Western Regional Air Partnership (WRAP) West-wide Jump Start Air Quality Study (WestJumpAQMS)
Response to Comments by Air Quality Stakeholder Review
Document: Draft WRF Application/Evaluation dated January 20, 2012
Response-to-Comments Dated February 29, 2012

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
Comments from US Environmental Protection Agency (USEPA) Region 8: Gail Tonnesen dated 2/13/12					
1	EPA R8	--	--	It might be interesting to see a comparison with the EPA/ORD 2008 12 km annual run if time and budget permits. Dennis has the EPA data, and Dennis or Ralph might have an opinion on whether this is worthwhile.	We agree that this would be interesting, but the WestJumpAQMS WRF modeling budget is stressed by the computer requirements of the large 4 km domain and the need to perform the large runs twice to improve model performance. Scientists at Alpine are attempting to procure the 2008 EPA/ORD WRF data and to conduct a comparison to be presented at the June 2012 Ad-hoc Modeling meeting.
2	EPA R8	--	--	It might be useful to compare WRF predicted cloud cover to satellite data. I don't know if this has been done previously, or how easy or useful the comparison would be. For summer ozone episodes, getting the cloud cover correct will be important.	Some comparison of the WRF and satellite cloud cover was done under the WRAP's Western Biogenics EI project. The results of that comparison can be found at: http://wrapair2.org/pdf/MGN210_sat_vs_wrf_DecCall14.Final2.ppt. A complete description of the analysis and results will be included in the Western Biogenics EI Final Report, to be completed in March 2012. Please contact Tom Moore mooret@cira.colostate.edu for more information.
3	EPA R8	--	--	It would be interesting to see plots of WRF simulated snow cover, as we now recognize the importance of this for winter photorates. We know this is a concern for winter inversion in oil and gas production areas, but could snow also affect model predictions of PM2.5 at Class I Areas in winter?	To simulate the highly episodic wintertime ozone events will require a different configuration of WRF including much higher vertical resolution to capture the shallow cold pooling inversions (e.g., see: http://epa.gov/scram001/adhoc/mcnally2010.pdf). We will revisit looking at the WRF snow cover when we generate the CAMx snow cover land use file later in the study.
4	EPA R8	--	--	This is purely a future research question: Is there any data analysis possible for the WRF output that would characterize how well WRF represents vertical exchange between the stratosphere/free troposphere/boundary layer? This might involve looking at potential vorticity and profiles of T and RH at specific sites.	Yes there are some data analysis tools that can be done to look at this issue. For example, a classic stratospheric ozone episode has been documented for April 2006 that included unusually high ozone and low RH observations at the Gothic CASTNet monitoring site among others. But such analysis is beyond the scope of this project at this time.
5	EPA R8		--	Page 3-45: First paragraph describes CPC data as 30 km resolution: "CPC analysis is based on a 0.25 x 0.25 degree (~30 x 30 km) grid" but the next two paragraphs refer to a 40 km resolution.	Since lat/long is an irregular grid there is no one resolution that can be specified in km. But the grid resolution will roughly range from18-28 km over the U.S., so text will be changed accordingly.

