ENVIRON

MEMORANDUM

To: WRAP Modeling Forum

From: Sue Kemball-Cook, Chris Emery, Yiqin Jia, Ralph Morris

Date: December 17, 2004

Subject: **Response to Reviewer Comments on the Draft Report:** "2002 ANNUAL MM5 SIMULATION TO SUPPORT WRAP CMAQ VISIBILITY MODELING FOR THE SECTION 308 SIP/TIP: MM5 Sensitivity Simulations to Identify a More Optimal MM5 Configuration for Simulating Meteorology in the Western United States

---

ENVIRON and UCR developed the subject draft report in September 2004 (hereafter referred to as the "Draft Sensitivity Study Document"), following initial annual MM5 applications for WRAP conducted on the National 36-km RPO domain. The purpose of the report is to document a series of in-depth sensitivity tests with MM5 to identify a more optimal model configuration specific to improve performance in the WRAP region for both 36-km and 12-km grid modeling. The goal of the tests was achieved.

The WRAP Modeling Forum distributed the draft report to other RPO's and agencies for peer review. WRAP received written responses from three groups: VISTAS (Mike Abraczinskas, George Bridgers, and Don Olerud), the State of Nevada (Frank Forsgren), and EPA/ORD (Tanya Otte). These responses were forwarded to ENVIRON.

This memorandum provides the WRAP RMC response to reviewer comments. Both the original comments and our response to each are provided below.

Response to "Comments on WRAP Draft MM5 Sensitivity Modeling" dated October 27, 2004.
Mike Abraczinskas, and George Bridgers, North Carolina Division of Air Quality (NCDAQ)
Don Olerud, Barons Meteorological Services (BAMS)

1) *The document mentions a plan to do 12-km modeling. Accordingly, why aren't the statistics/plots presented made on that domain? If 12-km results are the ultimate goal, we're particularly concerned about using the Betts-Miller cumulus scheme at that resolution, since the MM5 tutorial suggests that the Betts-Miller scheme is appropriate for grid scales larger than 30 km (see NCAR web site:* http://www.mmm.ucar.edu/mm5/documents/MM5_tut_Web_notes/MM5/mm5.html *Chapter 8.3.1). The Grell scheme would probably be a suitable option if the more sophisticated KF2 scheme is rejected.*

The reviewers raise several important points in this comment. The Draft Sensitivity Study

Document was originally intended to be read as part of a series of reports on the ongoing work of ENVIRON and UCR to identify a more optimal MM5 configuration for the 2002 WRAP Annual Run. However, the introduction to the Draft Sensitivity Study Document does not make this clear, nor does it make clear the role of the refinement of the 36-km simulation in producing improved boundary conditions for the 12-km grid, which may ultimately have a different physics setup than the 36-km grid. MM5's 36-km performance in the original WRAP 2002 Annual Run had some critical shortcomings, so we are taking an incremental approach to address the 36-km performance issues first and then address the 12-km issues. We have integrated the results of new 36-km and 12-km sensitivity studies into an updated version of the Modeling Protocol, "Draft WRAP 2002 Modeling Protocol," which is available from the WRAP website http://pah.cert.ucr.edu/aqm/308/mm5.shtml". The Modeling Protocol details the final 36-km and 12-km configurations, and also contains background information (METSTAT, domain definition, etc.), which are not given in the Draft or Revised Sensitivity Study Documents.

With regard to the reviewers' question on cumulus parameterization, we emphasize that all of the sensitivity studies in which the Betts-Miller scheme was used were at a grid resolution of 36 km. We are aware of no counter-indications for using the Betts-Miller scheme at this resolution. Although the Betts-Miller scheme was originally developed for use in large-scale models and does not have a parameterization of convective downdrafts, it gave better performance in our tests than the Kain-Fritsch I, Kain-Fritsch II, and Grell schemes.

We have also run additional tests to evaluate the effect of different cumulus parameterizations on the 12-km domain. At the 12 km scale, the problem of cumulus-parameterization is not well-posed, as there is no clear spectral gap between the resolved grid-scale process and the scale of the parameterized process (Arakawa and Chen 1987). Molinari (1993) suggests that parameterization of cumulus convection for grid sizes of 2-20 km cannot be addressed with either the fully explicit method (i.e. no cumulus parameterization) or the hybrid parameterization approach. However, our ultimate goal of performing CMAQ visibility modeling requires us to be pragmatic and configure MM5 in a physically reasonable way that produces the most accurate representation of the 2002 meteorology. Our plan for the selection of the 12 km grid cumulus scheme, therefore, was to conduct several sensitivity tests to determine which cumulus scheme (or no cumulus scheme at all) gives the best performance in terms of rainfall, surface temperature, humidity and wind.

It turns out that the using no cumulus scheme at all on the 12-km grid leads to the best performance overall for these variables. We recognize that this result could be serendipitous, i.e., leading to a better answer for the wrong reason, however the alternatives are obviously worse. Our sensitivity tests suggest that getting "acceptable" MM5 performance (i.e., within benchmark envelopes) for all variables in the western U.S. is probably beyond the skill of the MM5 model; the final configuration is probably as close to optimal as can be achieved within the time constraints of this project.

