



PM_{2.5} SIP

Appendix G

Model Performance Evaluations

{This page left blank for printing purposes}

Appendix G.1

WRF Model Performance Evaluation

{This page left blank for printing purposes}

Date

June 2016 (Rev. 1/18)

Prepared for:

Allegheny County Health Department

Prepared by:

Ramboll Environ US Corporation

Lynnwood, Washington

Novato, California

Project Number:

06-35892A

ALLEGHENY COUNTY HEALTH DEPARTMENT

PM_{2.5} STATE IMPLEMENTATION PLAN

FOR THE 2012 NAAQS

WRF MODEL PERFORMANCE EVALUATION

CONTENTS

1.	INTRODUCTION	2
1.1	Meteorological Modeling	2
1.2	ACHD Meteorological Modeling	2
2.	WRF MODELING METHODOLOGY	3
2.1	Model Inputs and Initial Configuration	3
2.1.1	WRF Version and Options	3
2.1.2	Domain Configuration	5
2.1.3	Model Vertical Resolution	8
3.	WRF MODEL PERFORMANCE EVALUATION	9
3.1	Quantitative Evaluation Using METSTAT	9
3.1.1	Quantitative Statistics	11
3.1.2	METSTAT Evaluation Using Integrated Surface Hourly Observations	12
3.2	Qualitative Evaluation Using Vertical Profile Analysis	15
3.2.1	Pittsburgh Temperature, Dew Point, and Wind Speed Soundings	15
3.3	Qualitative Evaluation Using PRISM Precipitation	23
4.	SUMMARY AND CONCLUSIONS	28
5.	REFERENCES	29

FIGURES

Figure 1. WRF (36/12/4/1.333/0.444 km) Modeling Domains for ACHD PM_{2.5} SIP. 3

Figure 2. WRF grid d04 (1.333 km, outer red rectangle) covering the Allegheny County nonattainment area. The blue rectangle shows the usable portion of the 1.333 domain..... 6

Figure 3. WRF grid d05 (0.444 km, red rectangle) covering the southeast portion of the Allegheny County nonattainment area. The blue rectangle shows the usable portion of the 0.444 km domain..... 7

Figure 4. Terrain map of 0.444 km domain with Liberty and airport meteorological stations (red stars). 10

Figure 5. ACHD WRF METSTAT d04 Wind Direction Performance 13

Figure 6. ACHD WRF METSTAT d04 Wind Speed Performance..... 13

Figure 7. ACHD WRF METSTAT d04 Temperature Performance..... 14

Figure 8. ACHD WRF METSTAT d04 Humidity Performance 14

Figure 9. Vertical profile soundings comparing WRF to upper-air data at KPIT on January 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.17

Figure 10. Vertical profile soundings comparing WRF to upper-air data at KPIT on March 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain..... 18

Figure 11. Vertical profile soundings comparing WRF to upper-air data at KPIT on May 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain..... 19

Figure 12. Vertical profile soundings comparing WRF to upper-air data at KPIT on July 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain..... 20

Figure 13. Vertical profile soundings comparing WRF to upper-air data at KPIT on September 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain..... 21

Figure 14. Vertical profile soundings comparing WRF to upper-air data at KPIT on November 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain. 22

Figure 15. PRISM precipitation (top) and WRF precipitation (bottom) for January 2011 in the 12 km domain 25

Figure 16. PRISM precipitation (top) and WRF precipitation (bottom) for January 2011 in the 4 km domain 26

Figure 17. PRISM precipitation (top) and WRF precipitation (bottom) for February 2011 in the 1.33 km domain 27

APPENDICES

A. Additional METSTAT Plots

Figure 18. ACHD WRF Liberty Station METSTAT d04 Wind Direction Performance A1

Figure 19. ACHD WRF Liberty Station METSTAT d04 Wind Speed Performance A1

Figure 20. ACHD WRF Liberty Station METSTAT d04 Temperature Performance..... A2

Figure 21. ACHD WRF Liberty Station METSTAT d05 Wind Direction Performance A2

Figure 22. ACHD WRF Allegheny Airport METSTAT d05 Wind Direction Performance... A3

Figure 23. ACHD WRF Liberty Station METSTAT d05 Wind Speed Performance A3

Figure 24. ACHD WRF Allegheny Airport METSTAT d05 Wind Speed Performance..... A4

Figure 25. ACHD WRF Liberty Station METSTAT d05 Temperature Performance..... A4

Figure 26. ACHD WRF Allegheny Airport METSTAT d05 Temperature Performance A5

Figure 27. ACHD WRF Allegheny Airport METSTAT d05 Humidity Performance..... A5

Figure 28. ACHD WRF Allegheny Airport METSTAT d03 Temperature Performance A6

WRF MODEL PERFORMANCE EVALUATION

B. Precipitation Plots

Figure 29. PRISM precipitation total (top) and WRF precipitation (bottom) for January 2011 in the 4 km domain	B1
Figure 30. PRISM precipitation total (top) and WRF precipitation total (bottom) for February 2011 in the 4 km domain	B2
Figure 31. PRISM precipitation total (top) and WRF precipitation total (bottom) for March 2011 in the 4 km domain	B3
Figure 32. PRISM precipitation total (top) and WRF precipitation total (bottom) for April 2011 in the 12 km domain	B4
Figure 33. PRISM precipitation total (top) and WRF precipitation total (bottom) for May 2011 in the 4 km domain	B5
Figure 34. PRISM precipitation total (top) and WRF precipitation total (bottom) for June 2011 in the 4 km domain	B6
Figure 35. PRISM precipitation total (top) and WRF precipitation total (bottom) for July 2011 in the 4 km domain	B7
Figure 36. PRISM precipitation total (top) and WRF precipitation total (bottom) for August 2011 in the 4 km domain	B8
Figure 37. PRISM precipitation total (top) and WRF precipitation total (bottom) for September 2011 in the 4 km domain	B9
Figure 38. PRISM precipitation total (top) and WRF precipitation total (bottom) for October 2011 in the 4 km domain	B10
Figure 39. PRISM precipitation total (top) and WRF precipitation total (bottom) for November 2011 in the 4 km domain	B11
Figure 40. PRISM precipitation total (top) and WRF precipitation total (bottom) for December 2011 in the 4 km domain	B12

Acronyms and Abbreviations

ACHD	Allegheny County Health Department
ASTER	Advanced Space-borne Thermal Emission and Reflection Radiometer
ESRL	Earth Systems Research Laboratory
FDDA	Four Dimensional Data Assimilation
ISHO	Integrated Surface Hourly Observations
KAGC	Surface Identifier for Allegheny County Airport
KPIT	Surface and upper-air identifier for Pittsburgh Airport
LCC	Lambert Conformal Conic map projection
LSM	Land-Surface Model
LST	Local Standard Time
MADIS	Meteorological Assimilation Data Ingest System observation archive
METSTAT	Meteorological Statistical Program
NAAQS	National Ambient Air Quality Standards
NCEP	National Center for Environmental Prediction
NCDC	National Climatic Data Center
NOAA	National Oceanographic and Atmospheric Administration
NOHRSC	National Operational Hydrologic Remote Sensing Center
NWS	National Weather Service
PBL	Planetary Boundary Layer
PRISM	Parameter-elevation Regressions on Independent Slopes Model
RMSE	Root Mean Square Error
RTG	Real-Time Global
SCAS-OSU	Spatial Climate Analysis Service at Oregon State University
SIP	State Implementation Plan
SNODAS	Snow Data Assimilation System
SODAR	Sonic Detection and Ranging
SST	Sea Surface Temperature
UTC	Coordinated Universal Time
WPS	WRF Pre-processing System
WRAP	Western Regional Air Partnership
WRF	Weather Research and Forecasting model

1. INTRODUCTION

Ramboll Environ US Corporation (Ramboll Environ) conducted meteorological modeling to provide meteorological fields for use in air-dispersion modeling within the Allegheny County, PA fine particle ($PM_{2.5}$) nonattainment area for the 2012 National Ambient Air Quality Standards (NAAQS), covering all of Allegheny County. The Weather Research and Forecasting (WRF) meteorological model (NCAR, 2015) was selected for high resolution meteorological modeling of the region for the entire year of 2011. WRF's research and operational application ensures state-of-the-science physics and adaptability to a wide range of environments, through a broad selection of physics options, allowing us to develop the best-performing configuration for simulating meteorology in the region.

1.1 Meteorological Modeling

Over the past decade, emergent requirements for numerical simulation of urban and regional scale air quality have led to intensified efforts to construct high-resolution emissions, meteorological, and air quality data sets. It is now possible, for example, to exercise sophisticated mesoscale prognostic meteorological models and Eulerian and Lagrangian photochemical/aerosol models, for multi-seasonal periods over near-continental scale domains, in a matter of weeks with the application tailored to a specific air quality modeling project.

The WRF model contains separate modules to compute different physical processes such as surface energy budgets and soil interactions, turbulence, cloud microphysics, and atmospheric radiation. Within WRF, the user has many options for selecting the different schemes for each type of physical process. There is a WRF Pre-processing System (WPS) that generates the initial and boundary conditions used by WRF, based on topographic datasets, land use information, and larger-scale atmospheric and oceanic models.

1.2 ACHD Meteorological Modeling

The ACHD study performed WRF meteorological modeling of the Allegheny, PA $PM_{2.5}$ nonattainment area using a nested-grid structure of five grids nested in a 3:1 nesting ratio (36/12/4/1.333/0.444 km) as pictured in Figure 1. Ramboll Environ performed sensitivity testing on several WRF model configurations to optimize a high-resolution, one-year dataset for use in air dispersion modeling.

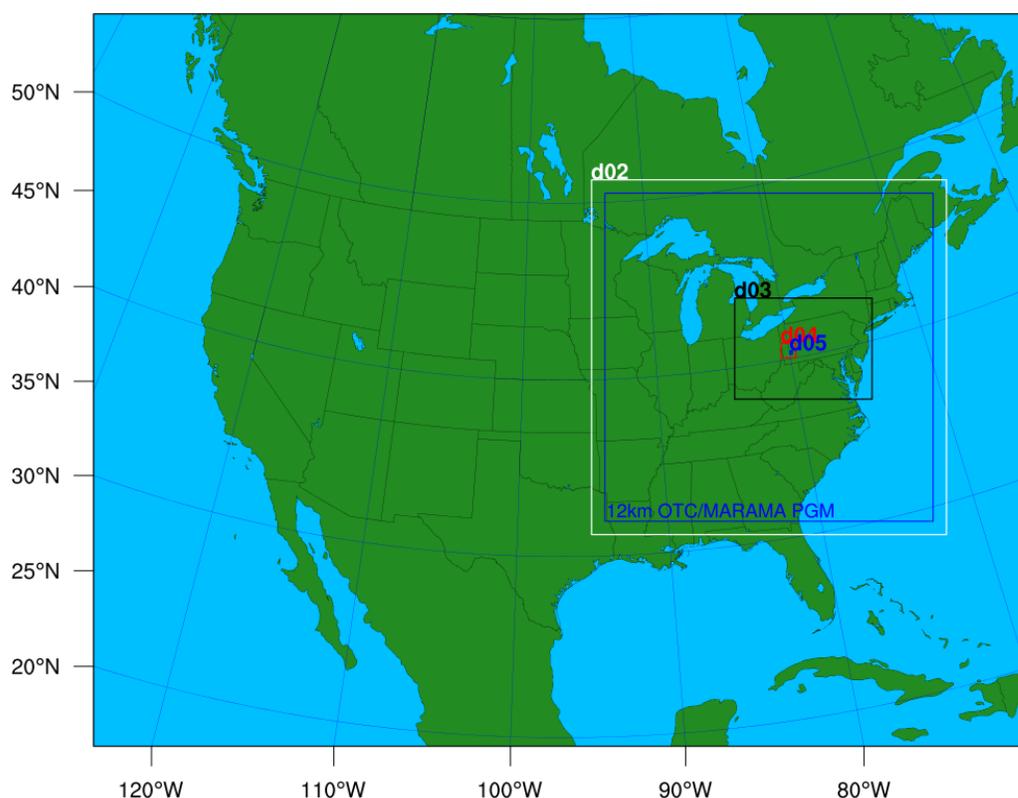


Figure 1. WRF (36/12/4/1.333/0.444 km) Modeling Domains for ACHD PM_{2.5} SIP.

2. WRF MODELING METHODOLOGY

This Section describes the methodology used for conducting the WRF simulation for the January through December 2011 modeling period, and describes the model configuration used in the final WRF dataset for the ACHD PM_{2.5} air dispersion modeling.

2.1 Model Inputs and Initial Configuration

A brief summary of the WRF configuration and input data used for this model performance evaluation is provided below.

2.1.1 WRF Version and Options

Model Selection: The publicly available version of WRF, version 3.7.1, was used for the ACHD simulation. The WPS pre-processor programs including GEOGRID, UNGRIB and METGRID were used to develop model inputs.

Topographic Inputs: Topographic information for all but the finest domain was developed using the standard WRF terrain databases. The 36 km domain was based on the 10 minute global data; the 12 km domain was based on the 2 minute data; and the 4 and 1.333 km domains were based on the 30 second data. The 0.444 km domain was based on high-resolution Advanced Space-borne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (DEM) 1 arc second ~ 90 meter data.

Vegetation Type and Land Use Inputs: Vegetation type and land use information was developed using the most recently-released databases provided with the WRF distribution. Standard WRF surface characteristics corresponding to each land use category were employed.

Atmospheric Data Inputs: WRF relies on other model or re-analysis output to provide initial and boundary conditions (IC/BC). The first guess fields were taken from the ~70 km European Center for Medium-Range Weather Forecasting Re-Analysis (ERA)-Interim dataset. These will be objectively re-analyzed using traditional observational data (meteorological towers) to the higher resolution of each WRF grid, using the OBSGRID program. These fields are used both to initialize the model, and used with analysis nudging (on selected domains) to guide the model to best match the observations.

Time Integration: Adaptive time stepping was used to maximize the time step that the model can use while keeping the model numerically stable. The model time step was adjusted based on the domain-wide horizontal and vertical stability Courant-Friedrichs-Lewy (CFL) target value of 0.8.

Diffusion Options: Horizontal Smagorinsky first-order closure with sixth-order numerical diffusion and suppressed up-gradient diffusion were used.

Lateral Boundary Conditions: Lateral boundary conditions were specified from the initialization dataset on the 36 km CONUS domain with continuous updates nested from each "parent" domain to its "child" domain.

Top and Bottom Boundary Conditions: The top boundary condition was selected as an implicit Rayleigh dampening for the vertical velocity. Consistent with the model application for non-idealized cases, the bottom boundary condition was selected as physical, not free-slip.

Water Temperature Inputs: The water temperature data were taken from the NCEP RTG daily global one-twelfth degree analysis, and was updated every 24 hours (as opposed to fixed for each WRF initialization).

Snow Cover: The model runs use the 1 km resolution snow data from the Snow Data Assimilation System (SNODAS). SNODAS is a modeling and data assimilation system developed by the NOAA National Weather Service's National Operational Hydrologic Remote Sensing Center (NOHRSC) to provide the best possible estimates of snow cover and associated variables to support hydrologic modeling and analysis. SNODAS includes procedures to ingest and downscale output from Numerical Weather Prediction (NWP) models and to simulate snow cover using a physically based, spatially-distributed energy and mass-balance snow model (NOHRSC, 2004). The product used is very similar to that used by NOAA's Rapid Refresh (RAP) assimilation/modeling system.