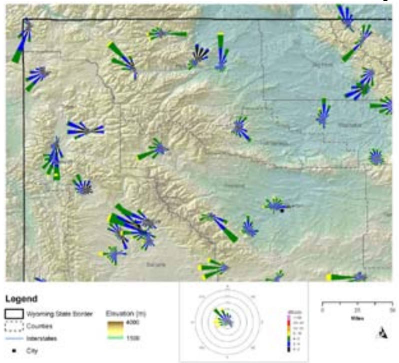
#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
6	EPA R8	4	4-1	Page 4-1: Double negative in this sentence: "Unfortunately not all of the previous model performance evaluations did not include the wind direction bias and error and wind speed RMSE calculations which were performed for this study so it is not possible to perform a comparison for winds.	Changed “did not include the” to “included.”
Comments from the Colorado Department of Health and Environment (CDPHE) APCD: Kevin Briggs dated 2/8/12.					
1	CDPHE	--	--	CDPHE/APCD pointed out several typographical errors	Typos corrected.
2	CDPHE	3.1	3-1	How about model results on those days that are conducive to high ozone or PM events which are sometimes on the extreme ends of the averages. Should this analysis be done at the regional level or at the project level?	We will revisit the WRF meteorological model performance evaluation for selected high ozone episodes and locations as part of the CAMx diagnostic ozone model performance evaluation.
3	CDPHE	3.1	3-1	Do the benchmarks for typical meteorology model performance need to be re-evaluated to include WRF model performance or have the benchmarks been updated to include the WYO WRF 2008 and ACHC WRF 2007 simulations? Are the current model performance criteria were produced using MM5 and RAMS.	Based on recent meteorological model performance evaluation in the Rocky Mountain region (e.g., WY DEQ and Denver RAQC) we are beginning to introduce the concept of the simple benchmarks for simple conditions, like the historical benchmarks based on ~30 MM5/RAMS simulation for ozone episode that are mainly flat terrain stagnation events versus complex benchmarks for complex terrain and more complex meteorological conditions (e.g., convection). As additional model simulations are completed and evaluated, there will need to be a reassessment of the benchmarks.
4	CDPHE	3.1.1	3-5	I can live with the temperature biases presented in Section 3.1.1, especially since the warm temperature bias is less in the summer (ozone season) than it is during the winter.	Comment noted.
5	CDPHE	3.1.3	3-9	Are the WRF modeled winds speeds truncated, rounded, or kept in decimal form when compared to the MADIS results?	The WRF wind speeds are kept in decimal form and there are no adjustments to them to account for this measurement artifact. A sentence has been added to the report stating this.
6	CDPHE	3.1.3	3-10	Shouldn't 33.96 degrees which is less than the 30 degree benchmark read as greater than?	Text has been changed accordingly.
7	CDPHE	3.1.3	3-10	Should there be some mention of erratic wind directions/speeds due to convective thunderstorms in the 4 km grid?	This concept has been added to the text.
8	CDPHE	3.3	3-46	With the overestimation of precipitation in the summer months simulated using WRF, is there a preferred ozone deposition scheme in CAMx/CMAQ that should be used?	Both CAMx and CMAQ have just one wet deposition scheme. The Zhang dry deposition scheme option will be used in CAMx. The precipitation overestimation is not so severe as to warrant fundamental changes to the photochemical model formulation (i.e. disabling wet deposition).

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
9	CDPHE	4	4-1	So based on the comparison of WRF and MM5 simulations, is it fair to say that WRF performs as good or better than MM5 and WRF can/should be used for future air quality studies?	Our experience is that WRF consistently performs better than MM5 for temperature and usually performs better for winds and mixing ratio. Precipitation is still a challenge for both models.
Comments from the Utah Department of Environmental Quality (UDEQ) DAQ: Lance Avery and colleagues, dated 2/16/12.					
1	UDEQ	--	--	Just wanted to drop you a note that Utah read the WRF technical memo, and it looks good. No comments or complaints here.	Comment noted.
Comments from the New Mexico Environment Department (NMED): Mark Jones and Rita Bates dated 2/17/12.					
1	NMED	--	--	NMED listed several typographical errors.	All typographical errors corrected except use of degree sign in front of the K for Kelvin.
2	NMED	--	--	Overall – the model performance appears to be adequate for the WestJump project	
Comments from the Wyoming Department of Environmental Quality (WDEQ): Ken Rairigh and Josh Nall dated 2/24/12					
1	WDEQ	2.2	2-3	The acronym for the National Oceanic and Atmospheric Administration is shown as "NOAH" and should be NOAA. Need closed parenthesis after NOAA.	Acronym has been corrected.
2	WDEQ	Global		Temperature units in Kelvin are typically referenced using the symbol (capital K) with no degree unit or degree symbols.	This is fairly standard practice and the presence of the degree symbol or nor does not change any interpretation for the reader (see: http://en.wikipedia.org/wiki/Kelvin).
3	WDEQ	3.1.1 Global	3-6	There may not be a "bright line" criteria for when (and when not) to use the complex terrain metrics for assessing model performance. However, the use of the complex terrain metrics for assessing model performance seems inappropriate when applied to states such as Oklahoma, Texas, Nebraska, which are located far from complex terrain. Perhaps there is a way to delineate such states, and not necessarily rely on the complex terrain metrics.	We present both the “simple” and “complex” model performance benchmarks in the model performance evaluation figures for each state. For “simple” states the reader can ignore the “complex” benchmarks.
4	WDEQ	3.1.1	3-6	Third sentence: the word verses is misspelled and should be versus.	Typo has been corrected.
5	WDEQ	3.1.1	3-7	The opening sentence is misleading. The use of more monitors doesn't degrade model performance, but instead indicates more areas where the model is (or is not) meeting the performance metrics. Suggest revising this sentence accordingly.	Because a majority of the observation sites in the 36 km domain are in “simple” terrain locations whereas the majority of the monitoring sites in the 4 km domain are in “complex” terrain locations we would expect better model performance across all sites in the 36 km domain than the 4 km domain.
6	WDEQ	3.1.1	3-7	The word "simplex" should be simple.	Typo corrected.