2) *The VISTAS modeling didn't find any pervasive problem with July 2002 temperatures over the WRAP region as a whole. BAMS showed in their annual modeling report that a significant artificial cold temperature bias exists over the western states if the elevation differences between observation sites and the model are not considered. It is recommended*

that this effect be accounted for in some manner prior to statistical evaluations. For more information on the VISTAS methodology, see pages 19-20 and especially Figure 17 on page 21 of the VISTAS 2002 Annual Simulation documentation at: http://www.baronams.com/projects/VISTAS/reports/VISTAS_TASK3f_final.pdf

We agree that complex terrain introduces an artificial temperature bias that can confound attempts to compare modeled and observed surface temperatures in the west. The resolved MM5 terrain averages over the real terrain so that in the Western U.S. (where most weather stations are in valleys) the gridded MM5 terrain can be much higher than monitoring station elevations. In the summer, during the day, there is a modeled cold bias because there is adiabatic cooling with height and so predicted surface temperatures are cooler than the observations. The BAMS/VISTAS methodology was to simply increment the temperature predictions according to the height differential between the local resolved terrain and the monitor elevation. At night, or during winter, however, the temperature profile changes. In these situations, there may be a temperature inversion, so that the valleys should be cooler than the predictions. So there is no simple temperature adjustment to account for model verses monitor elevation differences that will always work. We feel that adding a temperature correction to METSTAT requires careful consideration to develop a robust approach that can be used for any location at any time. Possibly a more direct approach for the temperature evaluation is to relax our error and bias performance benchmarks by a degree or more for analysis regions in complex terrain.

3) *The VISTAS sensitivity modeling also showed that nudging t/q in the boundary layer leads to excessive precipitation in the summer, but better surface statistics (see VISTAS documentation pg 73 at:* http://www.baronams.com/projects/VISTAS/reports/VISTAS_TASK2e_draft.pdf*). We judged the precipitation degradation to outweigh the temperature/moisture improvements. Given the all-encompassing effects of too much precipitation, the cumulus and LSM tests would be better served operating on a system with no surface nudging of t/q. The text states that the run with no surface nudging of t/q had little effect on precipitation, but the plots were not shown. Given the apparent differences between these results and the VISTAS results, it would be good to show those plots.*

We are incorporating the reviewer's suggestion regarding the removal of t/q nudging in the boundary layer into our MM5 configuration for the final 2002 annual run. Precipitation plots for the FDDA tests have been added to the Sensitivity Study Document.

4) *Speaking of precipitation, the scale for the accumulated precipitation plots is not well suited for light to moderate accumulations of precipitation. The scale as shown highlights the heaviest precipitation accumulated over a 5-day period. One-quarter of an inch (0.25 in.) of rain is not inconsequential, yet the scale treats it exactly like it does 0.00 inches of rain (white on color scale). QPF is notoriously challenging for a model, and most precipitation statistics are calculated based on measurable precipitation, which is 0.01 in. This can be easily remedied by using a non-linear scale.*

We have re-plotted the precipitation fields to better show light precipitation amounts.

5) *It appears there are no precipitation statistics generated in this study. We recommend doing more than just a qualitative evaluation for precipitation.*

We agree with the reviewers that this is worthwhile, and are looking into methods of quantitative precipitation analysis.

6) *Despite the text indicating good agreement between the observed and estimated January precipitation, Figures 2.9.a - 2.9.b clearly demonstrate a temporal mismatch between the model fields and the observations. It almost appears as if there is a ~24 hour mismatch.*

The April precipitation observations plot was for the wrong time period. This has been corrected in the revised document.

7) *Add an image of your 12 km domain.*

An image of the 12-km domain is included in the revised modeling protocol.

Response to "Peer review of Draft Report 2002 Annual MM5 Simulation to Support WRAP CMAQ Visibility Modeling for Section 308 SIP/TIP: MM5 Sensitivity Simulations to Identify a More Optimal MM5 Configuration for Simulating Meteorology in the Western U.S." dated November 2, 2004.
Tanya Otte, US Environmental Protection Agency, Office of Research and Development (EPA/ORD)

… The MM5 sensitivity study was performed exclusively using 36-km horizontal grid spacing, but the authors very weakly imply that the final MM5 configuration of the 36-km simulations will also be applied on a 12-km domain that focuses on the western United States. …

… These model options (that were ultimately chosen) are among those that could be considered "reasonable" choices for the annual simulation at 36-km. However, the justification used to select these options via the sensitivity studies is not rigorous or thorough, and it is only mildly convincing. Furthermore, the assessment of the relative performance of the MM5 configuration at 36-km is not analogous to 12-km, particularly in areas of complex terrain. …