FDDA Data Assimilation: The WRF model was run with a combination of analysis and observation nudging (i.e., Four Dimensional Data Assimilation [FDDA]). Analysis nudging was used for winds, temperature, and humidity on the 36 and 12 km domains. The nudging used both surface and aloft nudging, but nudging for temperature and mixing ratio was not performed in the lower atmosphere (i.e., within the boundary layer). The WRF simulation used observation nudging within the 4 km and 1.333 km domains for winds, temperature, and humidity. Observation nudging was performed using the Meteorological Assimilation Data Ingest System (MADIS, 2015) observation archive. There was no nudging of the 0.444 km domain, as only two sites were inside.

Physics Options: The WRF model contains many different physics options. The physics options chosen for the WRF configuration are presented in Table 1.

Table 1. Physics Options used in the ACHD WRF Dataset

Physics Scheme	Option
Longwave Radiation	RRTMG
Shortwave Radiation	RRTMG
Microphysics	Thompson
Cumulus Parameterization	Multi-scale Kain-Fritsch in 36, 12, 4 and 1.33 km
Planetary Boundary Layer (PBL)	Yonsei University scheme (YSU)
Land Surface Model (LSM)	Noah
Surface Layer	Monin-Obukhov

Application Methodology: The WRF model was executed in 5.5-day blocks initialized at 12Z every five days. Model results were output every 60 minutes and output files were split at twelve (12) hour intervals. Twelve (12) hours of spin-up were included in each 5-day block before the data were used in the subsequent evaluation. The model was run at the 36, 12, 4, 1.333, and 0.444 km resolution from December 16, 2010 through December 31, 2011.

2.1.2 Domain Configuration

The WRF Domain configuration is comprised of a system of simultaneous nested grids. Figure 1 shows the WRF modeling grids at 36/12/4/1.333/0.444 km. All WRF grids are defined on a Lambert Conformal Conic (LCC) projection centered at 40°N, 97°W with true latitudes at 33°N and 45°N (the so-called standard Regional Planning Organization (RPO) projection). The outermost domain with 36 km resolution includes the entire continental United States and parts of Canada and Mexico. The inner 12 km regional grid covers the northeastern portion of the United States. The 4 km domain covers the four state areas of Pennsylvania, Ohio, West Virginia, and Maryland. Figure 2 pictures the 1.333 km domain with grid cells, covering multiple counties in southwest Pennsylvania, as well as the nested 0.444 km domain. Figure 3 shows the inner 0.444 km domain with grid cells, covering the nonattainment portion of Allegheny County as pictured. Table 2 provides the domain specifications used in the modeling study.

Table 2. WRF Domain Specifications

Grid Spacing (km)	Number of Points (West-East)	Number of Points (South-North)	Starting Point (West-East)	Starting Point (South-North)
36	165	129	-	-
12	187	187	88	38
4	217	160	76	72
1.333	70	70	74	66
0.444	31	46	43	19

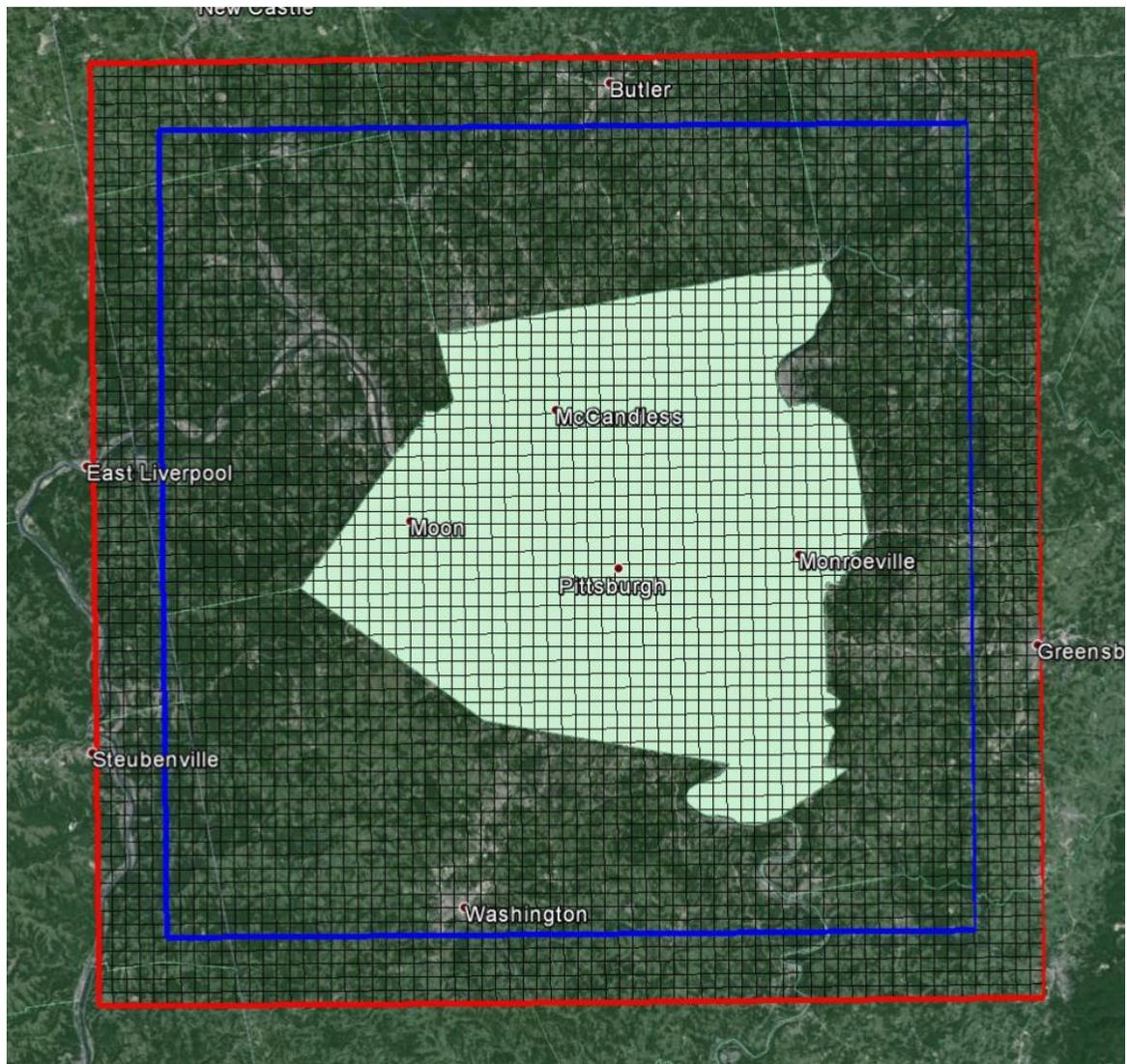


Figure 2. WRF grid d04 (1.333 km, outer red rectangle) covering the Allegheny County nonattainment area. The blue rectangle shows the usable portion of the 1.333 domain.

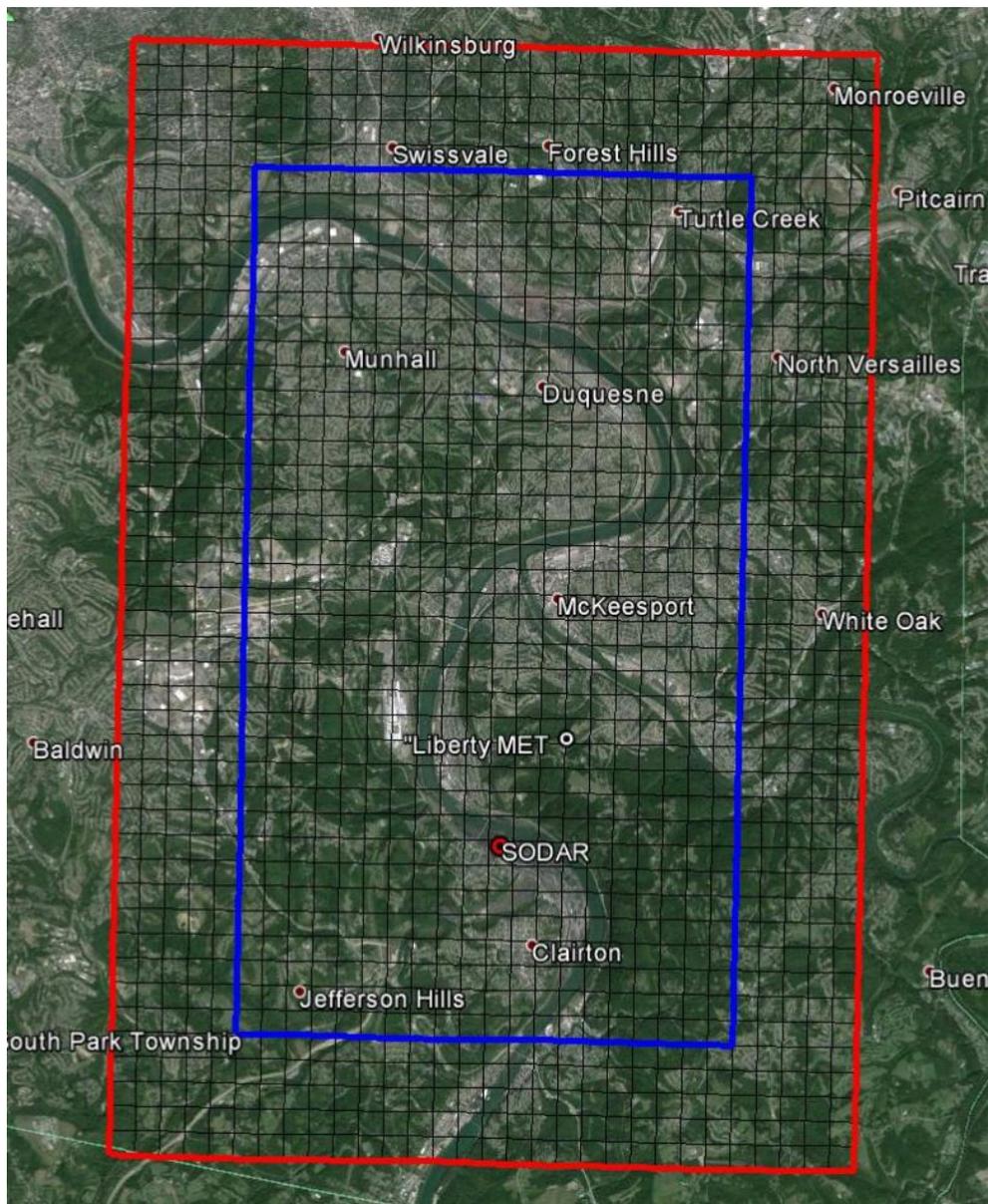


Figure 3. WRF grid d05 (0.444 km, red rectangle) covering the southeast portion of the Allegheny County nonattainment area. The blue rectangle shows the usable portion of the 0.444 km domain.

2.1.3 Model Vertical Resolution

High vertical resolution enables the model to more accurately capture the convective updraft velocities in summer and low-level temperature inversions frequently present during the fall. The ACHD WRF modeling was based on 37 vertical layers with an approximately 11 meter thick lowest layer. Table 3 illustrates the vertical layer structure used in this modeling project.

Table 3. ACHD WRF Dataset Model Levels

Level	Eta	Pressure (mb)	Height (m)	ΔZ (m)
1	1.0000	1013	0.0	
2	0.9985	1012	11.2	11.2
3	0.9970	1010	22.4	11.2
4	0.9955	1009	33.7	11.2
5	0.9940	1008	44.9	11.2
6	0.9925	1006	56.1	11.3
7	0.9910	1005	67.4	11.3
8	0.9895	1003	78.7	11.3
9	0.9870	1001	97.5	18.8
10	0.9845	999	116.4	18.9
11	0.9820	997	135.3	18.9
12	0.9795	994	154.2	18.9
13	0.9770	992	173.2	19.0
14	0.9745	990	192.2	19.0
15	0.9720	987	211.2	19.0
16	0.9690	985	234.1	22.9
17	0.9660	982	257.1	22.9
18	0.9610	977	295.4	38.4
19	0.9510	968	372.6	77.1
20	0.9360	955	489.4	116.8
21	0.9210	941	607.6	118.2
22	0.9010	923	767.4	159.8
23	0.8810	904	929.7	162.4
24	0.8600	885	1103.1	173.4
25	0.8200	849	1441.9	338.8
26	0.7600	794	1973.0	531.1
27	0.7000	739	2534.6	561.6
28	0.6000	648	3550.4	1015.8
29	0.5000	557	4689.7	1139.3
30	0.4000	465	5991.9	1302.2
31	0.3000	374	7520.7	1528.8
32	0.2200	301	8979.7	1459.0
33	0.1500	237	10514.7	1535.0
34	0.1000	191	11832.4	1317.6
35	0.0600	155	13084.9	1252.5
36	0.0270	125	14313.8	1228.9
37	0.0000	100	15513.7	1199.9

3. WRF MODEL PERFORMANCE EVALUATION

Both a quantitative and qualitative evaluation of the ACHD WRF simulation were conducted. The quantitative evaluation compared both integrated surface hourly and on-site meteorological observations with WRF predictions matched by time and location. The qualitative evaluation compared twice daily vertical profiles with upper-air data and on-site hourly SODAR profiles with WRF predictions matched by time and location. Additionally, monthly total spatial precipitation fields based on observations (PRISM data) were compared with the WRF gridded monthly total precipitation fields. Below we summarize the main features of the WRF simulation model performance evaluation.

3.1 Quantitative Evaluation Using METSTAT

A quantitative model performance evaluation of the ACHD WRF simulation was performed using integrated hourly surface and on-site meteorological measurements and the publicly-available METSTAT software (Ramboll Environ, 2015) evaluation tool. METSTAT calculates statistical performance metrics for bias, error and correlation for surface winds, temperature, and mixing ratio (i.e., water vapor or humidity). To evaluate the performance of a meteorological model simulation for air quality model applications, a number of performance benchmarks for comparison are typically used. Table 4 lists the meteorological model performance benchmarks for simple (Emery et al., 2001) and complex (Kemball-Cook et al., 2005) situations. The simple benchmarks were developed by analyzing well-performing meteorological model evaluation results for simple, mostly flat terrain conditions and simple meteorological conditions (e.g., stationary high pressure) that were mostly conducted to support air quality modeling studies (e.g., ozone SIP modeling). The complex benchmarks were developed during the Western Regional Air Partnership (WRAP) regional haze modeling and are performance benchmarks for more complex conditions, such as the complex terrain of the Rocky Mountains and Alaska (Kemball-Cook et al., 2005). McNally (2009) analyzed multiple annual runs that included complex terrain conditions and suggested an alternative set of benchmarks for temperature under more complex conditions. The purpose of the benchmarks is to understand how good or poor the results are relative to other model applications run for the United States.

In this section, Ramboll Environ compared the initial WRF meteorological variables to the benchmarks as an indication of WRF model performance. These benchmarks include bias and error in temperature, wind direction and mixing ratio as well as the wind speed bias and Root Mean Squared Error (RMSE) between the models and databases.