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
7	WDEQ	3.1.1	3-7	The reference to WRF is spelled as "WR" and should be WRF.	Typo corrected.
8	WDEQ	3.1.1	3-7	… all three grid resolution; add "s" to make plural.	Typo corrected.
9	WDEQ	3.1.2	3-8	… 2.0 g/kg mixing ration; remove the "n" on ration.	Typo corrected.
10	WDEQ	3.1.2	3-8	first sentence: … "IMWD region is"; change to " regions are".	Typo corrected.
11	WDEQ	3.1.3	3-10	The value of 33.96 degrees is stated as being less than the +/- 30 degree benchmark. Please revise as needed here and as needed throughout in the document.	Text has been revised as suggested.
12	WDEQ	3.1.3	3-11	… "benchmark on many months" should be revsied to "benchmark for many months".	Did not see this issue in the document.
13	WDEQ	Global		In the tables of results (e.g., Table 3-17), the top row reflects "ALL". Looking at January (as an example), the average value of 25.51 does not appear to be reflective of the values provided for the 13 states in the WRAP region. Please provide the basis used to define "ALL", and the rationale for comparing the model performance for "ALL" states.	"ALL" refers to all monitors contained in the modeling domain. The state statistics include only states either entirely, or nearly entirely, contained within the domain.. So for Table 3-17 "ALL" includes all states within the 12 km domain that includes WRAP, CENRAP and portions of MRPO states (see Figure 2-2). Rationale is that we are looking at all observations in the domain so it gives an overall indication of model performance averaged across the entire domain using all data. It is also important to note that the number of monitors varies greatly by state so the RPO average is not a simple arithmetic average of the state values.
14	WDEQ	4	4-1	This paragraph employs double negatives and is awkward; please revise as needed.	Sentence fixed.
15	WDEQ	1	1	Ist sentence needs a period at the end	Period added.
Comments from the National Park Service (NPS): Mike Barna dated 2/17/12					
1	NPS	2.1	2-2	p. 2-2: With regard to analysis FDDA, typical WRF (and, previously, MM5) simulations would employ analysis nudging at the coarsest (36km) domain, and obs nudging at the finest (4km) domain. This simulation uses analysis nudging on both the 36km and 12km domains (using 12km-resolved NAM), which suggests that the NAM data will provide a better solution than letting WRF's internal dynamics do the job. Is this true? Doesn't this overly-constrain WRF's performance at the 12km domain? Is NAM-12 used on the 12km domain since its spatial scale is the same?	Fairly weak nudging coefficients were used (e.g., $5x10^{-4}$ for winds and temperature on the 36 km domain) so we do not believe that the nudging will overly constrain WRF's solution. The nudging coefficients were further relaxed for the 12 km domain (e.g., $3x10^{-4}$) making the nudging even weaker for the 12 km domain allowing WRF's model algorithms to be the primary determination of its performance.