The reviewer raises several important points in this section. The Draft Sensitivity Study Document was intended to be read as part of a series of reports on the ongoing work of ENVIRON and UCR on finding a more optimal MM5 configuration for the 2002 WRAP Annual Run that are available on the WRAP website (http://pah.cert.ucr.edu/aqm/308/mm5_reports04.shtml). However, the introduction to the Draft Sensitivity Study Document does not make this clear, nor does it make clear the role of the refinement of the 36-km simulation in producing improved boundary conditions for the 12-km grid, which may ultimately have a different physics setup than the 36-km grid. MM5's 36-km performance in the original WRAP 2002 Annual Run had some critical shortcomings, so we are taking an incremental approach in which we address the 36-km performance issues first and then address the 12-km issues. We have integrated the results of new 36-km and 12-km sensitivity studies into an updated version of the Modeling

Protocol, "Draft WRAP 2002 Modeling Protocol," which is available from the WRAP website http://pah.cert.ucr.edu/aqm/308/mm5.shtml". The Modeling Protocol details the final 36-km and 12-km configurations, and also contains background information (METSTAT, domain definition, etc.), which are not given in the Draft or Revised Sensitivity Study Document.

Applicability at 12-km: ENVIRON and UCR suggest in their letter requesting a review of the Draft Report (from WRAP Modeling Forum to Fifth Generation Mesoscale Model (MM5) Users, dated 29 September 2004) that "The RMC team … plan[s] to use this proposed MM5 configuration] for the final 2002 base year simulation that will be used as input to CMAQ for base year and 2018 projection year 36-km/12-km simulations, pending comments from the peer review." However, it is only weakly implied in §1.0 that the results of the 36-km sensitivity studies in the Draft Report will also be applied at 12-km. ENVIRON and UCR suggest in multiple places in the Draft Report that the influence of the topography in the western United States is largely subgrid-scale at 36-km (and, thus, not well captured by MM5). Although it is implied in the Draft Report that the 36-km simulations are still inadequate for simulating the western United States, it is unclear from the Draft Report whether or not a 12-km simulation for the western United States would be sufficient, even if the 36-km simulations have been improved. As stated above, the results at 36-km are not necessarily analogous at 12-km, and it is fairly common for MM5 users to have different model configurations at 36-km and 12-km, particularly with regard to the FDDA configurations (cf. Stauffer and Seaman, J. Appl. Meteor., 1994; Hogrefe et al., Atmos. Environ., 2001). The sensitivity studies of the primary model configurations should be repeated with the 12-km domains to verify that the MM5 options selected generate the desired quality of meteorological fields.

As stated above, we are restructuring our series of reports so that the function of the 36-km sensitivity study is clearer and the physics configuration for the 12-km simulation is discussed in a separate section. Since the writing of the 36-km Draft Sensitivity Study Document, we have carried out a series of 12-km sensitivity tests, the results of which are reported in the Modeling Protocol and the supporting 36-km and 12-km Revised Sensitivity Study Documents.

The justification for not using the soil moisture nudging option with the Pleim-Xiu land- surface model is weak. The authors state in §1.0 of the Draft Report that soil moisture nudging "was not used based on poor performance reported by IDNR and LADCo" with no references provided to back this statement. The modeling results achieved by IDNR and LADCo for their region(s) of interest will not necessarily translate into similar performance for the western United States (for example, compare Figures 3-1 and 3-4 in the Draft Report). Furthermore, the authors also state in §1.0, "We found that removal of soil moisture nudging improved temperature and humidity performance for the short summertime tests," again without providing a reference or additional information to back this statement. While it is plausible that less favorable results were achieved for the western United States when soil moisture nudging is used, Pleim and Xiu (J. Appl. Meteor., 2003) show that the soil moisture nudging can be advantageous. The authors should consider providing some substantiation for deviating from the preferred method for running the land-surface model, or include a sensitivity test that demonstrates that the simulations are improved without the soil moisture nudging. If a sensitivity test is performed, note that the appropriate spin-up time should be used (see Pleim and Xiu 2003).

A sensitivity test showing that soil moisture nudging degrades performance was performed as part of an earlier series of tests and is described in the Draft WRAP 2002 Modeling Protocol. References to other studies showing that soil moisture nudging does not improve performance are provided as well. This topic has been discussed at great length and detail for several past years at the annual Ad-hoc Meteorological Modeling Group meetings, at which EPA modeling representatives have been in attendance. The finding that removal of soil moisture nudging improves performance has been reported by LADCo and IDNR, among several other groups, at these meetings.

The authors state in §1.0 of the Draft Report that "RMC opted to use the Reisner II mixed-phase cloud microphysics package according to suggestions from EPA/ORD," without providing a reference. While Reisner II can be considered among the state-of-the-science microphysics packages in MM5, it has provided inferior results when compared to simulations that use the (not obsolescent) Reisner I scheme (R. Dennis, personal communication, 2004). The authors need to provide a reference for where in EPA/ORD the suggestion originated (either published document or personal communication), or find a different justification for using the scheme without sensitivity.

