APPENDIX A

Model Performance Evaluation of the 2002 36 km MM5 Meteorological Model Simulation used in the CENRAP Modeling and Comparison to VISTAS Final 2002 36 km MM5 and WRAP Interim 2002 36 km MM5 Simulations
The CENRAP 2002 36 km MM5 simulation (Johnson, 2007) was evaluated against observed surface and upper-air meteorological observations and observed precipitation amounts and its performance was compared against the VISTAS final and the WRAP interim 2002 36 km MM5 simulations. The CENRAP, VISTAS and WRAP 2002 36 km MM5 simulations used several common science options:

- Lambert Conformal Projection with center at (97°, 40°) and standard parallels at (33°, 45°).
- 164 by 128 36 km by 36 km horizontal grids covering the continental U.S. and adjacent regions.
- 34 vertical layers up to 100 mb (~15 km AGL).
- Pleim-Xiu Land Surface Module (LSM).
- Asymmetric Convective Mixing (ACM) Planetary Boundary Layer (PBL) model.
- RRTM long-wave radiation.
- Dudhia short-wave radiation.
- No Shallow convection.

However, there were some differences in the choice of science options:

- VISTAS and CENRAP MM5 simulations used the Kain Fritsch 2 cumulus parameterization, whereas WRAP MM5 used Kain Fritsch 1.
- VISTAS and CENRAP MM5 simulations used the Reisner 1 moist physics while WRAP MM5 used Reisner 2.
- All three MM5 simulations used Four Dimensional Data Assimilation (FDDA analysis nudging at the surface for winds, but WRAP also used surface analysis nudging to temperature and moisture.
- All three MM5 simulations used analysis nudging FDDA above the PNL to winds, temperature and moisture.

Much of the difference in the model performance for the three MM5 simulations was related to the surface temperature and moisture analysis nudging used in the interim WRAP MM5 simulations that resulted in better surface temperature model performance, but caused instabilities resulting in degradation in meteorological model performance above the surface. The final WRAP 2002 36 km MM5 simulation did not use the surface temperature and moisture FDDA and used the Betts-Miller cumulus scheme instead of Kain Fritsch that resulted in much improved meteorological model performance in the western States (Kemball-Cook et al., 2005).

### A.1 Surface Meteorological Model Performance

The performance of the three MM5 simulations at the surface was evaluated through comparisons against observed surface wind, temperature and humidity measurements from the ds472 observational database. The METSTAT program was used to evaluate the MM5 simulations for each month of 2002 and across the 11 subdomains shown in Figure A-1. These subdomains are as follows:
Emery and Tai (2001) have developed model performance benchmarks by analyzing over 30 MM5/RAMS meteorological model simulations and tabulating the typical level of performance that a good meteorological model achieves. These performance benchmarks are not intended to be pass/fail grades; rather they provide a framework to evaluate the model performance against past applications. Since many of the past MM5/RAMS meteorological model simulations that the benchmarks were developed from were in support of urban ozone modeling that are typically fairly stagnant conditions with little or no precipitation and involved multiple iterations to achieve the final base case simulation. Thus, we may not expect the 2002 annual MM5 simulations to achieve a similar level of performance given the complicating factors of precipitation and complex terrain associate with many Class I areas in the west. Table A-1 lists the meteorological model performance benchmarks for wind speed, wind direction, temperature and humidity.

**Table A-1.** Meteorological model performance benchmarks (Source: Emery et al., 1999).

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Wind Speed</th>
<th>Wind Direction</th>
<th>Temperature</th>
<th>Humidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>≤ 2 m/s</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Bias</td>
<td>≤ ±0.5 m/s</td>
<td>≤ ±10°</td>
<td>≤ ±0.5 K</td>
<td>≤ ±1.0 g/kg</td>
</tr>
<tr>
<td>Index of Agreement</td>
<td>≤ 0.6</td>
<td></td>
<td>≤ 0.8</td>
<td>≤ 0.6</td>
</tr>
<tr>
<td>Gross Error</td>
<td></td>
<td>≤ 30°</td>
<td>≤ 2.0 K</td>
<td>≤ 2.0 g/kg</td>
</tr>
</tbody>
</table>

Below we present the evaluation of the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations against surface meteorological observations for the four seasonal months of January, March, July and October and the CENRAP North (CenrapN) and CENRAP South (CenrapS) subdomains (i.e., subdomains 5 and 6 in Figure A-1). The surface evaluation of the three MM5 2002 36 km simulations outside of the CENRAP subdomains can be found in Kemball-Cook et al., (2004).
Figure A-1. Eleven subdomains where monthly evaluation of the MM5 simulations surface model performance was evaluated.
A.1.1 Temperature

Figure A-2 displays the surface temperature model performance for the CENRAP, VISTAS and WRAP 2002 36 km MM5 simulations in the CenrapN and CenrapS subdomains and the months of January, March, July and October. The WRAP MM5 simulations are performing best for January temperature in both CENRAP domains exhibiting low bias and the lowest error that are within the benchmark. The VISTAS MM5 runs is performing next best with bias well within the benchmark and error within but close to the error benchmark. The CENRAP MM5 simulation performs well for the CenrapS domain with zero bias and error within, but approaching the benchmark. However, the CENRAP performance for the CenrapN domain does not achieve the performance benchmarks due to a too cold bias.