Table 4. Meteorological Model Performance Benchmarks for Simple and Complex Conditions

Parameter	Emery et al. (2001)	Kemball-Cook et al. (2005)	McNally (2009)	Resulting Criteria
Conditions	Simple	Complex	Complex	Complex
Temperature Bias	$\leq \pm 0.5$ K	$\leq \pm 2.0$ K	$\leq \pm 1.0$ K	$\leq \pm 1.0$ K
Temperature Error	≤ 2.0 K	≤ 3.5 K	≤ 3.0 K	≤ 3.0 K
Temperature IOA	≥ 0.8	(not addressed)	(not addressed)	≥ 0.8
Humidity Bias	$\leq \pm 1.0$ g/kg	$\leq \pm 0.8$ g/kg	$\leq \pm 1.0$ g/kg	$\leq \pm 1.0$ g/kg
Humidity Error	≤ 2.0 g/kg	≤ 2.0 g/kg	≤ 2.0 g/kg	≤ 2.0 g/kg
Humidity IOA	≥ 0.6	(not addressed)	(not addressed)	≥ 0.6
Wind Speed Bias	$\leq \pm 0.5$ m/s	$\leq \pm 1.5$ m/s	(not addressed)	$\leq \pm 1.5$ m/s
Wind Speed RMSE	≤ 2.0 m/s	≤ 2.5 m/s	(not addressed)	≤ 2.5 m/s
Wind Speed IOA	≥ 0.6	(not addressed)	(not addressed)	≥ 0.6
Wind Dir. Bias	$\leq \pm 10$ degrees	(not addressed)	(not addressed)	$\leq \pm 10$ degrees
Wind Dir. Error	≤ 30 degrees	≤ 55 degrees	(not addressed)	≤ 55 degrees

The output from the ACHD WRF simulation was compared against on-site meteorological data obtained from the Liberty monitoring station operated by Allegheny County and the National Climate Data Center's (NCDC) global-scale, quality-controlled DS3505 integrated surface hourly

observational (ISHO) data as verification data (NOAA-NCDC, 2015). Global hourly and synoptic observations are compiled from numerous sources into a single common ASCII format and common data model. The DS3505 database contains records of most official surface meteorological stations from airports, military bases, reservoirs/dams, agricultural sites, and other sources dating from 1901 to the present. Figure 4 is a terrain map of the 0.444 km domain, showing the geography of the ridgelines and river valley, with the locations of both the Liberty monitoring station and closest DS3505 station at the Allegheny Airport (KAGC).

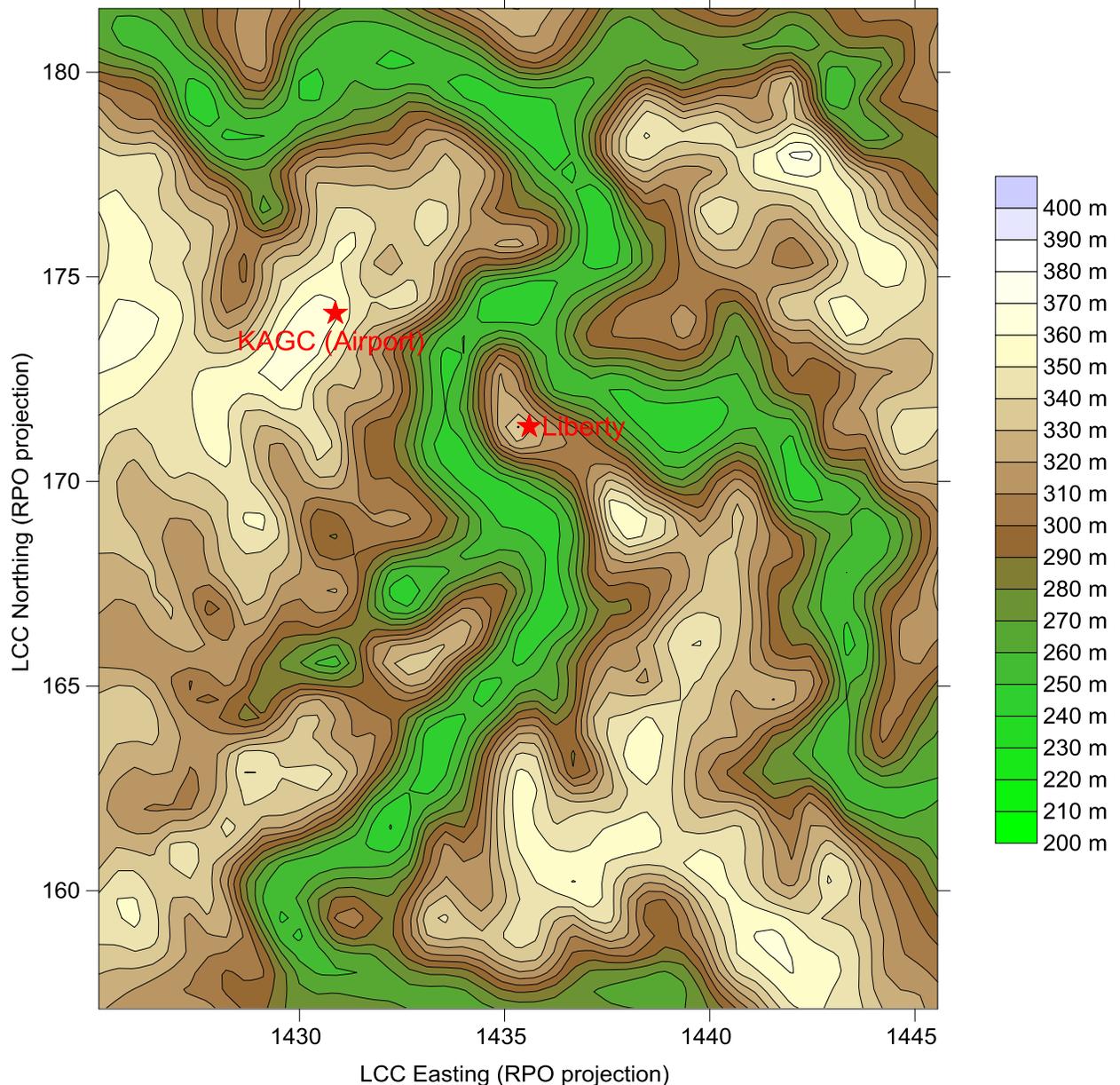


Figure 4. Terrain map of 0.444 km domain with Liberty and airport meteorological stations (red stars).

A standard set of statistical metrics from the METSTAT package was used. These metrics were calculated on hourly, daily and monthly time frames for wind speed, wind direction, temperature, and humidity at the surface, using all available observational weather data. The WRF surface meteorological model performance metrics were compared against the simple and complex model performance goals using "soccer plots." Soccer plots use two WRF performance metrics as X-axis and Y-axis values (e.g., temperature bias as X, and temperature error as Y) along with the performance benchmarks. The closer the symbols are to the zero origin, the better the model performance. These

plots also make it easy to see when the two WRF performance metrics fall within the benchmark lines. We present monthly surface meteorological model performance across the 1.333 km (d04) county area in section 3.1.2. METSTAT plots of the surface performance at the Liberty station for the 0.444 km domain (d05) are included in Appendix A.

3.1.1 Quantitative Statistics

The quantitative analysis was conducted using METSTAT. Statistical measures calculated by METSTAT include observation and prediction means, prediction bias, and prediction error that are given as follows.

Mean observation (M_o) is calculated using values from all sites for a given time period by

Eq. (1):

$$M_o = \frac{1}{IJ} \sum_{j=1}^J \sum_{i=1}^I O_j^i \quad (1)$$

where O_j^i is the individual observed quantity at site i and time j , and the summations are over all sites (I) and over time periods (J).

Mean Prediction (M_p) is calculated from simulation results that are interpolated to each observation used to calculate the mean observation for a given time period by Eq. (2):

$$M_p = \frac{1}{IJ} \sum_{j=1}^J \sum_{i=1}^I P_j^i \quad (2)$$

where P_j^i is the individual predicted quantity at site i and time j . Note the predicted mean wind speed and mean resultant direction are derived from the vector-average (for east-west component u and north-south component v), from which the mean wind speed and mean resultant direction are derived.

Bias (B) is calculated as the mean difference in prediction-observation pairings with valid data within a given analysis region and for a given time period by Eq. (3):

$$B = \frac{1}{IJ} \sum_{j=1}^J \sum_{i=1}^I (P_j^i - O_j^i) \quad (3)$$

Gross Error (E) is calculated as the mean *absolute* difference in prediction-observation pairings with valid data within a given analysis region and for a given time period by Eq. (4):

$$E = \frac{1}{IJ} \sum_{j=1}^J \sum_{i=1}^I |P_j^i - O_j^i| \quad (4)$$

Note that the bias and gross error for winds are calculated from the predicted-observed residuals in speed and direction (not from vector components u and v). The direction error for a given prediction-observation pairing is limited to range from 0 to $\pm 180^\circ$.

Root Mean Square Error (RMSE) is calculated as the square root of the mean squared difference in prediction-observation pairings with valid data within a given analysis region and for a given time period by Eq (5):

$$RMSE = \left[\frac{1}{IJ} \sum_{j=1}^J \sum_{i=1}^I (P_j^i - O_j^i)^2 \right]^{\frac{1}{2}} \quad (5)$$

The RMSE, as with the gross error, is a good overall measure of model performance. However, since large errors are weighted heavily (due to squaring), large errors in a small sub-region may produce a large RMSE even though the errors may be small and quite acceptable elsewhere.

3.1.2 METSTAT Evaluation Using Integrated Surface Hourly Observations

Figure 5 through Figure 8 presents soccer plots of WRF performance in the 1.333 km domain for all months of 2011, evaluated against all DS3505 observations within the domain. Wind direction performed well, with all months falling within the simple conditions threshold. There is a slight positive wind direction bias in the model, around 2-7 degrees, most likely from the terrain influence of the river valley. Soccer plots of wind speed performance are presented in Figure 6. WRF performs well in wind speed too, with all months again falling within the complex conditions thresholds. The model reflects only a small positive bias during the warmer months of roughly 0.5 m/s. WRF temperature performance, shown in Figure 7, was satisfactory in d04. All months fall within the complex conditions threshold for temperatures, but many exhibit a cold bias between -0.75 and 1.25 degrees K. Winter months tended to have the largest cold bias, while the spring and early summer months were the most accurate. WRF humidity performance in Figure 8 shows strong accuracy, with 9 of the months falling within the threshold for simple conditions. There is a small tendency toward positive humidity bias in the model, with all months exhibiting a positive bias. However, the largest WRF humidity biases are still less than 1.0 g/kg.

Figure 18 through Figure 20 in Appendix A display METSTAT performance plots for only the Liberty met station in the 1.333 km (d04) domain for all months in 2011. Note that the Liberty on-site station database did not contain a humidity variable. Wind direction performance by WRF at the Liberty site was strong, with about half of the months meeting the criteria for simple conditions and all months falling well within the threshold for complex conditions. There was no obvious wind direction bias at Liberty. Wind speed performance at Liberty was excellent, as all months fell within the threshold for simple conditions. There is a slight positive bias in wind speed performance, up to about 0.5 m/s in May. Temperature performance was satisfactory at Liberty, with a cold bias averaging around 1.0 K and up to about 1.75 K in February. All months fell within the temperature performance thresholds for complex conditions, and June through July temperatures were within the simple conditions threshold.

Figure 21 through Figure 26 in Appendix A compare METSTAT performance plots alternating between the Liberty met station and the Allegheny County Airport met station (KAGC) observation-based evaluations in the 0.444 km (d05) domain for all months in 2011. Wind direction performance was satisfactory for both stations, with a positive wind direction bias for all months and no seasonal trend in higher biases. One month at each station fell slightly outside of the complex conditions threshold. Wind speed performance was satisfactory at Liberty and even better at KAGC. Modeled wind speeds at both sites were within the complex conditions threshold for all months, and most of the months fell within the simple conditions threshold at KAGC. Higher wind speed and wind direction biases for the Liberty station is most likely due to the terrain channeling effects of the valley area. Temperature performance was also satisfactory with most months meeting the threshold for complex, but not simple, conditions at both sites. A consistent negative temperature bias characterized both sites, with maximum cold biases of about 2 K in February at Liberty and in November at KAGC. In Figure 27, The Allegheny County Airport humidity performance in the 0.444 km domain is similar to composite performance in the 1.333 km domain with an increased positive humidity bias, especially in the warmer months. Finally, Figure 28 shows METSTAT temperature performance by WRF in the 4 km domain, considering all meteorological station data within that domain. This plot indicates that the cold bias is limited to the smaller domains for most months. The impact of this local cold bias on WRF vertical profiles is discussed further in Section 3.2.

Overall, WRF performed well in both the 1.333 km and 0.444 km domains. The METSTAT performance benchmarks were originally developed involving statistics averaged over a large number of surface observations sites. Considering the limited amount of integrated surface hourly stations within the higher-resolution domains (four in the 1.333 km domain, one for each of the 0.444 km domain datasets) the results are acceptable for the Study area.