The new versions of CMAQ (v4.4) and MCIP (v2.3) now allow the graupel mixing ratio to be passed from MM5, so we chose Reisner II in order to take advantage of this new capability in future visibility modeling. Furthermore, this was strongly suggested to us by several EPA representatives at past National RPO meetings. Our concerns presently, as documented in both Draft and Revised Sensitivity Study Documents, is the over prediction of precipitation by the sub-grid cumulus parameterizations and associated effects on temperature and humidity performance. These problems far outweigh any issues associated with sensitivity to the choice of explicit grid-resolved microphysics packages. Given a tight project schedule, we have chosen certain MM5 physics options by default without further investigation.

The nudging weights should be mentioned for both the surface and 3D and for each variable, as well as whether or not those weights will change at 12-km. Also, it is unclear which observations are being nudged with the observation nudging technique, and whether or not those same observations were used to verify the model performance. The authors should mention how the FDDA configuration will change at 12-km. Unless significant modifications were made to MM5, water vapor mixing ratio is nudged, not relative humidity. The authors should consider discussing their nudging configuration with D. Stauffer at Penn State to increase their confidence.

Nudging weights for the 36-km and 12-km grids are listed in the Modeling Protocol. The reference to the nudging of relative humidity was an error and has been replaced throughout the Revised Sensitivity Study Document, with references to nudging of water vapor mixing ratio. We have been using MM5 for nearly 10 years, and during that time have directly consulted with Dr. Nelson Seaman at Penn State extensively and in all facets of MM5 application.

The authors should at least mention which radiation option was selected. The authors should consider updating the frequency of the radiation calculations as the horizontal grid spacing decreases.

The radiation scheme selection is discussed in the Modeling Protocol.

The authors state in §2.0 that the "observed precipitation data set came from the CPC … on a grid which covers the U.S. mainland at a resolution of 0.25° x 0.25° … [which] has a reasonably high resolution, which is especially important when it comes to resolving the effects of orography on precipitation over the western U.S." A precipitation data set with that resolution could be considered quite coarse. The authors could consider using the 4-km National Precipitation Analysis (NPA).

The CPC data set has sufficiently high resolution for our purpose here, which is qualitative evaluation of the precipitation field. This data set has also by been used for this purpose by other researchers, for example, Olerud and Sims (2003). For future work, the NPA dataset would also be useful, particularly when we move to a more quantitative evaluation of precipitation performance. We appreciate this suggestion.

Although Figure 3-4 focuses on the "regions" in the eastern United States, there is little else in the Draft Report that suggests that the model performance is reasonable there. The authors could consider expanding other plots and tables to show that the performance in the east is not degraded during other periods of the year.

The performance of MM5 in the 2002 Annual WRAP simulation in areas outside the west is discussed in the document "2002 Annual MM5 36 KM Simulation to Support WRAP CMAQ Visibility Modeling for the Section 308 SIP/TIP: Preliminary Report on the Initial 2002 36 km MM5 Simulation Performed during 2003" available from the WRAP website http://pah.cert.ucr.edu/aqm/308/mm5.shtml. After the initial MM5 2002 Annual Run, further sensitivity studies were undertaken in order to improve performance in the west, with an emphasis on the 12-km WRAP domain. A section of the revised report is included to show that in the process of achieving better performance in the west through different model configuration we have not degraded the performance in the east.

All of the references listed in §4.0 are very gray literature with no refereed scientific publications shown. For the gray literature, the authors should minimally consider listing where the project reports can be obtained (if they are available online or by request). The authors should also consider using published scientific works as references, where appropriate.

In the Revised Sensitivity Study Document and the Modeling Protocol, we list the address of the web site where project reports can be obtained.

1) *Exec. Summary: For those who are not familiar with "past work", the authors could explain what the "initial 2002 36-km WRAP simulation" is and where to read about it.*

Context for this report has been addressed above.

2) *Exec. Summary: The opening paragraph makes it sound like MM5 is a bad model. The authors should be careful not to shoot themselves in the foot here.*

We do not intend to imply that MM5 is a "bad" model, however, its performance over the western U.S. in the initial 2002 Annual Simulation has some deficiencies that must be addressed before the MM5 meteorological fields can be supplied to CMAQ. We believe that a clear-headed assessment of MM5's performance, with its strengths and weaknesses, is not an indictment of the model or of its suitability for our intended purpose.

3) *Exec. Summary: The 3rd paragraph should be clarified to suggest that the bulleted items apply to simulations that were compared for brief modeling periods to the aforementioned initial run.*

This suggestion has been incorporated.

4) *p. 1-1: MM5 is not maintained by PSU. It is only maintained by NCAR.*

This has been changed.

5) *p. 1-1: EPA does not own SMOKE; CEP does.*

This has been changed.

6) *p. 1-1: SMOKE is the Sparse Matrix Operator Kernel Emissions system.*

This has been changed.

7) *p. 1-1: The authors should explicitly state in the 3rd paragraph that the results will be applied also at 12-km.*

This was addressed in our response above.