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
2	NPS	2.1	2-2	p. 2-2: With regard to observational FDDA, were satellite and aircraft observations retained within MADIS? In the past, NPS has found that these two data sets could cause problems with obs nudging. Also, the obs nudging coefficient seems low at 1e-4.	Just the surface wind observations in MADIS were used in the obs nudging of the 4 km WRF simulation, satellite and aircraft observations were not used. Even at 4 km resolution WRF will not be able to resolve all terrain effects in the Rocky Mountains so should not use too strong a nudging coefficient as the observed winds may be affected by subgrid-scale phenomena.
3	NPS	2.2	2-3	p. 2-3: Should "NOAH" really be "NOAA" in the second paragraph?	Yes, typo has been corrected.
4	NPS	2.2	2-3	p. 2-3: Is the model's configuration consistent throughout the simulation year, e.g., is there a different set of physics options, vertical layer definition, etc., to better resolve wintertime vs. summertime conditions?	The same WRF physics options were used for the entire 2008 year. The vertical layer definition includes shallow 12 m layers near the surface and was based on the WDEQ WRF study for both summer and winter ozone events.
5	NPS	App A	A-5-8	p. A-5 – A-8: It's somewhat discouraging that the wind speed/direction RMSE is outside the "goal" of the soccer plots for the Impact Assessment Domains, but the results in terms of model bias look encouraging.	The Impact Assessment Domains all include very complex terrain conditions, where getting low RMSE is a challenge for WRF. We were encouraged that the wind speed RMSE was right outside of the simple benchmark given the complex conditions.
Comments from the Bureau of Land Management (BLM): Susan Bassett dated 2/16/12 received 2/27/12					
1	BLM	3	3	Extending the IMWD into Montana and North Dakota will be a great help. Extension of the MT_ND IAD is also much appreciated.	We appreciate the quick feedback so we can design the modeling study to best serve BLM's needs.
2	BLM	3	7	This statement (temperature degradation due to complex terrain) may be true for some areas. However, the Northern Plains states (simple terrain) have poorer temperature performance. In Tables 3-1 and 3-2, the greatest positive temperature bias within the WRAP modeling states occurs in North Dakota and South Dakota and in the MT_SE_4km_rev and MT_ND_rev domains.	That is an interesting observation that we cannot explain at this time. The temperature performance in this region is improved when going to finer grid resolution, but the overestimation still occurs especially in the summer and some of the transition months (e.g., Apr and Sep).
3	BLM	3	10	The mean wind direction error for the 4 km domain was 33.96 degrees which is more than the ±30 degree benchmark.	Typo has been corrected.
4	BLM	3	15	Can statistics for the modeled portions of ND and SD be included in this table and the other 4km domain result tables?	Unfortunately we received this comment after we had already regenerated and updated the tables and doing that again would interfere with us meeting the hard 2/29/12 deadline for this deliverable. Hopefully we'll develop the online meteorological evaluation tool that will allow the user to drill down and address these kinds of specific performance questions.