Allegheny County Annual PM2.5 SIP d04 Wind Direction Performance FINAL - all

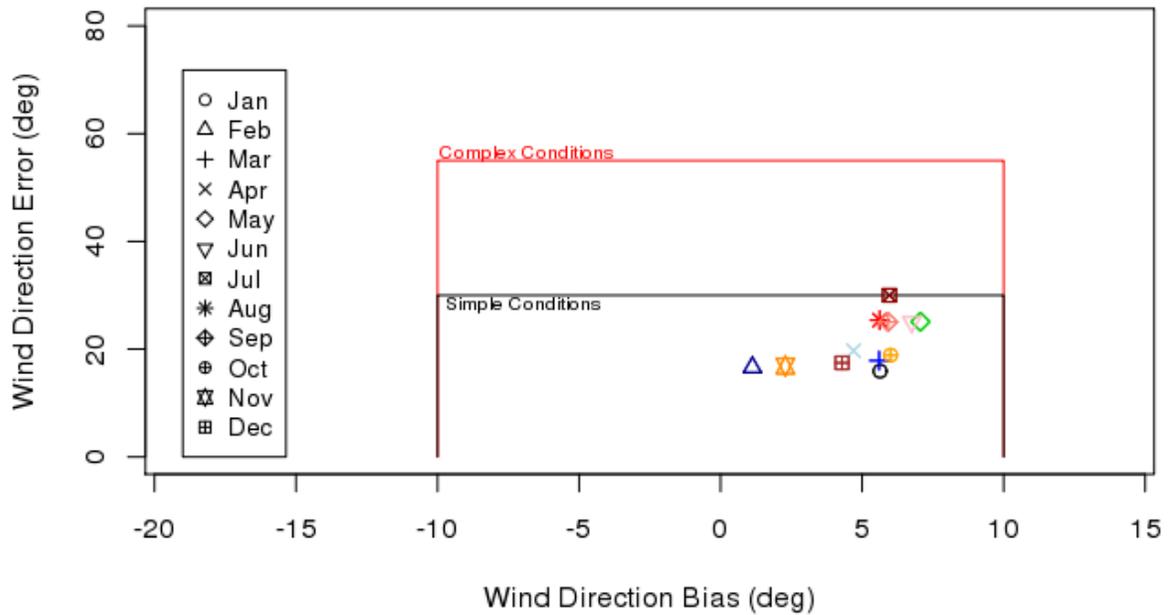


Figure 5. ACHD WRF METSTAT d04 Wind Direction Performance

Allegheny County Annual PM2.5 SIP d04 Wind Speed Performance FINAL - all

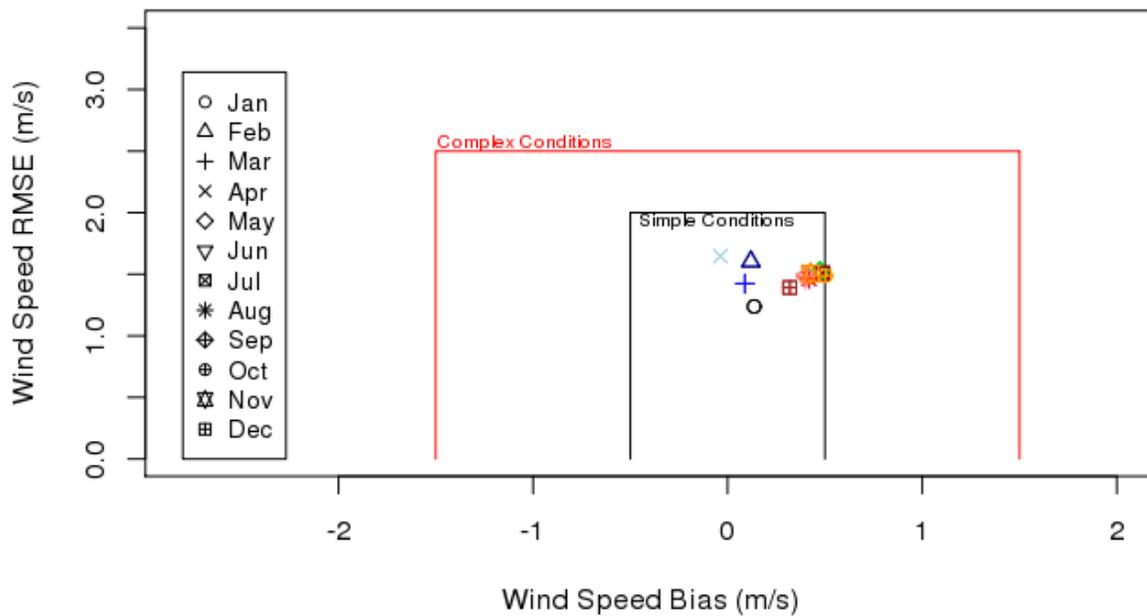


Figure 6. ACHD WRF METSTAT d04 Wind Speed Performance

Allegheny County Annual PM2.5 SIP d04 Temperature Performance FINAL - all

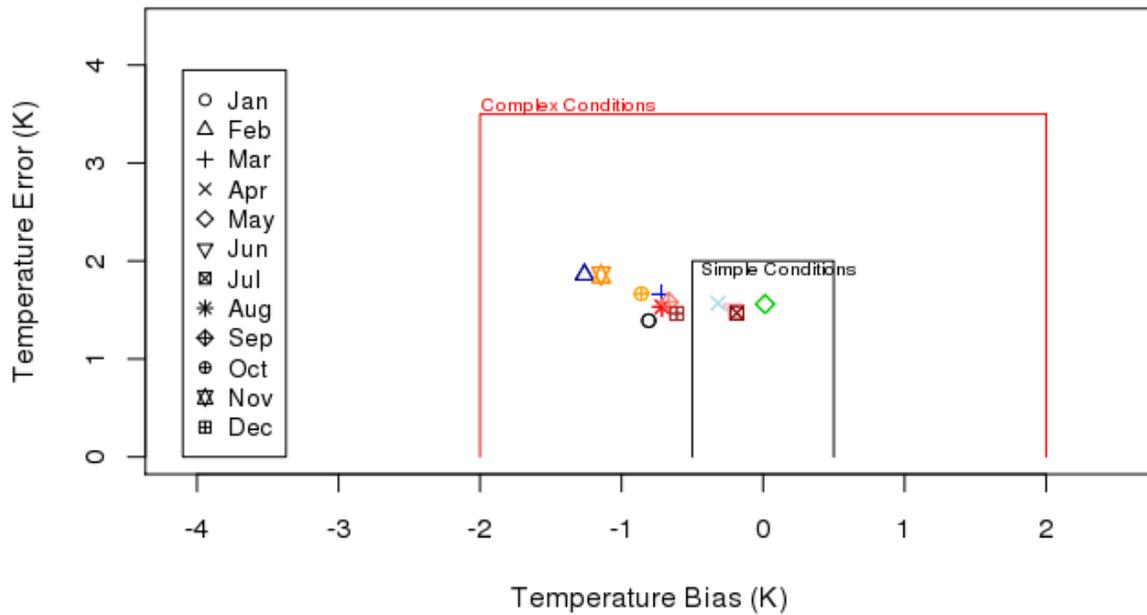


Figure 7. ACHD WRF METSTAT d04 Temperature Performance

Allegheny County Annual PM2.5 SIP d04 Humidity Performance FINAL - all

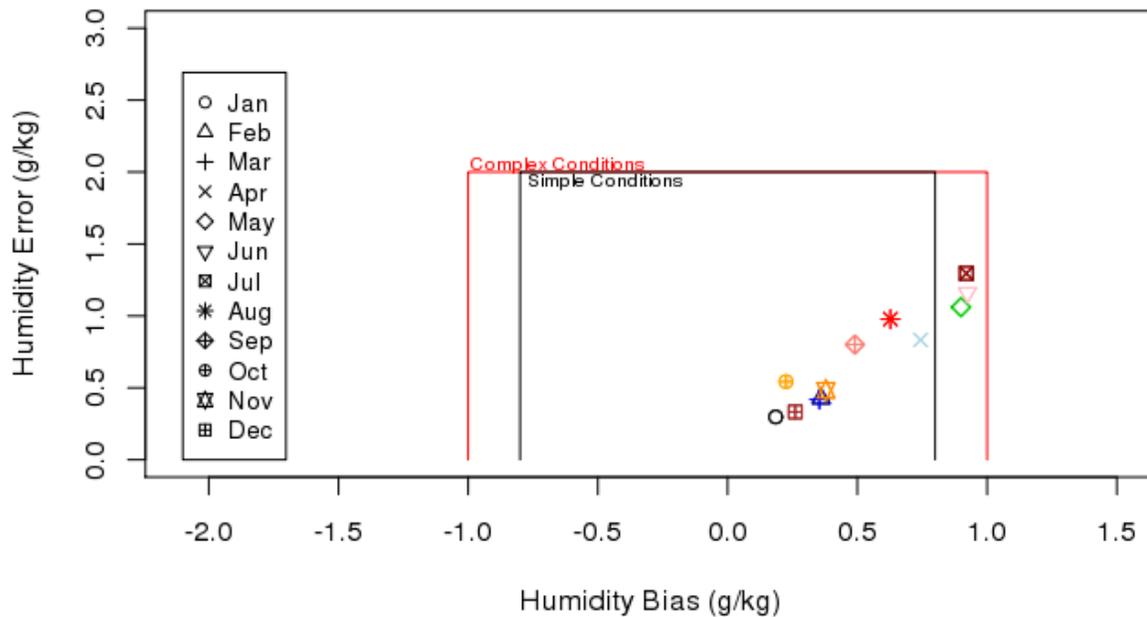


Figure 8. ACHD WRF METSTAT d04 Humidity Performance

3.2 Qualitative Evaluation Using Vertical Profile Analysis

Upper-air data from the Pittsburgh radiosonde (KPIT) dataset were used to evaluate WRF vertical profiles of predicted temperature, moisture and wind speeds above that location. The KPIT radiosonde dataset is collected and maintained by the National Weather Service (NWS). Radiosondes are launched from KPIT twice per day, at approximately 0 and 12 UTC (7 AM and 7 PM LST). Radiosondes provide high-resolution vertical profiles of temperature, humidity, wind speed and wind direction throughout the troposphere. The data are made publicly available by NOAA on the Earth System Research Laboratory (ESRL) Radiosonde Observation website (www.esrl.noaa.gov/raobs). Ramboll Environ downloaded and stored the radiosonde data from 2011 in FSL format for use in WRF model dataset comparisons.

The advantages of traditional radiosonde datasets are the length of the dataset and that they are direct measurements, as opposed to remote sensing. Disadvantages include the low frequency at only twice per day, and that the lowest reported level comes from a different sensor than the rest of the profile and can lead to un-physical profiles. The measurement from the stationary sensor may be warmer than the lowest layer reported by the radiosonde, which is an unstable situation that could not actually occur as the air would rapidly mix and the unstable temperature gradient would be erased.

3.2.1 Pittsburgh Temperature, Dew Point, and Wind Speed Soundings

Vertical profile plots showing WRF modeled data from the 1.333 km domain and observed upper-air soundings were created in order to evaluate the performance of the vertical atmospheric structure. The upper-air observation location from the Pittsburgh National Weather Service office (KPIT, WBAN 94823) was selected as the closest sounding site to the Allegheny County PM_{2.5} study area. At a distance of about 24 miles to the northwest (outside of the 0.444 km domain) the KPIT balloon-borne radiosonde launch site provides the nearest representative NWS upper-air data for the study and vicinity.

Figure 9 through Figure 14 display examples of vertical profile comparisons of WRF temperatures and dew points (blue lines) against actual upper-air temperatures and dew points (red lines) on the top panels and comparisons of WRF wind speeds (blue lines) against actual upper-air wind speeds (red lines) on the bottom panels. Each figure includes vertical profile comparisons for the two daily soundings: 0 UTC or 6:00 LST (left panels) and 12 UTC or 18:00 LST (right panels). As a small sampling, vertical profile plots and the corresponding analyses and corresponding analyses are provided below for the first days of January, March, May, July, September, and November of 2011.

On the left panel of Figure 9, WRF forecasts a small surface-based temperature inversion in the morning of January 1st which is not present in the observations. In the evening, WRF over-predicts surface dew point by roughly 2 degrees Celsius, which would correspond to the slight positive bias in surface-level moisture as shown in the METSTAT evaluation in Section 3.1.2. Other than these discrepancies, WRF replicates the vertical profiles of temperature and dew point relatively accurately. On the bottom panels of Figure 9, wind speeds are represented very well by WRF from the surface to 4000 feet.

On the morning of March 1st, shown in Figure 10, WRF represents the vertical profiles of temperature and dew point with considerable accuracy. WRF also represents the evening temperatures well. Evening dew point is represented very accurately up to an elevation of about 750 meters, with a maximum overestimation of about 2 degrees Celsius at around 750 meters. Above 1200 meters, WRF temperature begins to diverge from the observations to an under-estimation of about 5 degrees by 1800 meters. At around 500 meters, WRF over-predicts morning winds by 2-3 m/s and under-predicts afternoon winds by about the same margin. Near the surface, WRF wind speeds are more accurate at these times.

In Figure 11 on the morning of May 1st, WRF over-predicts temperature up to around 900 meters by about 2 degrees Celsius. WRF also over-predicts evening temperature near the surface by about 3-4 degrees. Dew point profiles are represented well by WRF at lower altitudes, but WRF significantly underestimates dew point at around 1500 meters. However, there are limited dew point observation data at other high altitudes that morning which makes it difficult to assess the extent of the dry layer that WRF failed to model. Meanwhile, WRF accurately forecasts morning and afternoon wind speeds near the surface and at most elevations. In the evening, however, WRF overestimates wind speed aloft, at around 500 meters, by about 5 m/s.

On the morning of July 1st, shown in the left panel of Figure 12, WRF again shows a surface-based inversion up to about 100 meters, which is not represented in the observed profile. Otherwise, morning temperatures and dew points are forecasted very accurately. In the evening of July 1st, WRF over-estimates surface temperature by about 4 degrees Celsius, but converges to match observed temperatures as elevation increases to about 900 meters. WRF represents the morning wind speed profile very well, but over-estimates the calm evening surface wind by 2-3 m/s.

In Figure 13, WRF again generates a very small surface-based temperature inversion up to about 50 meters in the morning of September 1st, which does not exist in the observations. Besides this, WRF performs considerably well at modeling both morning and evening temperatures through the atmosphere. The modeled dew point is accurate up to about 700 meters at both times of day. In the morning, WRF underestimates dew point in the upper atmosphere, while overestimating upper-atmosphere dew point in the evening. In both the morning and evening, WRF overestimates the observed, calm wind speeds by about 5 m/s. The remainder of the morning wind profile is modeled more accurately, but WRF significantly underestimates evening wind speeds above 50 meters by 5-10 m/s.

On November 1st, Figure 14 indicates that WRF overestimates morning temperatures while underestimating dew point each by a couple degrees Celsius in the lower atmosphere. This creates drier conditions in the model than what were observed. In the evening of November 1st, WRF again simulates a surface-based temperature inversion, up to about 200 meters, that was not observed at KPIT. While the evening surface temperatures match, the inversion modeled by WRF creates an over-prediction of temperature up to about 5 degrees Celsius at around 300 meters altitude. WRF simulates morning winds speeds well, but overestimates evening wind speeds near the surface and underestimates evening wind speeds aloft, around 300 meters, by about 5 m/s each.

Overall, WRF performs well in forecasting vertical profiles of temperature, dew point, and wind speeds. Note that WRF vertical profiles tend to appear more smoothed than the observed profiles due to missing KPIT observation data at various heights. Wind speed forecasting accuracy varies but shows no obvious hourly or seasonal bias above the surface. WRF often overestimates wind speed at the surface compared to observations, which could be a reflection of higher surface roughness near the specific monitoring site, compared to the roughness modeled through the domain. The most notable discrepancy between modeled and observed vertical profiles is the tendency of WRF to generate shallow, surface-based temperature inversions in the mornings, when they are not present in the observations. In such cases, WRF usually underestimates surface temperature but matches observed temperatures at higher altitudes. Based on METSTAT analysis, this trend appears to be a result of a cold bias in the surface energy budget near the KPIT sounding site. Figure 28, in Appendix A, shows METSTAT monthly temperature in the 4 km domain, d03. In this larger domain, considering data from many more meteorological stations, the negative temperature bias no longer exists during most months. Thus, it is apparent that WRF's negative temperature bias is mostly specific to the KPIT meteorological station in Allegheny County, and that the erroneously modeled surface-based temperature inversions are likely a result of this surface energy deficit.