The temperature performance in March is similar to January with both the VISTAS and WRAP MM5 simulations achieving the benchmark for both CENRAP subdomains. Again the CENRAP MM5 simulation has a near zero bias and achieves the error benchmark in the CenrapS subdomain, but is too cold in the CenrapN domain falling out of the bias benchmark range.

In July the three simulations achieve the temperature benchmark in both CENRAP subdomains, although the WRAP MM5 simulations is cooler with the CenrapS bias right at the -0.5 K lower bound benchmark. The CENRAP MM5 simulation is slightly warmer than the VISTAS MM5 simulation.

In October, all three MM5 simulations achieve the temperature performance benchmarks. The WRAP MM5 simulation performs best with near zero bias and lower error than either the VISTAS or CENRAP simulations. The VISTAS and CENRAP MM5 simulations exhibit nearly identical temperature performance in October with a near zero bias for the CenrapS subdomain and a cool bias for the CenrapN subdomain.

In conclusion, the WRAP MM5 simulation is always performing best for surface temperature with the lowest bias and usually the lowest error. The VISTAS MM5 simulations is performing next best as the CENRAP MM5 simulations exhibits a cool bias for the CenrapN subdomain in January and March that exceed the performance benchmarks.
Figure A-2a. Temperature performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and January (top) and March (bottom).
Figure A-2b. Temperature performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and July (top) and October (bottom).
A.1.2 Humidity

The humidity performance for the three MM5 simulations is comparable and always achieves the performance benchmarks. The humidity bias is always near zero for all three runs and four months. In January, March and October the humidity error is at or less than half of the 2.0 g/kg benchmark. However, in July there is more error in the humidity with it within but approaching the benchmark value for all three models.

In conclusion, all three MM5 simulations achieved the humidity benchmark performance goals for all months studied. No model simulation exhibited superior performance over another.
Humidity performance is comparable for all three runs.

Performance is comparable for all three runs.

Figure A-3a. Humidity performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and January (top) and March (bottom).
Figure A-3b. Humidity performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and July (top) and October (bottom).
A.1.3 Winds

The model performance for wind speed and direction and January is almost identical and within the benchmarks for all three models and both CENRAP subdomains. In fact, the performance is so close the CenrapS symbols are plotted over and obliterate the CenrapN performance symbols.

In March, the wind performance is within the benchmark for all three MM5 simulations, which exhibit similar performance statistics. The wind performance in the CenrapS subdomain is slightly better than CenrapN with the CENRAP MM5 simulations showing the largest wind speed RMSE in the CenrapN subdomain, although still within the benchmarks.

Slight degraded wind direction performance is seen in July with the error increases to just below 20 degrees to just below the 30 degree benchmark value for all three models. Similar wind speed RMSE is seen for all three models.

The October wind performance is within the benchmarks for all three models with performance between that seen for January/March and July.

In summary, the models exhibited similar model performance for surface wind speed and direction.
Figure A-4a. Wind Speed and Wind Direction performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and January (top) and March (bottom).
Figure A-4b. Wind Speed and Wind Direction performance for the CENRAP, VISTAS and interim WRAP 2002 36 km MM5 simulations, the CenrapN and CenrapS subdomains and July (top) and October (bottom).
A.2 Upper-Air Meteorological Evaluation

Figure A-5 displays an example comparison of the vertical profile of predicted and observed winds and temperature for Midland, Texas and January 7 2002 at 12 GMT (6am LST) and for July 16, 2002 at 00 GMT (6pm LST). Above the surface, all three models do a good job in replicating the observed temperature, dew point temperature and winds at 6a on January 7, 2002. Although the WRAP MM5 simulation predicts the surface temperature better than the other two simulations, the vertical structure of the temperature and the surface temperature inversion is not reproduced as well.

All three models understate the afternoon PBL depth on July 16, 2002 at Midland Texas. This phenomenon was seen at other sites as well.

The upper-air meteorological model evaluation found that all three models had difficulty reproducing the observed nocturnal inversion. The day time convective mixing depths were also typically underestimated.