8) *p. 1-3: Why was 1-5 July 2002 selected for the sensitivity tests? Was there some meteorological justification, or was it randomly the first "summer" period? Also, since that period was particularly dry for the WRAP subregions, why was this period chosen to evaluate the convective parameterization schemes? Was there a problem with prediction of precipitation that did not verify in the previous annual run?*

The 5 day period was chosen simply because it was the first 5-day period in July, which was the month of worst performance in the initial annual 2002 run. We agree that the period was particularly dry in the west. The initial MM5 simulations documented in earlier WRAP reports do show that the convective parameterization scheme was drastically over predicting precipitation for this period. Hence, this period was ideal for studying MM5 response to alternative cumulus parameterization schemes.

9) *p. 1-3: What is the nature of the modifications to MRF?*

The MRF scheme is known to have a too high daytime PBL height bias. A review of the MRF formulation by ATMET suggested two potential improvements to the PBL scheme: 1) reduce the critical bulk Richardson number used by the scheme and 2) reduce or eliminate the "scaled virtual temperature excess" that is designed to account for a near-surface temperature that is warmer than the lowest-level model temperature. The reference for the M-MRF scheme is given below, and this reference has been added to the Sensitivity Document.

ATMET. 2003c. "MM5 Simulations for TexAQS 2000 Episode, Task 3: Sensitivities to modifications of the MRF PBL scheme, Draft Final Report." Prepared for the Houston Advanced Research Center, The Woodlands, TX, and the Texas Commission on Environmental Quality, Austin, TX, by ATMET, LLC, Boulder, CO (30 September, 2003).

10) p. 1-3: Where is Pleim-Xiu recommended by the EPA?

The Pleim-Xiu dry deposition scheme is technically superior and more current than the Wesely dry deposition scheme. In the VISTAS Phase I testing, CMAQ performed better using the Pleim-Xiu scheme than the Wesely scheme (Morris et. al 2004; VISTAS Phase I final report is available at http://pah.cert.ucr.edu/vista/docs.shtml). Personnel at EPA have stated at the annual CMAS workshop, that the Pleim-Xiu dry deposition scheme in CMAQ is the preferred approach.

11) p. 1-3: Is there a reference that states that the Pleim M3Dry model is technically superior to and outperforms the Wesely RADM scheme? Technically, M3Dry can be used in CMAQ without using the Pleim-Xiu land-surface model in MM5, although the approximation of the resistances is crude.

Addressed in the response to Specific Comment #10.

12) p. 1-4: The examination of the non-WRAP subdomains is very weak, and literally appears to be an afterthought. The evaluation of the MM5 performance for all of the sensitivities should include the entire domain (or at least the continental United States, where observational data are available), and it should not be limited to the four WRAP subdomains.

The performance of MM5 over the entire U.S. was examined in the report documenting the initial WRAP 2002 annual run. The present study focused on optimizing the MM5 configuration to produce a better simulation over the WRAP region. Once we had achieved this, we tested the configuration over the eastern U.S. to see whether the new configuration degraded performance there relative to the original run. This section was not an afterthought; it simply was not the primary focus of this work.

13) p. 2-1 (and throughout the document): Nudging-based FDDA in MM5 uses water vapor mixing ratio, not relative humidity.

This has been changed throughout the document.

14) p. 2-1: Why was Reisner I not considered?

Reisner II was used because of its ability to predict the graupel mixing ratio. Since the newest versions of CMAQ (v4.4) (available at www.cmascenter.org) and MCIP (v2.3) can accept graupel, we wanted to use a microphysics scheme that was able to supply it. Several attendees at past National RPO meetings had suggested this.

15) p. 2-2: What is the METSTAT evaluation software? Is it proprietary?

The METSTAT software was developed by ENVIRON and is freely available. METSTAT and its application to model performance evaluation are described in the Modeling Protocol.

16) p. 2-2: It would be very helpful if the authors would include a table that reiterates the relevant Emery and Tai benchmarks so they do not have to be inferred from the soccer plots.

This table is included in the Modeling Protocol along with a discussion of how the benchmarks were derived.

17) p. 2-3: Are these subdomains prescribed by METSTAT? If not, why is there so much overlap, particularly of areas 4 and 6, and of area 7? This overlap could affect the regional statistics by double-counting or triple-counting model performance (both good and bad).

Subdomains are not prescribed by METSTAT, but are instead user specified. Definitions of the METSTAT subdomains for this sensitivity study were adapted from:

Johnson, M. 2003: Meteorological modeling protocol: IDNR Annual MM5 Application. Prepared for the Iowa Department of Natural Resources Air Quality Bureau.

We do not fully understand what is meant by "double-" or "triple-counting" model performance. Since quantitative model performance is based on statistics among specific surface observation sites provided in the ds.472 dataset, sub-domain overlap might cause a few sites to be used for two or more sub-domains. However, we believe this number to be quite minimal, and thus does not significantly impact our findings.

18) p. 2-4, para. 2: The statement that singles out the Betts-Miller performance for the DesertSW region is really a stretch. The fact that all of the schemes are outside of the benchmarks for the DesertSW is not surprising given the predominant complex terrain in that region.

We believe that the data support the statement that the Betts-Miller performs best of the schemes tested, although all schemes did perform poorly.