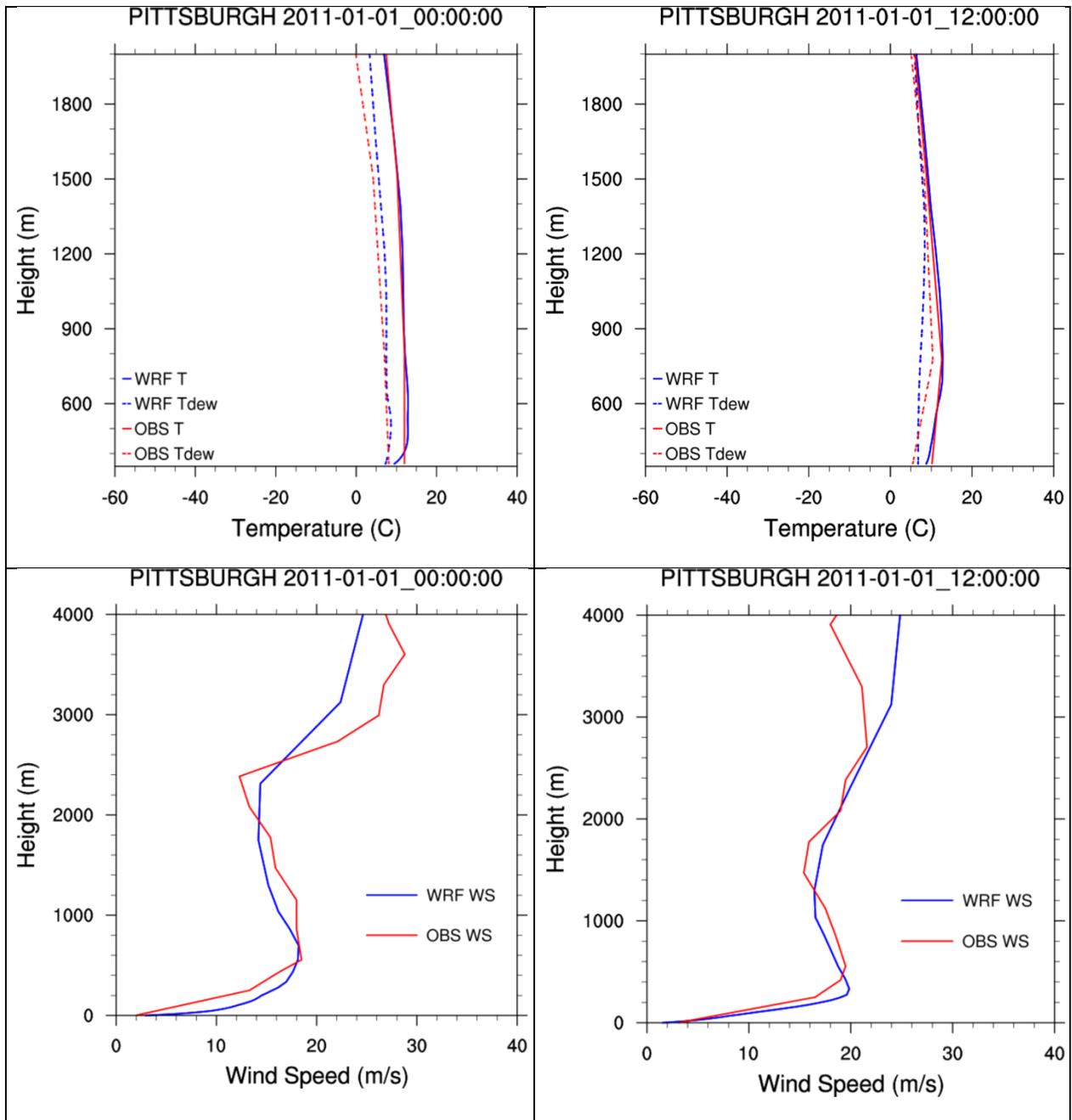


Figure 9. Vertical profile soundings comparing WRF to upper-air data at KPIT on January 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

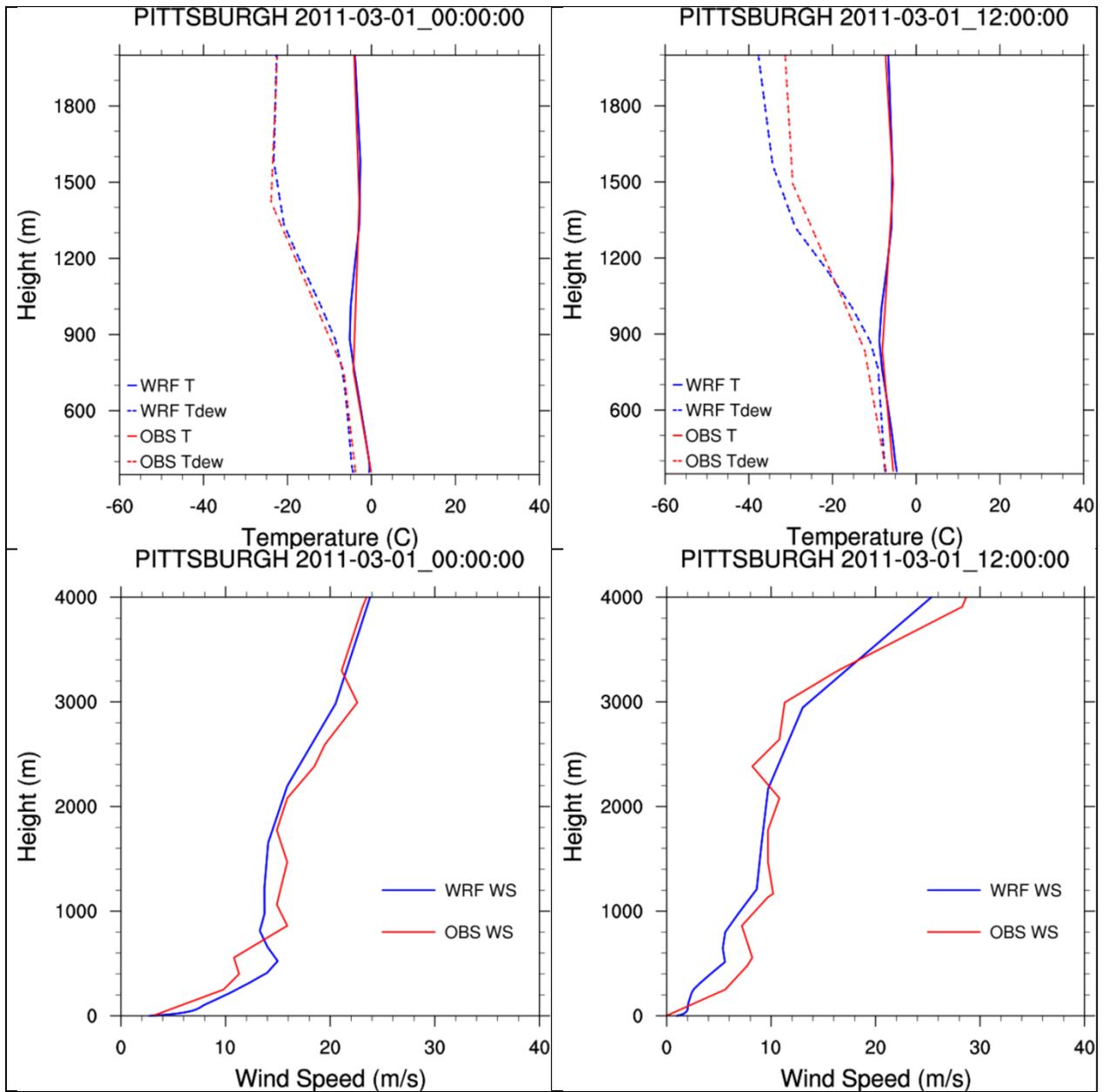


Figure 10. Vertical profile soundings comparing WRF to upper-air data at KPIT on March 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

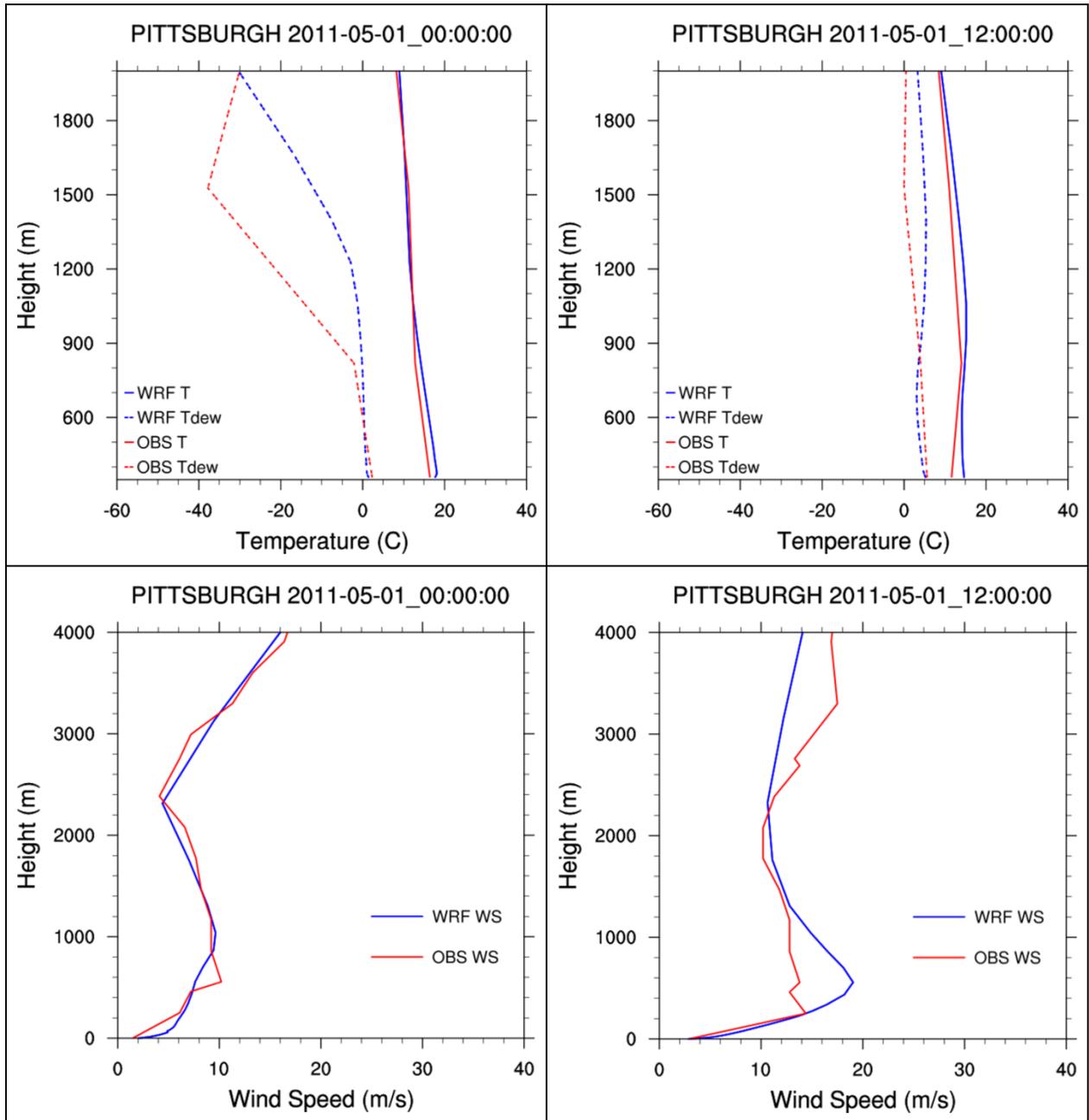


Figure 11. Vertical profile soundings comparing WRF to upper-air data at KPIT on May 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

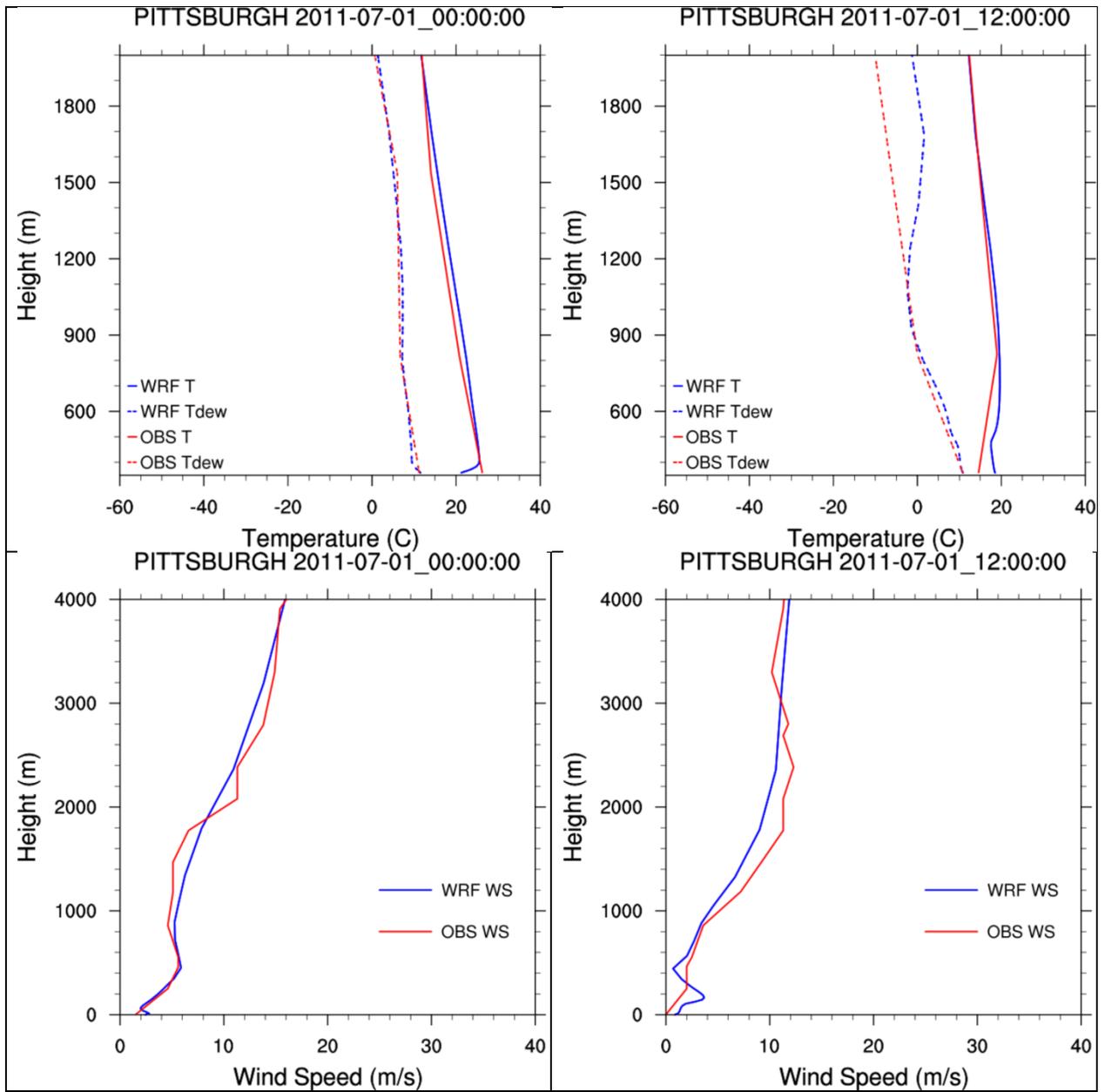


Figure 12. Vertical profile soundings comparing WRF to upper-air data at KPIT on July 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

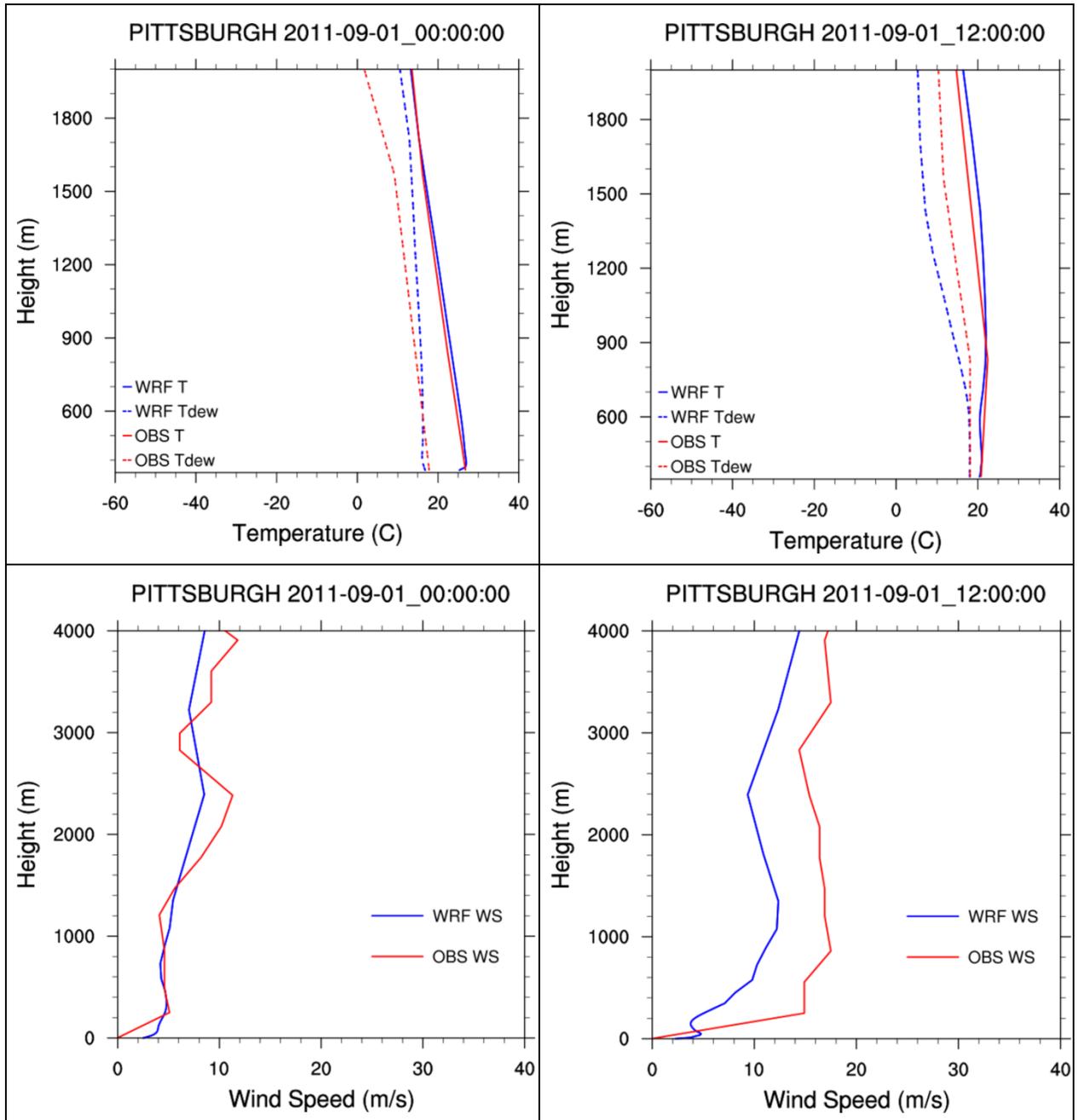


Figure 13. Vertical profile soundings comparing WRF to upper-air data at KPIT on September 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