Although the WRAP MM5 simulation reproduced the surface temperature the best of the three models, it was worst at reproducing the observed vertical temperature structure and resultant level of mixing. These results are likely due to the surface data assimilation of temperature employed by the WRAP interim MM5 simulation and resulted in WRAP eliminating the surface temperature and humidity FDDA in their final simulation.
Figure A-5. Comparison of predicted and observed vertical temperature, dew point and winds profiles for the CENRAP (left), VISTAS (middle) and WRAP (right) at Midland Texas on January 7, 2002 at 12 GMT (top) and July 16, 2002 at 00 GMT (bottom).
A.4 Precipitation Model Performance Evaluation

The three MM5 model simulation precipitation estimates were evaluated by comparing the monthly average spatial distributions and amounts with observed values from the observed CPC 0.25 by 0.25 degree (approximately 28 km by 28 km) gridded analysis fields. The CPC analysis fields are gridded from on U.S. land-based observations, consequently the gridded observed fields are not available over the oceans and Canada and Mexico. The CPC observed monthly average precipitation fields were displayed using the MM5 modeling domain. The MM5 total precipitation estimates were accumulated for a month and plotted. Here total precipitation includes both explicit large scale synoptic precipitation as well as the subgrid-scale convective precipitation from the cumulus parameterization (Kain Fritsch 1 or 2).

Figures A-6 through A-9 display the monthly average precipitation fields for the months of January, March, July and October and the CPC observed and CENRAP, VISTAS and interim WRAP MM5 simulations. In January (Figure A-6), all three models reproduce the observed monthly average precipitation well with enhanced predicted and observed precipitation over the Pacific Northwest and the Appalachian Mountains. The MM5 simulations also estimated enhanced precipitation in off-shore areas north of Seattle, over the Atlantic Ocean and in the Gulf of Mexico that can not be either confirmed or refuted by the CPC observations. MM5 does overstate the amount of precipitation in January over the northern CENRAP region including over Minnesota, Iowa and Nebraska.

The three models also do a good job in reproducing the observed spatial distribution and amounts of the precipitation in March 2002 (Figure A-7). Elevated precipitation areas in the Pacific Northwest and across the lower Midwest from Arkansas and up into the Ohio River Valley and adjacent areas. The MM5 simulations do understate the highest observed precipitation amounts in Arkansas. The MM5 simulations also overstate the amount of precipitation in the desert southwest (Four Corners) area in March.

The MM5 monthly average precipitation performance is dramatically worse in July 2002 (Figure A-8). Precipitation is overstated by all three MM5 simulations throughout the U.S. and particularly in the southern states, from Arkansas across Texas to the southeastern U.S. particularly Florida South and North Carolina. This over-prediction bias is due to convective precipitation from the cumulus parameterization (either Kain Fritsch 1 or 2). This overactive precipitation is the result of the over-prediction bias I humidity seen in many subdomains (see Table A-3b and Kemball-Cook et al., 2004a).

In October 2002, the three MM5 simulations reproduced the observed monthly average rainfall fairly well across the U.S. (Figure A-9). The models predict the location of the maximum precipitation in southern Louisiana well, but under-predict the magnitude, which may be due to a slight spatial displacement offshore in the Gulf of Mexico. The MM5 simulations understate the precipitation over the CENRAP region, which explains the dry humidity bias in the CenrapS subdomain in October (Figure A-3b).
In conclusion, the three MM5 simulations do a good job in simulating the observed precipitation when it is due to synoptic weather systems. However, when precipitation is due to convective activity as seen in July that is simulated by the MM5 cumulus parameterization, MM5 greatly overstates the precipitation amounts. This is particularly pronounced in the southern states from the Four Corners area to Florida with the interim WRAP simulation exhibiting the largest over-prediction bias. In the final WRAP MM5 simulation the Betts-Miller cumulus parameterization was used that greatly reduced the convective precipitation amounts resulting in better model performance (Kemball-Cook et al., 2005). However, an overestimation bias under convective precipitation conditions still was present.

Figure A-6. Comparison of January 2002 observed monthly average precipitation (top left) with predicted values for the CENRAP (top right), VISTAS (bottom left) and WRAP (bottom right January 2002 simulation (note: observed precipitation not valid over water due to lack of measurements).
Figure A-7. Comparison of March 2002 observed monthly average precipitation (top left) with predicted values for the CENRAP (top right), VISTAS (bottom left) and WRAP (bottom right January 2002 simulation (note: observed precipitation not valid over water due to lack of measurements).
Figure A-8. Comparison of July 2002 observed monthly average precipitation (top left) with predicted values for the CENRAP (top right), VISTAS (bottom left) and WRAP (bottom right) (note: observed precipitation not valid over water due to lack of measurements).
Figure A-9. Comparison of October 2002 observed monthly average precipitation (top left) with predicted values for the CENRAP (top right), VISTAS (bottom left) and WRAP (bottom right) (note: observed precipitation not valid over water due to lack of measurements).