19) p. 2-4, para. 4: The sentence that suggests that Betts-Miller is best for humidity is really over-selling the case. It is not as clear-cut as the authors suggest. Consider removing the 2 nd sentence of that paragraph entirely.

We believe the data support our conclusion regarding the Betts-Miller scheme's humidity performance. For example, the wet bias in the desertSW and north subdomains is clearly smaller in the Betts-Miller case than in the other two runs.

20) *p. 2-4, para. 5: It does not appear that Kain-Fritsch II "has a strong wet bias". Figure 2-3 shows that the SW and DesertSW regions have similar simulations for all three cases. With the color scale that was used in Figure 2-3, it is difficult to discern whether the magnitudes of the differences are significant. Furthermore, the discussion of the precipitation maximum over Texas is out of context unless the evaluation is based on the entire domain, not just the WRAP subdomains (which do not include the area of Texas with the precipitation maximum).*

The color scale has been changed. The precipitation maximum over Texas is discussed because, although not in the WRAP domain, it is the only strong convective event during this period, and is useful for evaluating how well the cumulus schemes handle intense convective rain.

21) *p. 2-5, para. 1: The logic regarding Betts-Miller performance over both the southwest and Texas is sketchy, at best. Again, if Texas is brought in for the precipitation verification, it should also matter for the other meteorological fields.*

See response to comment 20.

22) *p. 2-5, para. 2: The first sentence should be clarified to mention that the analysis was limited to four subregions, and also that it was not the best for wind.*

Changed as suggested.

23) *p. 2-5, para.2: What about the overforecasts in Georgia, South Carolina, Utah, and Florida?*

This is beyond the scope of work for this project. Precipitation performance outside the WRAP region is discussed in the report documenting the initial WRAP 2002 annual run.

24) *p. 2-7, Fig. 2-3: Would be interesting to see stats for subdomains 6 and 9…where the precipitation is.*

This is beyond the present scope of work.

25) *p. 2-7, para. 3: Should be run 1ba, not 16a.*

Changed as suggested.

26) *p. 2-8, para. 3: Consider clarifying opening sentence with "For this case" rather than "Overall", and stating that it is for the WRAP subdomains.*

Changed as suggested.

27) *p. 2-8, para. 4: Consider replacing "equivalent" with "similar".*

Changed as suggested.

28) *p. 2-8, para. 6: M3Dry can be used in CMAQ without the Pleim-Xiu land-surface model, but M3Dry is tailored for the output from P-X. Where is P-X endorsed by the EPA? Consider replacing "is endorsed by" with "was developed by".*

See response to comment 10.

29) *p. 2-12, para. 3: In the last sentence, why is temperature deemed to be more important than winds? The "terrain ... may prevent good agreement [for] winds", but complex terrain is not resolved at 36-km. Winds are still very important for transport. If anything, this seems to be an argument for higher horizontal resolution more than anything else.*

We agree that winds are important for transport and that higher resolution is necessary to resolve the terrain of the western U.S. This was one of the main motivations for using the nested 12 km grid in the WRAP annual run. We did intend to suggest that surface wind performance is unacceptable, nor that we are choosing to ignore it. One of our conclusions from the initial 2002 WRAP MM5 simulation was that wind performance quickly improves with height, and is actually quite good aloft (above the characteristic terrain scale and at altitudes associated with long-range transport) due to nudging to meteorological analyses. We conclude that surface wind performance impacts on regional transport may not be as significant as poor temperature, humidity, and precipitation performance impacts on critical chemistry and removal pathways.

30) *p. 2-16: The tests in other seasons seem to be an afterthought. Realistically, the earlier sensitivities that had been limited to July should have been extended to the other seasons.*

The full range of tests for other seasons was not feasible given the time constraints of the project. The final configuration was tested for 5-day periods in each season to investigate changes in performance; the final configuration was able to generally improve performance across-the-board in all seasons.

31) *p. 3-1, para. 5: Without a much more thorough analysis, the claims with regard to CAPE should be excluded. This is treading into very dangerous territory.*

This statement has been withdrawn.

32) *§3.0: The soccer plots for the 36-km simulation in Fig. 3-4 show that the non-WRAP subdomains can generally achieve the Emery and Tai benchmarks for wind, temperature, and humidity, but Fig. 3-1 shows that the WRAP subdomains largely cannot achieve the benchmarks for anything but humidity. Why is this? Is this acceptable? Will this also hold at 12-km? If not, will 4-km be required, or perhaps a different model configuration?*

This is addressed above and also the conclusion of the revised sensitivity document.

33) *§3.0: The title indicates that there should be a discussion of performance in the non-WRAP subdomains, yet most of the discussion is still on the WRAP subdomains.*

See response to comment 12. The title has been modified.

34) *p. 3-9, para. 2: The discussion of precipitation in the CENRAP domains first appears in the summary. This probably should have been developed prior to the summary subsection.*

Changed as suggested.