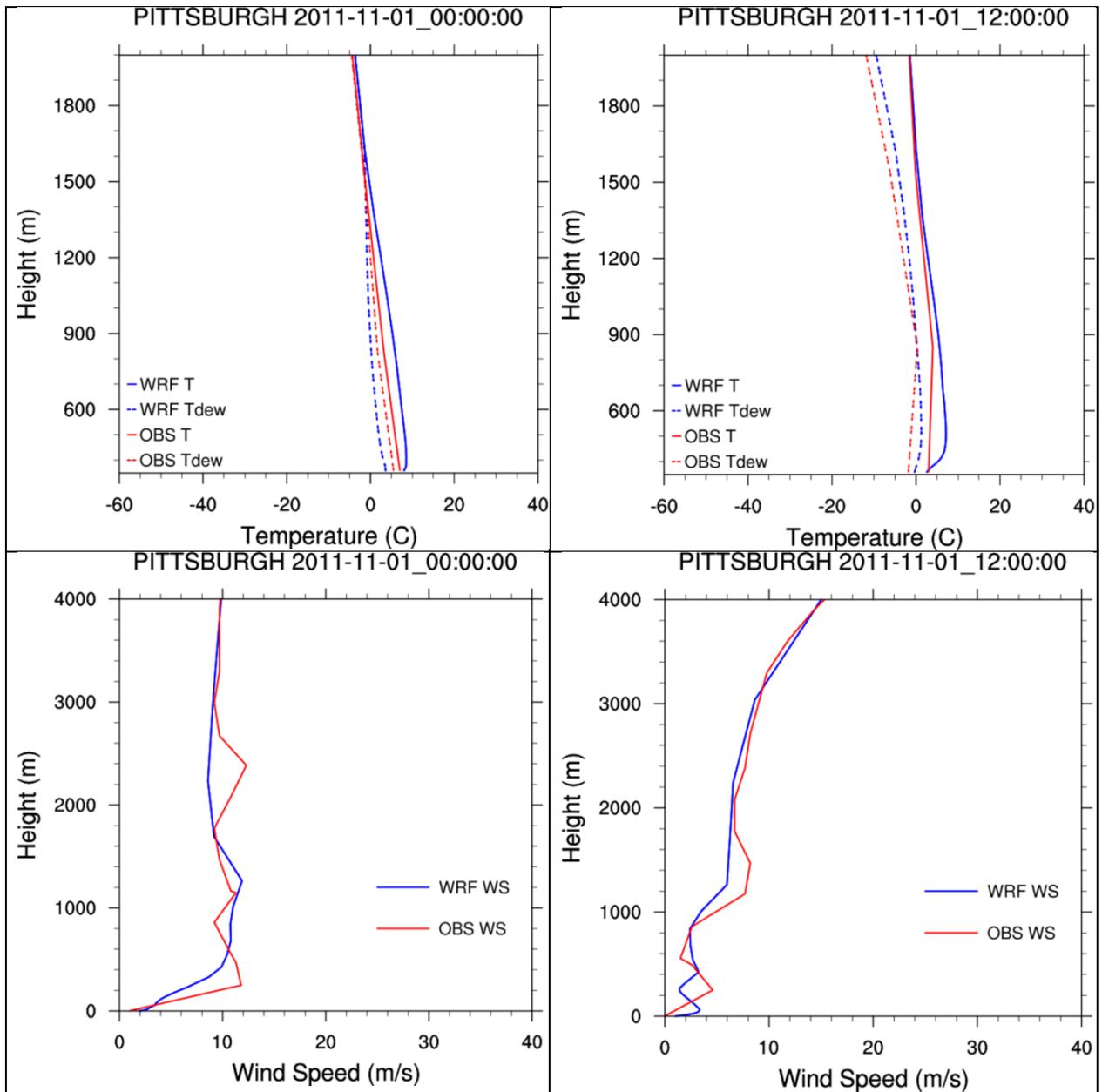


Figure 14. Vertical profile soundings comparing WRF to upper-air data at KPIT on November 1st, 2011 at 0 and 12 UTC (7 AM and 7 PM LST) for the 1.33 km domain.

3.3 Qualitative Evaluation Using PRISM Precipitation

This section presents a qualitative comparison of WRF-simulated monthly precipitation with monthly PRISM (Daly, 2008) analysis fields based on observations. Monthly precipitation plots for 2011 were constructed from ACHD WRF output and compared to PRISM monthly plots for the months of January through December in the 4 km domain (d03) and 1.33 km domain (d04).

The Parameter-elevation Regressions on Independent Slopes Model (PRISM) datasets are spatial maps of climate elements across the United States built by the Spatial Climate Analysis Service at Oregon State University (SCAS-OSU). Gridded maps of mean monthly and annual precipitation and temperature were built using meteorological station measurements and a set of statistical weighting procedures and corrections based on distance, elevation, topographic orientations and influences, and land-surface type (SCAS-OSU, 2001). The process and results have been extensively peer-reviewed and generally accepted by the climatological community as state-of-the-art representations.

When comparing the WRF and PRISM precipitation data, note that the PRISM analysis fields only cover the continental United States and do not extend into Canada, over the Great Lakes or any oceanic portions of the domain. The WRF fields, on the other hand, cover the entire domain. To address this, the WRF precipitation plots are masked to emphasize rainfall over the CONUS and thus better facilitate visual comparison of the plots. Additionally, as the precipitation monitoring sites tend to be located at lower elevations (e.g., airports), the PRISM observation fields may not fully capture the enhanced precipitation at high elevations due to orographic effects that could be present in the WRF simulations. However, unlike early analysis fields based on observations, PRISM does include an elevation effect as one of the parameters in its regression model.

Figure 15 displays a 12 km domain example comparison of PRISM precipitation data on the top panel, to masked WRF precipitation output on the bottom panel for April 2011. On a larger scale, WRF accurately simulates the large band of high precipitation extending northeast from Arkansas and Missouri. However, there are some smaller-scale discrepancies between the two plots, including different distributions of the highest precipitation levels within Pennsylvania.

Figure 16 displays an example comparison in the 4km domain of PRISM precipitation data to masked WRF precipitation for January 2011. WRF slightly misplaces some of the precipitation in western Pennsylvania, but predicts the general placement and volume of the precipitation bands west and east of the Appalachian Mountains with relative accuracy. Overall, WRF under-predicts the total January precipitation in this domain, but the average and median values are close to those of PRISM.

Figure 17 displays an example comparison in the 1.33 km domain of PRISM precipitation data to masked WRF precipitation for July. In this case WRF under-predicts the average and median rainfall in Southwestern Pennsylvania, mostly due to under-prediction just to the north of Allegheny County. WRF does simulate the higher precipitation levels on the eastern edge of Allegheny County, though it is more concentrated in the middle latitudes of the domain than the PRISM data indicates. As these plots demonstrate, modeling precipitation patterns with high precision becomes increasingly unlikely on smaller and finer domains. Precipitation plots for the smallest domain of this study, 0.444 km, are less relevant than the others and will not be included in the remainder of this evaluation.

Figure 29 through Figure 32 in Appendix B display the WRF precipitation datasets and the PRISM data for January through April in the 4 km domain. In terms of monthly precipitation totals, WRF varies between over-prediction and under-prediction in this domain. However, WRF represents the spatial extent of the precipitation well in each of the months, reflecting the enhanced rainfall rates along the Appalachian Mountains from mechanical uplift and along the eastern shores of Lake Erie from convective uplift conditions.

Figure 33 through Figure 37 in Appendix B show WRF and PRISM average precipitation comparisons for May through September in the 4 km domain. WRF accurately represents the spatial extent of rainfall maximums through portions of West Virginia and western Pennsylvania in the months of spring through summer. However, WRF consistently overestimates the amount of total precipitation across the domain. This is likely due to the excess moisture in the spring and summer WRF simulations as reflected by the positive humidity bias in the METSTAT performance plots.

WRF and PRISM precipitation comparisons for October through December in the 4 km domain are shown in Figure 38 through Figure 40 in Appendix B. In October, WRF represents the large-scale

spatial distribution well throughout the domain. However, WRF mostly under-predicts total rainfall across Pennsylvania, particularly in the northwest corner of the state. This under-prediction is not severe. WRF generates an accurate spatial distribution of precipitation for November, but again overestimates precipitation through the Appalachian Mountains. In December, WRF under-predicts precipitation on the eastern side of the mountain range, but simulates an accurate spatial distribution on the western side.

Total PRISM Precipitation for 2011-04
 Contiguous U.S. Statistics: 10th=2.59 Median=5.75 Average=6.27 90th=11.21

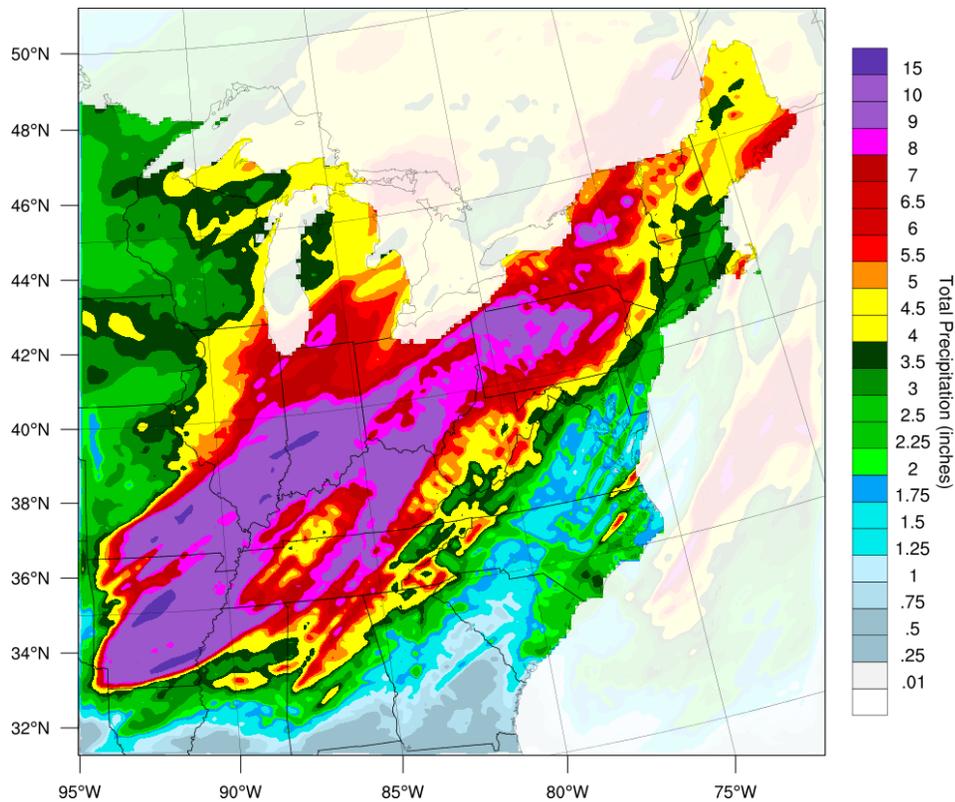
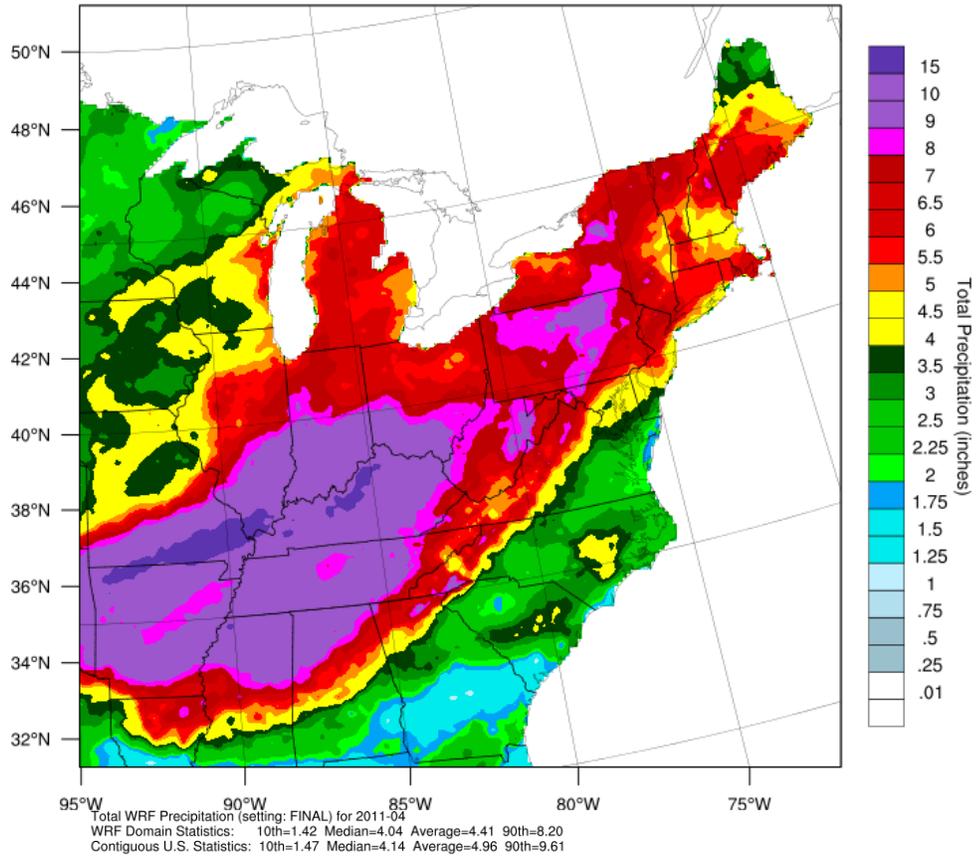


Figure 15. PRISM precipitation (top) and WRF precipitation (bottom) for January 2011 in the 12 km domain