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
5	BLM	4	5	I agree with the following statement. However, it would be useful to see some form of comparison in order to put the WestJump met modeling into perspective. "Comparisons for the 4 km domain are problematic since historically the MM5/WRF has not typically been applied to such a large 4 km domain that encompasses whole state which would facilitate intercomparison with the WestJumpAQMS 4 km WRF simulation."	Chapter 3 has lots of comparisons of the WestJumpAQMS 2008 WRF 36, 12 and 4 km simulation results with the WRF 4 km simulation usually performing better than the WRF 36 and 12 km simulation results. Because the WestJumpAQMS WRF modeling is doing something new (very large 4 km domain), there is no historical analog to directly compare with in the past. But the fact that the WestJumpAQMS WRF 36/12 km results exhibit comparable or better than past MM5 36/12 km simulations and the WestJumpAQMS WRF 4 km results are better than the WRF 36/12 km results gives some perspective.
6	BLM	4	5	In this table, WestJump WRF 2008 temperature bias performance is worse than all but one other CENRAP study. Is there a reason why Northern Plains temperature bias is high for this study? Since the other high bias study also modeled 2008, is there some reason why this baseline year was a difficult year to model these states?	We don't know the answer to the question why the WestJumpAQMS WRF temperature overestimation bias for the Northern Plains is due to unusual conditions for 2008, model physics options or other analysis. This would require further analysis. However, 2008 was an unusually wet and cool year, so it could be related to the year being modeled.
Comments from Doug Blewitt (DB) dated 2/22/12					
1	DB	--	--	**Wind Speed and Wind Direction Accuracy** The Environ report presents a comparison of meteorological performance using a previously developed statistical matrix. Unfortunately, the model evaluation approach evaluates model performance over a large geographical region (a single state) and averages modeling results over a minimum of one month. Such an evaluation procedure does not address the critical issue of how well the meteorological model performs in simulating flow for a critical air pollution episode over a sub region (smaller than a single state). Establishing the accuracy of such meteorological flow from source regions to receptor regions is critical in defining air quality impacts over large geographical regions. The WRF model performance evaluation presented in the report does not provide any information on how accurate WRF represents hourly plume transport over a large region and at specific locations. It is important that such an evaluation be conducted in addition to the model performance evaluation that is presented in the report. It is recommended that an alternate model evaluation technique be conducted by comparing predicted versus observed wind roses. This can be done by extracting hourly meteorological model results from WRF for a grid cell that contains a meteorological tower and then plot predicted versus observed hourly wind rose for that grid square. Included in these comments is a presentation I gave at the 9th EPA Modeling	We agree that more can be done in evaluating the WRF model performance and the WRF evaluation presented provides a summary of the WRF overall model performance across wide domains and down to an individual state level. The idea of comparing observed and WRF predicted wind rose at observation sites is another evaluation approach that could be done. But given there are thousands of MADIS observation sites across the WRF domain such comparisons are not feasible in a printed report. However, we are considering the development of an interactive online tool where the user could select a wind monitoring site from a map and generate observed and WRF predicted wind roses for selected time periods. The tool could also interact with the METSTAT tool and calculate wind speed and direction, temperature and mixing ratio time series and statistical comparisons for a user selected site and time period. This would allow the user to drill down into the WRF simulation and examine model performance for subregions and subperiods of the year.

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
				Conference regarding this analysis technique and Environ provided some of the data that was used in the analysis. The suggested analysis technique is very useful because it indicates that the meteorological model is providing the correct flow in the region which is critical to correctly defining source receptor relations in air quality modeling. Figure 1 presents locations where meteorological data are currently being collected in Wyoming along with the annual wind rose for each site. As indicated in this figure, this region has very complex flow and the use of a geographic monthly mean statistic to quantify meteorological model accuracy provides no useful information on how well the meteorological model can replicate what is observed in the atmosphere. Figure 1. Measured Wind Roses in Southwest Wyoming  Figure 2 presents predicted versus observed wind roses for the Jonah, WY meteorological tower for 36, 12 and 4 km grids (MM5). The 36 kilometer grid MM5 run was performed by Earth Tech as part of the SWWYTAF analysis. The 12 kilometer and 4 kilometer MM5 analyses were conducted by Environ as part of the Jonah and CDC EISs. Figure 2 presents the significant differences in predicted winds for alternate grid resolution.	The current WestJumpAQMS scope of work, budget and schedule does not include the development and implementation of such an online tool. But such a development is attractive and would be an overall benefit to the user community so will pursue funding for it. An important element of the WestJumpAQMS project will be a lessons learned and recommendations so we welcome these types of comments and will address them in subsequent WestJumpAQMS deliverables.