35) *p. 3-9, para. 3: The sentence about "perfect MM5 performance" is technically accurate for all domains, not just the western United States. However, explicitly stating that it is "probably beyond the skill of the MM5 model" casts doubt on using MM5 for your work. Also, "as close to optimal as can be achieved" is probably inaccurate given the amount of work the authors have put into this sensitivity exercise. Four five-day test periods are insufficient to draw those conclusions, especially considering the influence of horizontal grid spacing and the verification focus on one part of the domain. There is much more that can be done to optimize performance aside from switching model options within MM5. The authors could consider reading various papers in J. Appl. Meteor. from the past ten years that have focused on mesoscale modeling of the western United States to support air quality modeling to see what others have done. If this is infeasible, consider removing that paragraph altogether.*

This paragraph has been modified to read "as close to optimal as can be achieved within the time constraints of this project".

36) *§4.0: Typo in link for Emery et al. (2004). Also, should probably be listed as Wang (2004).*

Changed as suggested.

37) *§4.0: Olerud (2003) should probably be Olerud and Sims (2003).*

Changed as suggested.

Response to Review of "MM5 Sensitivity Simulations to Identify a More Optimal MM5 Configuration for Simulating Meteorology in the Western U.S." dated December 13, 2004. Frank Forsgren, Nevada Division of Environmental Protection, Bureau of Air Quality Planning

I want to preface my review with the following comments:

- *My modeling and meteorological experience is limited, so some of my comments may be born of ignorance.*
- *Previous MM5 sensitivity analyses were not reviewed for background.*
- *I understand this to be a stand-alone document subject to additional scrutiny by a variety of agencies.*

The reviewer raises an important point in the preface. The Draft Sensitivity Study Document was originally intended to be read as part of a series of reports on the ongoing work of ENVIRON and UCR on finding a more optimal MM5 configuration for the 2002 WRAP Annual

Run. However, the introduction to the Draft Sensitivity Study Document does not make this clear, nor does it make clear the role of the refinement of the 36-km simulation in producing improved boundary conditions for the 12-km grid, which may ultimately have a different physics setup than the 36-km grid. MM5's 36-km performance in the initial WRAP 2002 Annual Run had some critical shortcomings, so we are taking an incremental approach in which we address the 36-km performance issues first and then address the 12-km issues. We have integrated the results of new 36-km and 12-km sensitivity studies into an updated version of the Modeling Protocol, "Draft WRAP 2002 Modeling Protocol," which is available from the WRAP website http://pah.cert.ucr.edu/aqm/308/mm5.shtml". The Modeling Protocol details the final 36-km and 12-km configurations, and also contains background information (METSTAT, domain definition, etc.), which are not given in the Draft or Revised Sensitivity Study Document.

1) *Numerous of acronyms in the text, many unknown to me and not explained or identified in the text.*

Acronyms are explained in the Modeling Protocol.

2) *The MM5 sensitivity analysis was conducted for the 36-km grid (?).*

The MM5 sensitivity analysis was conducted for the 36-km grid. We address this comment in our response to the reviewers preface above.

3) *A brief explanation of how METSTAT calculates error and bias would be helpful (i.e. predicted vs observed compared for each station every hour?). Can precipitation bias/error be calculated?*

A complete explanation of how METSTAT calculates error and bias is included in the Modeling Protocol. The Revised Sensitivity Study Document now contains a reference to the Modeling Protocol. Precipitation bias and error were not calculated, but we are investigating methods for adding a quantitative analysis of the precipitation results.

4) *Explain why the NCAR ds472 surface observations data set was chosen for the model evaluation benchmark?*

An explanation of the rationale for choosing the NCAR ds472 surface observations data set for the model evaluation benchmark is given in the Modeling Protocol, which is now referenced in the Revised Sensitivity Study Document.

5) *What do the red and green areas on Figure 2-1 represent?*

The red and green areas represent overlap among the subdomains. This is how the PAVE visualization software represents such areas by default.

6) *Explanations of some general modeling protocols would be helpful, particularly the nudging methodologies employed (i.e. soil nudging, surface nudging, nudging aloft, and observation nudging), the FDDA schemes tested, and the link between FDDA and nudging.*

Explanation of some general modeling protocols is given in the Modeling Protocol referenced above. The nudging methodologies employed (i.e. soil nudging, surface nudging, nudging aloft, and observation nudging) and the link between FDDA and nudging are all discussed in the Modeling Protocol.

7) *The document suggests wind performance improves quickly with height, representing regional transport speeds/directions. Can the document include some support of this? This is a concern for dispersion of haze components generated nearby Class I areas. Can scatter plots of the winds aloft be prepared?*

During the initial WRAP 2002 annual run, the upper air performance of MM5 was evaluated in two ways. First, MM5 upper air fields were qualitatively compared with upper air analysis charts on selected days. This amounted to a spot check, as looking at twice-daily analyses for multiple fields and levels for all of 2002 would be prohibitively time-consuming. The results of the upper air analysis are described in the initial annual run document "2002 Annual MM5 36 KM Simulation to Support WRAP CMAQ Visibility Modeling for the SECTION 308 SIP/TIP: Preliminary Report on the Initial 2002 36 km MM5 Simulation Performed during 2003", which is available from the WRAP website at: http://pah.cert.ucr.edu/aqm/308/mm5.shtml. Also described there are the results of spot checks of modeled and observed January and July upper air soundings. We do not currently have software to generate scatterplots for upper air data. However, since we are nudging the wind fields to the ETA analysis above the boundary layer, and since the results of the upper air analysis of the initial 2002 run were encouraging, this is probably unnecessary for the current application.