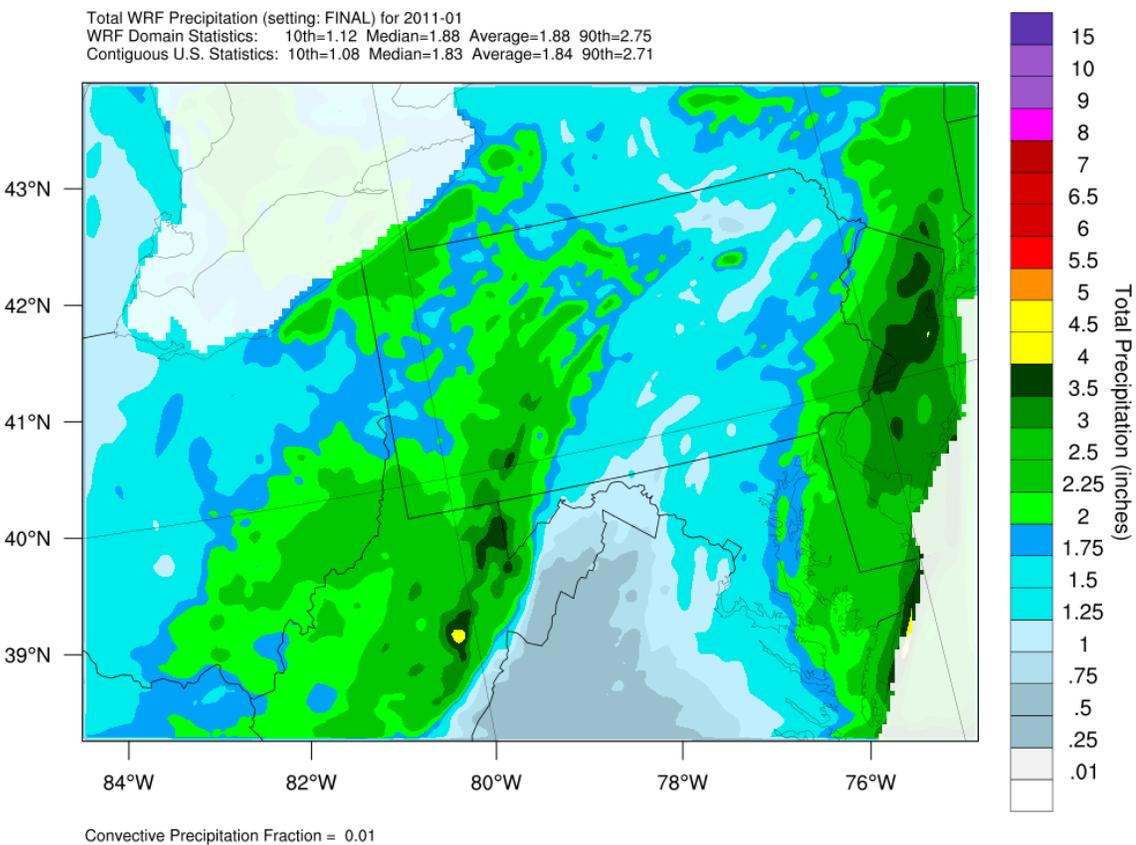
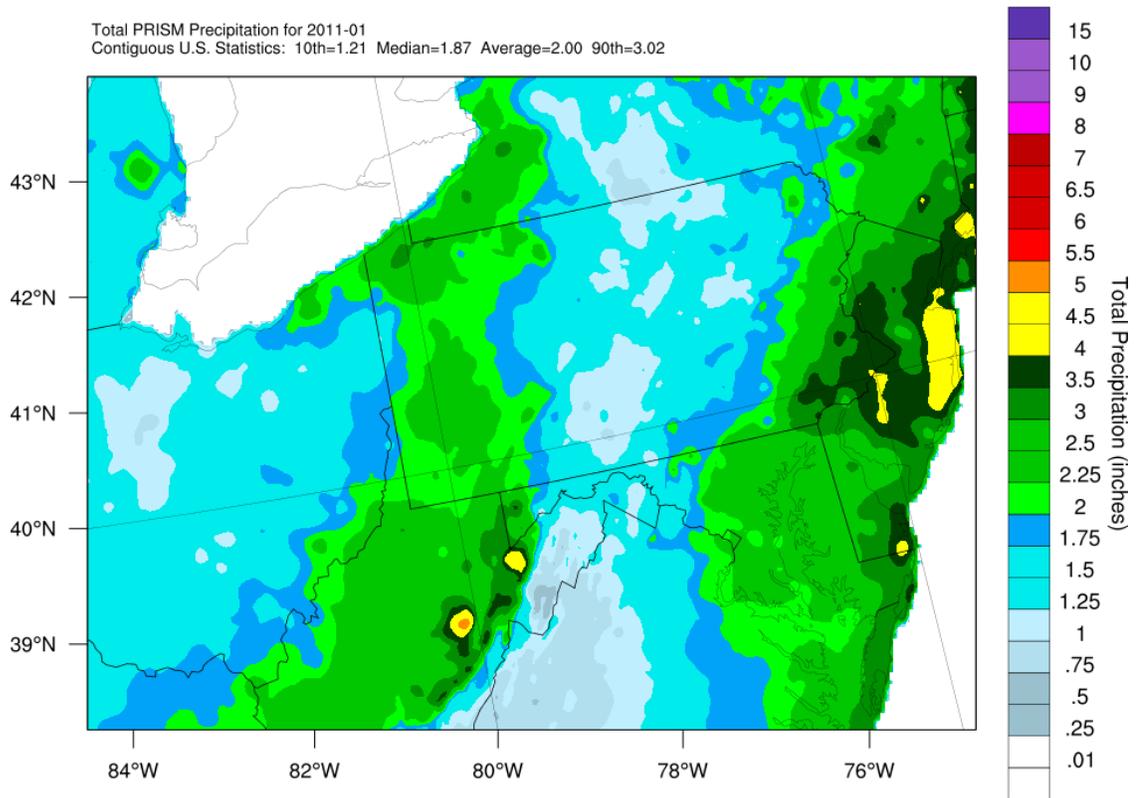
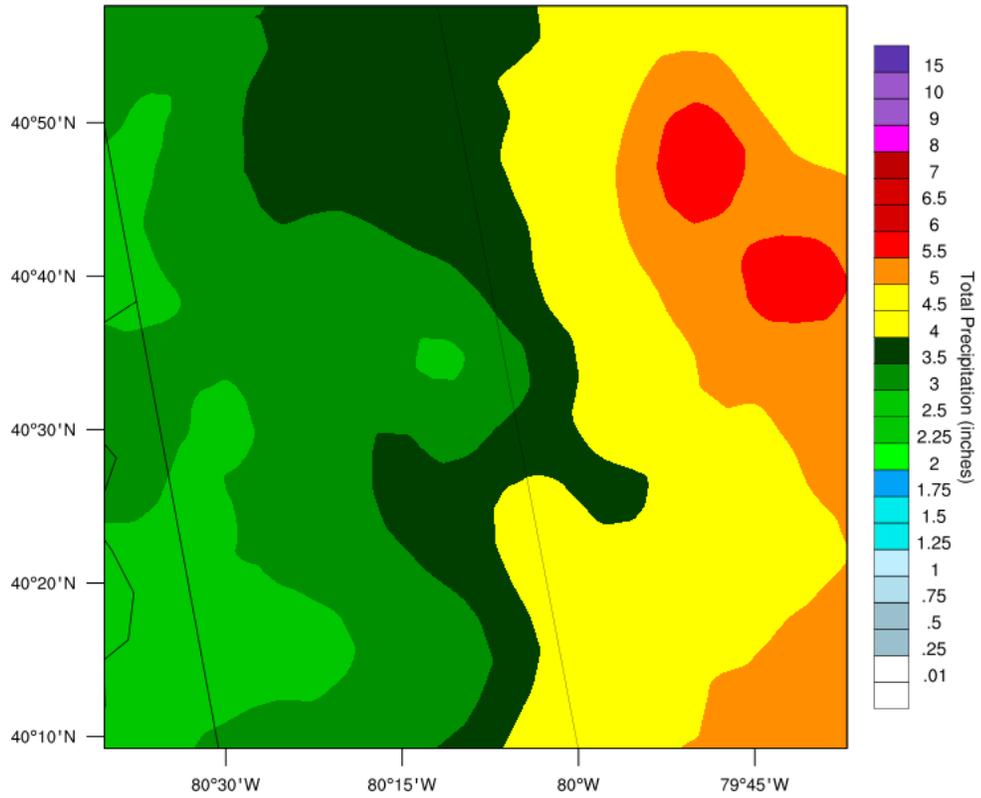


Figure 16. PRISM precipitation (top) and WRF precipitation (bottom) for January 2011 in the 4 km domain

Total PRISM Precipitation for 2011-07
 Contiguous U.S. Statistics: 10th=3.00 Median=3.79 Average=3.94 90th=5.14



Total WRF Precipitation (setting: FINAL) for 2011-07
 WRF Domain Statistics: 10th=1.94 Median=3.39 Average=3.52 90th=5.17

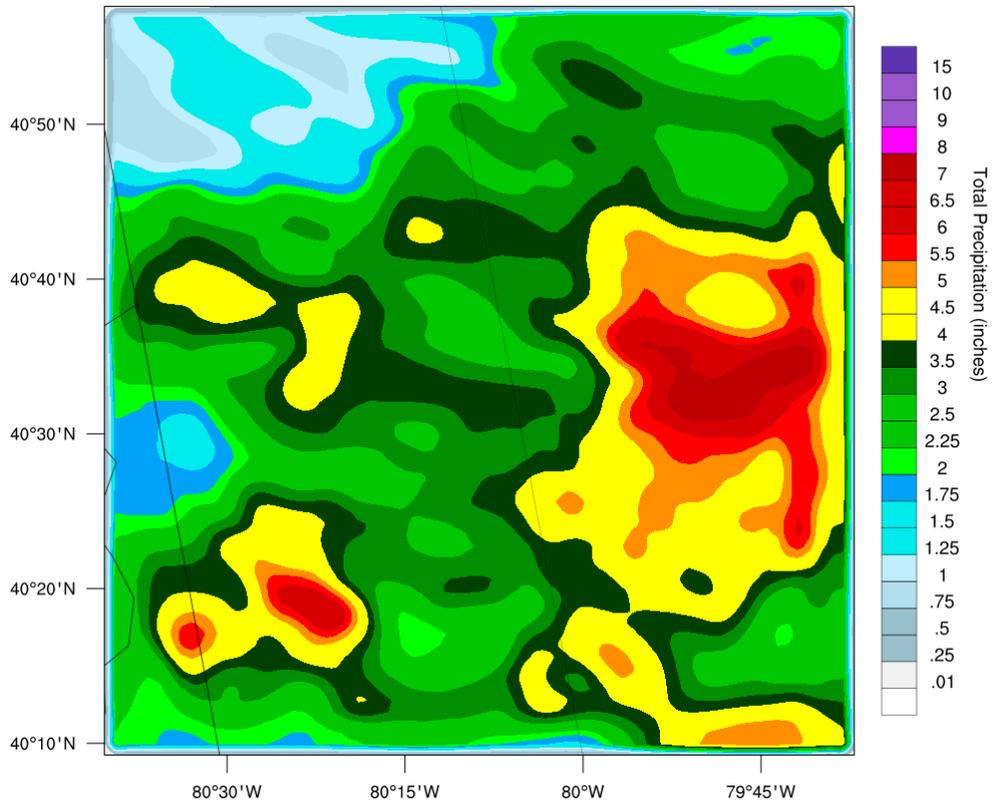


Figure 17. PRISM precipitation (top) and WRF precipitation (bottom) for February 2011 in the 1.33 km domain

4. SUMMARY AND CONCLUSIONS

The ACHD WRF meteorological model simulation for January through December of 2011 reproduces the observed surface and upper-air meteorological variables adequately. The 1.333 km (d04) METSTAT performance for wind direction, wind speed, and temperature in the region shows a strong agreement between the model and surface observations. The METSTAT performance for humidity does reflect a slight positive bias in the warmer months, leading to additional moisture in the model.

Upper air performance in the 1.333 km (d04) domain shows acceptable agreement between WRF and radiosonde data at the KPIT station. WRF vertical profiles of temperature, dew point, and wind speeds were especially accurate for the evening soundings. The most notable error in vertical profiles is the tendency of WRF to simulate shallow, morning, surface-based temperature inversions when they are not present in the observations. This discrepancy is likely due to a negative bias in the surface energy budget in the 1.33km and 0.444km domains, as evidenced by the METSTAT plots.

The precipitation analysis for the 4 km (d03) domain indicates there is a slight understatement of precipitation in the fall and winter months, while higher humidity biases create excess moisture and results in an overstatement of precipitation in the summer months. Modeled precipitation accumulations in southwest Pennsylvania are more accurate compared to other regions in the northeast.

Based on our experience, the ACHD PM_{2.5} WRF modeling's performance provides a sound basis for developing meteorological inputs for air quality dispersion modeling.

5. REFERENCES

- Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, P.P. Pasteris
2008. Physiographically sensitive mapping of climatological temperature and precipitation across the
conterminous United States. *Int. J. Climatol.*,
(http://prism.nacse.org/documents/Daly2008_PhysiographicMapping_IntJnlClim.pdf)
- Emery, C.A., E. Tai, and G. Yarwood. 2001. Enhanced Meteorological Modeling and Performance
Evaluation for Two Texas Ozone Episodes. Prepared for the Texas Natural Resource Conservation
Commission (now TCEQ), by ENVIRON International Corp, Novato, CA. Available at:
[http://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/EnhancedMe
tModelingAndPerformanceEvaluation.pdf](http://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/mm/EnhancedMe
tModelingAndPerformanceEvaluation.pdf)
- Kemball-Cook, S., Y. Jia, C. Emery, and R. Morris, 2005. *Alaska MM5 Modeling for the 2002 Annual
Period to Support Visibility Modeling*. Prepared for the Western Regional Air Partnership, by ENVIRON
International Corp., Novato, CA. Available at:
http://pah.cert.ucr.edu/aqm/308/docs/alaska/Alaska_MM5_DraftReport_Sept05.pdf
- MADIS. 2015. Meteorological Assimilation Data Ingest System. Retrieved from
<http://madis.noaa.gov>
- McNally, D. E., 2009. *12 km MM5 Performance Goals*. Presentation to the Ad-Hoc Meteorology Group.
June 25, 2009. Available at: <http://www.epa.gov/scram001/adhoc/mcnally2009.pdf>
- NCAR. 2015. *Weather Research & Forecasting (WRF) ARW Version 3 Modeling System User's Guide*.
Nat. Center for Atmos. Research Mesoscale & Meteorology Division.
- NOHRSC, 2004. National Operational Hydrologic Remote Sensing Center. *Snow Data Assimilation
System (SNODAS) Data Products at NSIDC, Version 1*. Boulder, Colorado USA. NSIDC: National
Snow and Ice Data Center.
- NOAA. 2015. National Weather Service. National Centers for Environmental Prediction.
<http://www.ncep.noaa.gov>
- NOAA-NCDC. 2015. NOAA NCDC Integrated Surface Database. Retrieved from
www.ncdc.noaa.gov/isd
- Ramboll Environ US Corp. 2015. *METSTAT*. Retrieved from [http://www.camx.com/download/support-
software.aspx](http://www.camx.com/download/support-
software.aspx)
- SCAS-OSU, 2001. Climate Mapping with PRISM. Oregon State University (OSU)