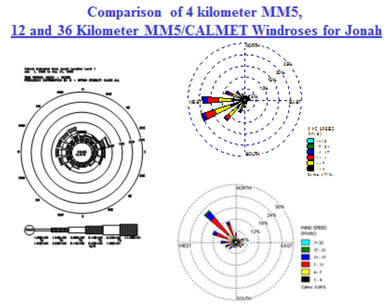
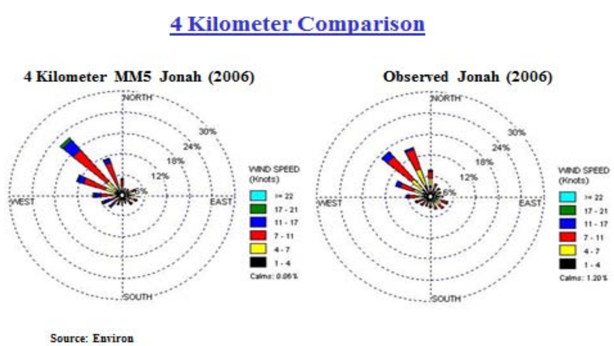
<table>
<tr><th>#</th><th>Commen-ter</th><th>Sec-tion</th><th>Page</th><th>Draft WRF Report Comment</th><th>Response for Final WRF Report</th></tr>
<tr><td></td><td></td><td></td><td></td><td>


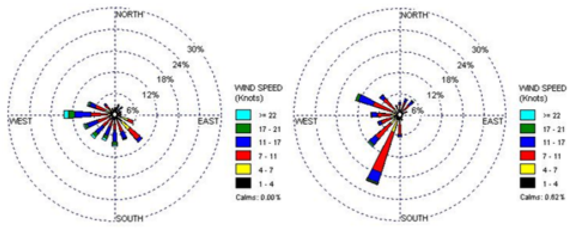
Figure 2.

Figure 3 presents a comparison of 4 kilometer MM5 results for the Jonah tower and the agreement between predicted and observed is quite good. The MM5 analysis was nudged with all available local meteorological data. Figure 4 presents 4 kilometer MM5 results for the Wamsutter meteorological tower. Again, local nudging was used; however, the predicted wind rose is different than the observed wind rose.

Another point that needs to be stressed is that both the 12 kilometer and 4 kilometer MM5 modeling achieved the performance objectives outlined in the WRF WRAP report. While the performance objectives are useful, they do not provide a complete picture of prognostic model accuracy and may lead to inaccurate plume transport in air quality modeling.

Figure 3


</td><td></td></tr>
</table>

#	Commen-ter	Sec-tion	Page	Draft WRF Report Comment	Response for Final WRF Report
				Figure 4 **4 Kilometer Comparison** **4 Kilometer MM5 4 Wamsutter (2006) Observed Wamsutter Monitor (2006)**  Source: Environ It is imperative that the WESTJump project spends the resources to provide a more thorough evaluation of WRF model accuracy and may need to consider nudging WRF using additional local surface data.	
2	DB	--	--	More detail is needed in the report regarding how the model evaluations of WRF were conducted. Specifically, what surface stations were used to evaluate model performance within a region? In addition, how was the comparison performed? These are important issues that must be discussed in the report.	The quantitative evaluation of the WRF using surface meteorological data was performed using ENVIRON's METSTAT tool. Sentences have been added on how METSTAT works and where it can be downloaded from. To address the locations of the MADIS surface meteorological observation sites, we have added three figures to the WRF report showing their locations in the 36, 12 and 4 km WRF modeling domains.
3	DB	3.1	3-5	It is recommended that Section 3 (Page 3-5) provide the mathematical definitions for performance objectives.	Mathematical formulas for bias, error and RMSE have been added to the report.
4	DB	2.1	2-2	A more complete discussion of what data were used to nudge WRF is needed. Also, were the meteorological data used to nudge the WRF modeling results also used in the model evaluation?	The WRF nudging (FDDA) approach and description is provided on page 2-2. It provides the nudging datasets used, when are where nudging was performed, type of nudging used and the nudging coefficients so is a complete description. The MADIS surface winds used to nudge the 4 km WRF surface winds were also used in the quantitative evaluation. However, with the relatively weak nudging coefficients, we do not feel that the performance metrics are overly influenced by the inclusion of the observations.