8) *Table 1-1 and 2-1 could be combined to present a clearer overview of the modeling runs conducted. It may be helpful to add notation indicating the initial, interim, final, and optimal configurations. Run 1bb is not shown on either table. Add column headings on Table 2-1.*

The information from Table 1-1 has been integrated into Table 2-1 to present a clearer overview of the modeling runs conducted. Notation indicating the initial, interim, final, and optimal configurations has been added. Run 1bb is now included and column headings have been added to Table 2-1.

9) *Page 2-1, first paragraph. Add states to first sentence (…in the western states are shown…).*

States was added to the first sentence.

10) *Reisner II was the only microphysics option tested (at the suggestion of EPA/ORD). Runs 1a and 3a used Simple Ice microphysics according to Table 2-1. Should these two runs be eliminated from Table 2-1 and the discussion on page 2-1 since they were not tested or should they be more fully discussed in the text? In addition, the Eta and Blackadar PBL schemes shown on Table 2-1 are not identified, tested, or discussed in the text.*

Reisner II was the only microphysics option tested. The new versions of CMAQ (v4.4) and MCIP (v2.3) allow the graupel mixing ratio to be passed from MM5, so we chose Reisner II (which predicts the graupel mixing ratio) in order to take advantage of this new capability in future CMAQ visibility modeling.

Runs 1a and 3a, which used Simple Ice microphysics, were eliminated from Table 2-1 and the discussion on page 2-1. In addition, the Eta and Blackadar PBL schemes shown on Table 2-1 were removed.

11) Some inconsistencies in the references to the various METSTAT domains (i.e. PacificNW, PacNW, Pacific Northwest).

We have fixed the inconsistencies in the references to the various METSTAT domains (i.e. PacificNW, PacNW, Pacific Northwest).

12) The y-axis on the soccer plots for wind speed is RMSE, but not for the other parameters. Please explain why RMSE is better for wind speed.

This is explained in the Modeling Protocol.

13) Page 2-7, second paragraph. Replace Run 16a with Run 1ba (i.e. ...for the Run 16ba and 3b.).

Changed as reviewer suggested.

14) Page 2-8, third paragraph. The NOAH/M-MRF configuration was compared to the PX/ACM configuration. Are the LSM and PBL for these schemes linked (i.e. is the NOAH LSM always used with the M-MRF PBL?)? Explain why the NOAH LSM wasn't tested with the ACM PBL and the PX LSM tested with M-MRF PBL. In this discussion, it would be helpful to reference the runs with the currently used annotation (i.e. PX/ACM) but also include the run number for easy reference to Table 2-1.

The PX LSM scheme must be used with the ACM PBL. This is explained in the Modeling Protocol. The run number for each LSM/PBL test is now included for easy reference to Table 2-1.

15) Several places, the first on page 2-12, first partial paragraph, state that the differences in precipitation fields are small. Provide the supporting figures to allow the reader to reach their own conclusions.

Precipitation plots from the FDDA tests are now included.

16) Page 2-16, fifth paragraph. Change at to and in last sentence (…Runs 2ae and 1bb are…).

Changed as reviewer suggested.

17) MM5 performance was optimized for summer conditions due to the incidence of haze incidents during this time. Can the document include some evidence of the occurrence of *haze incidents through the year? Anecdotal evidence suggests numerous haze incidents occur in the intermountain west due to winter temperature inversions.*

This document focuses on MM5 performance, and a discussion of the frequency of haze events is beyond the present scope of work.

18) How does MM5 treat temperature inversions important to local haze issues in the West? Which parameters are most critical to allow CMAQ to model these situations?

MM5's simulation of temperature inversions important to local haze issues in the West is discussed in the initial 2002 run results document "2002 Annual MM5 36 KM Simulation to Support WRAP CMAQ Visibility Modeling for the Section 308 SIP/TIP: Preliminary Report on the Initial 2002 36 km MM5 Simulation Performed during 2003", which is available from the WRAP website at: http://pah.cert.ucr.edu/aqm/308/mm5.shtml. In order for CMAQ to simulate the effects of inversions, it must receive accurate vertical temperature gradients from the meteorological model.

19) Section 3.0 Summary of Results. Table 3-1 and Figures 3-1 are for the modeled period in July 2002?

Table 3-1 and Figures 3-1 are for the modeled period in July 2002. This is now made clear in the text.

20) What do the highlighted values shown in Table 3-1 indicate?

Highlighted values shown in Table 3-1 indicate particularly important changes in performance statistics. This is now made clear in the text.

21) References. Second Emery reference was presented in 2004 not 2005.

Changed as reviewer suggested.