtmg_ipxwrf | 194801 | -31.7 | -48.0 | 16.4 | 38.5 |
| 5-myj.wsm5 | 194801 | -23.6 | -48.0 | 24.4 | 38.2 |
| 6-myj.wsm6 | 194801 | -22.8 | -48.0 | 25.2 | 38.2 |
| 7-ysu.wsm5 | 194801 | -71.8 | -48.0 | -23.8 | 38.4 |
| 8-ysu.wsm6 | 194801 | -71.7 | -48.0 | -23.7 | 38.4 |
| 9-OTC | 194801 | -77.6 | -48.0 | -29.6 | 38.2 |

The production run, and closely-related sensitivity run number three, exhibit two of the smallest mean bias values during the winter when none of the bias values lie within the target of less than 10 degrees. For the summer, given that the observations for METAR sites are provided only to the nearest 5 degrees, we consider the variability in the mean bias values to be unimportant. For both time periods, MAE errors lie outside the target value of 30 degrees. The inter-run variability is very small and there is no clear “best” choice based solely on that metric.

4.4 Summary

AER does not disagree with the process by which the “best” sensitivity run was selected or the final choice of the “best” sensitivity run. When compared to all other sensitivity runs, the model parameters used by this “best” run: a) generally produced modeled weather fields that exhibited some of the smallest errors, and b) did not exhibit any apparent egregious errors.

Based on experience, however, clearly identifying a single model configuration that minimizes model error for all meteorological fields for large geographical areas and for long periods of time is frequently a difficult exercise. Analysis of the errors from the nine sensitivity runs

presented to AER bear out the expectation that there is simply no *clearly* identifiable 'best' model configuration that minimizes errors in *all* fields for *all* times and in *all* regions.

5 Meteorological Modeling Website and Data Availability

The project web site (<http://sesarm.aer.com>) serves as the primary means for distributing and evaluating the substantial meteorological modeling results/summaries and communicating the corresponding documentation to SESARM. This web site satisfies the requirements for 'Task 5 Technical Meteorological Modeling Web Site' of contract S-2009-09-01. All content on this web site has been approved by the SESARM Project Officer prior to posting. The website presents objective evaluation results as itemized by Amendment 2 of the project Contract, signed 4 August 2011. All WRF-MET point_stat output files that represent the full set of hourly statistics files have been provided on an external drive to SESARM.

6 Summary and Conclusions

Do the model results fit our conceptual understanding?

Meteorological fields generated during this modeling effort agree with our conceptual understanding of the physical processes of the atmosphere, and, as such, are generally representative of the observed weather conditions.

Are diurnal (i.e., day-to-night cycle) features adequately captured?

The WRF model adequately captures the diurnal cycle of temperature, moisture and wind. Overall, though, forecast errors are minimized during periods when the surface winds are driven by the familiar transient low pressure areas. During spring and summer when strong sunshine causes warm air near the surface to mix up vertically above the surface, the model tends to underpredict surface wind speeds during the afternoon. However, this is a common model limitation and the overall effect on the value of the weather fields is relatively small.

Do the wind fields agree with the observations?

Overall surface wind fields are well-represented in the model and provide valuable estimation of the flow fields in the troposphere. Error values for wind fields well above the Earth's surface far removed from the effects of complex land surface features were noticeably smaller.

Wind direction forecasts did not exhibit substantial biases, however, mean absolute errors for both domains were at or slightly above the critical threshold of 30 degrees. In combination, these two sets of errors suggest that the direction errors were somewhat random in nature, which is not unexpected when compared with other modeling studies.

Do the temperature and moisture fields generally match the observations?

Overall the temperature and moisture fields generally match the observations. Values of mean bias and gross error for temperature were at or below the critical thresholds.

Both mean bias and gross errors for specific humidity, a measure of the absolute content of the moisture in the atmosphere, were well within acceptable thresholds, even during the summer when levels of moisture in the atmosphere are significantly higher than in the cold winter months.

Are the meteorological fields likely to produce acceptable air quality results? Are all of the above features captured in a spatially- and temporally-realistic manner that is consistent with model grid spacing.

AER recommends the future use of these WRF runs for the following reasons: 1) the fields were generated from the latest version of a state-of-the-art meteorological model, 2) proper configuration of the model was followed, 3) error magnitudes and patterns are consistent with those in the literature, and 4) there is an absence of obvious, profound, systematic modeling errors. Despite the broad recommendation, it is suggested that future users of subsets of these raw meteorological fields, and subsequent air quality fields, closely inspect the quality of the individual WRF runs, since the quality of any one run can vary considerably.