Appendix A

ADDITIONAL METSTAT PLOTS

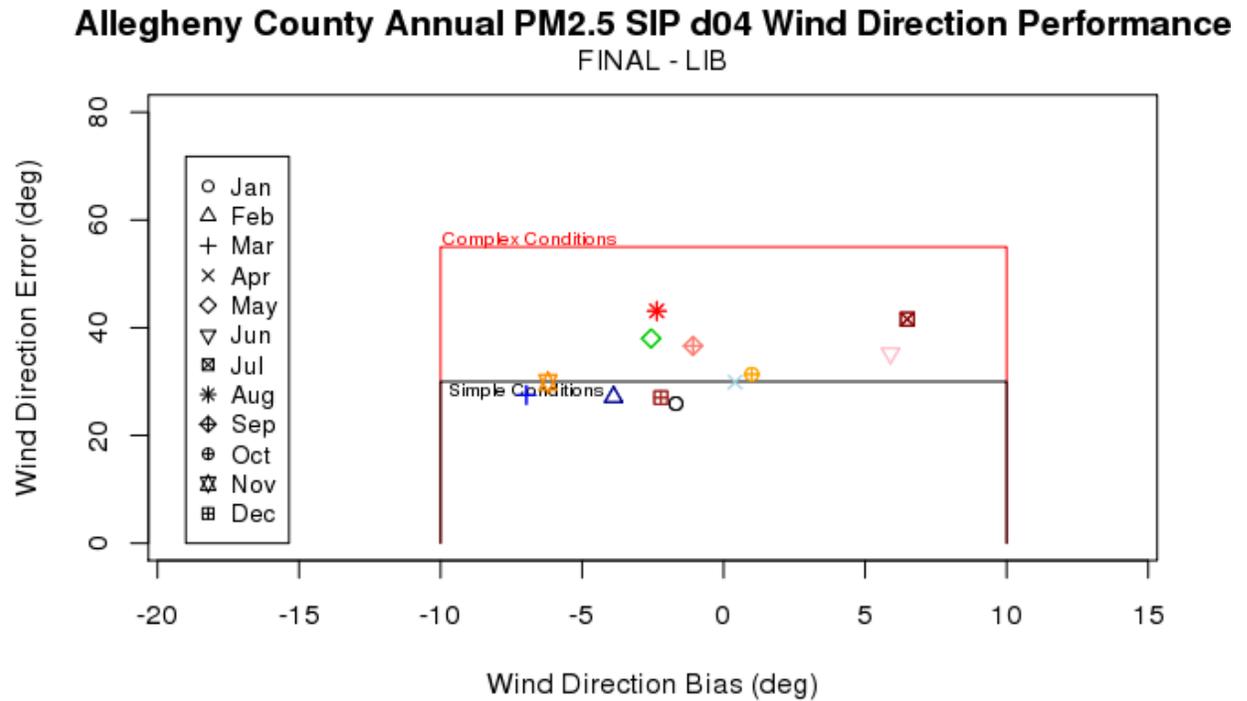


Figure 18. ACHD WRF Liberty Station METSTAT d04 Wind Direction Performance

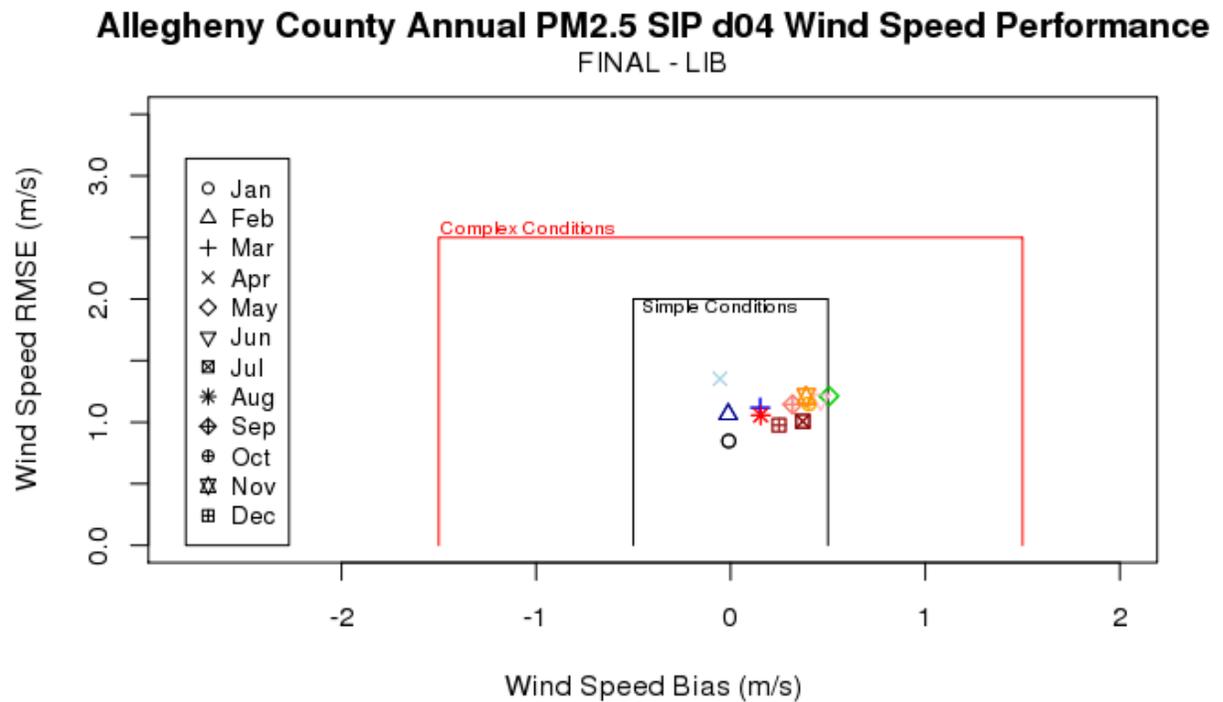


Figure 19. ACHD WRF Liberty Station METSTAT d04 Wind Speed Performance

Allegheny County Annual PM2.5 SIP d04 Temperature Performance
FINAL - LIB

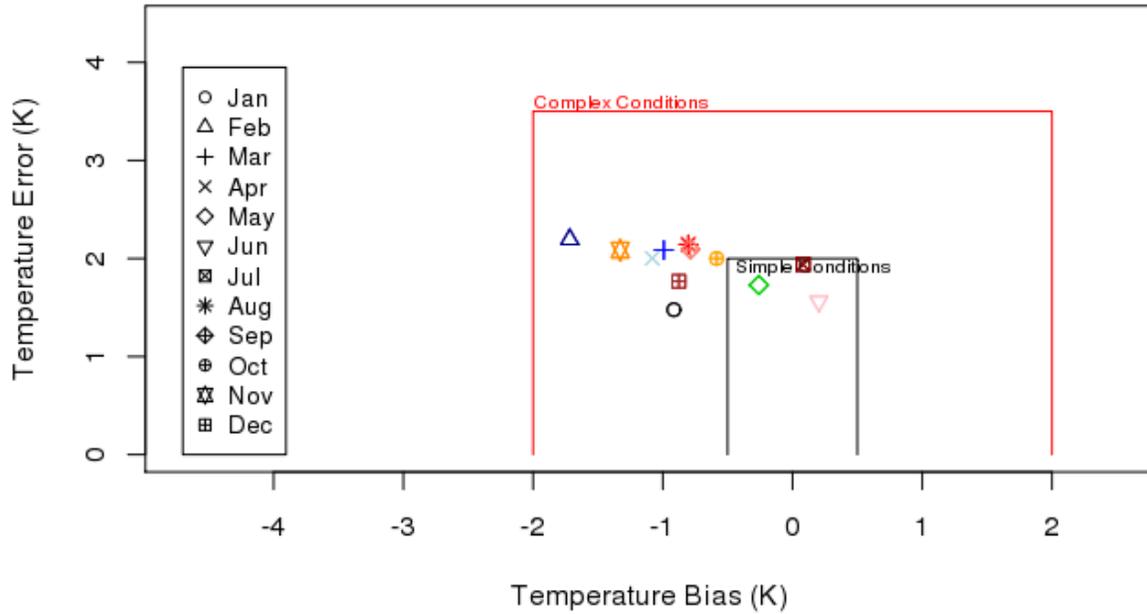


Figure 20. ACHD WRF Liberty Station METSTAT d04 Temperature Performance

Allegheny County Annual PM2.5 SIP d05 Wind Direction Performance
FINAL - LIB

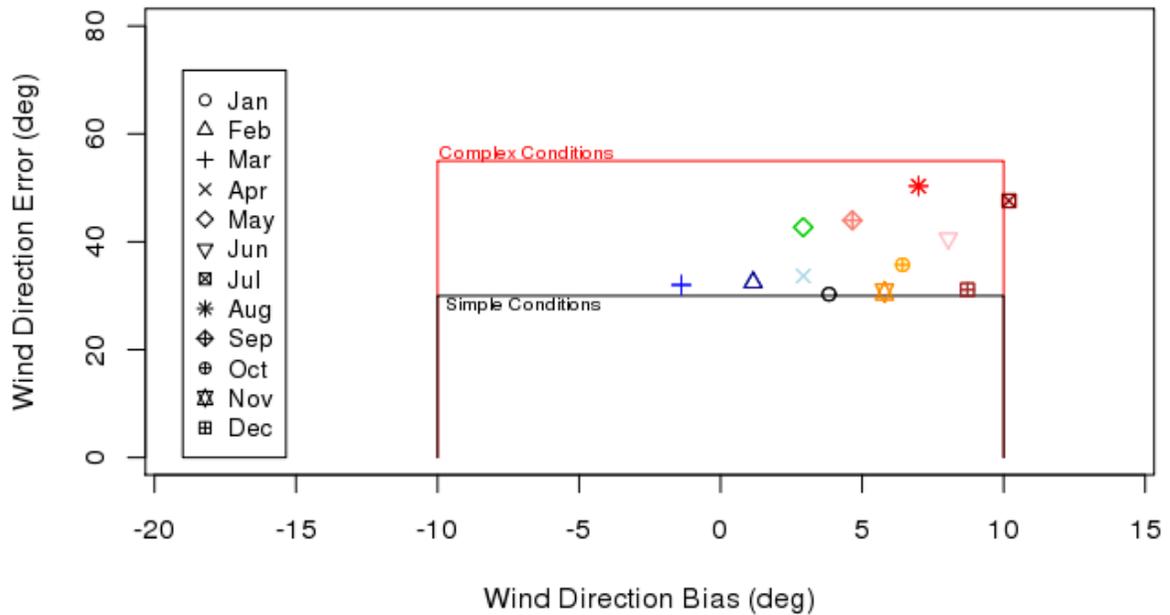


Figure 21. ACHD WRF Liberty Station METSTAT d05 Wind Direction Performance

Allegheny County Annual PM2.5 SIP d05 Wind Direction Performance
FINAL - all

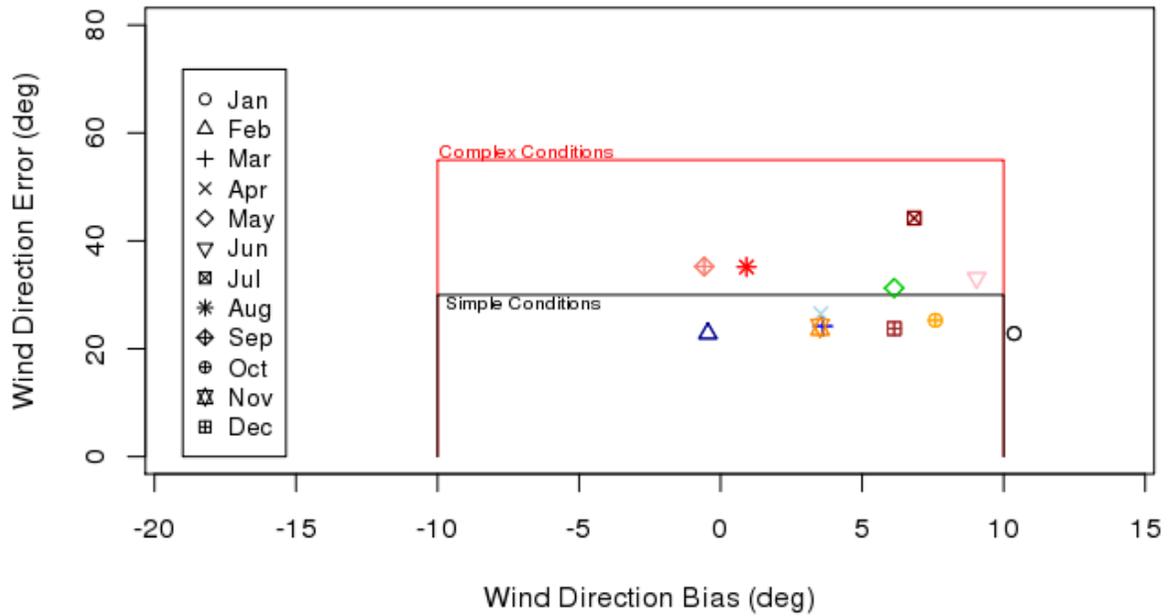


Figure 22. ACHD WRF Allegheny Airport METSTAT d05 Wind Direction Performance

Allegheny County Annual PM2.5 SIP d05 Wind Speed Performance
FINAL - LIB

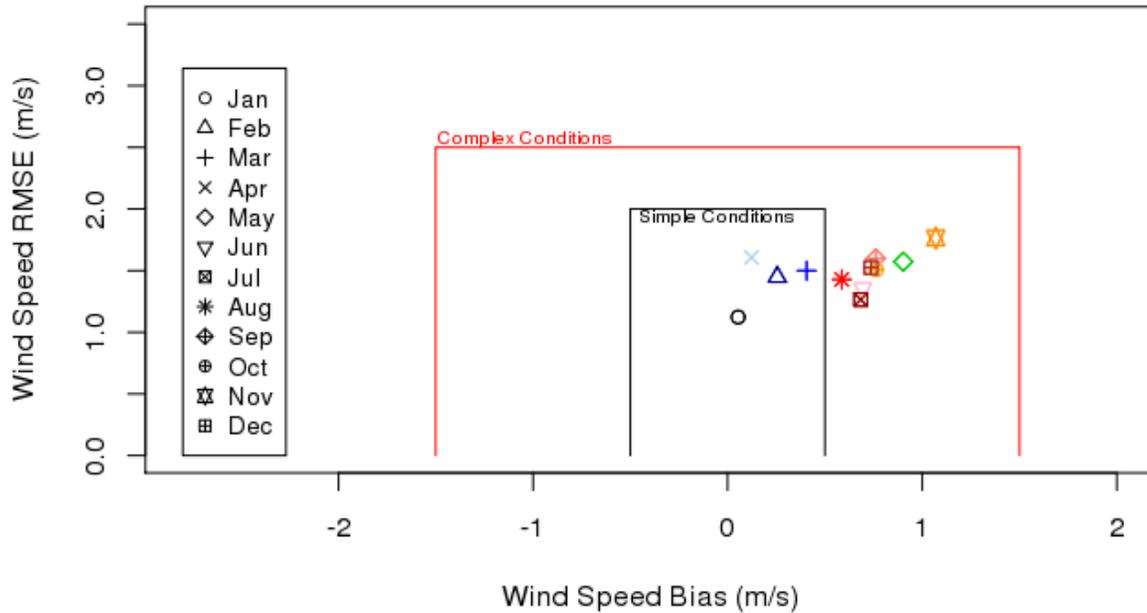


Figure 23. ACHD WRF Liberty Station METSTAT d05 Wind Speed Performance

Allegheny County Annual PM2.5 SIP d05 Wind Speed Performance
FINAL - all

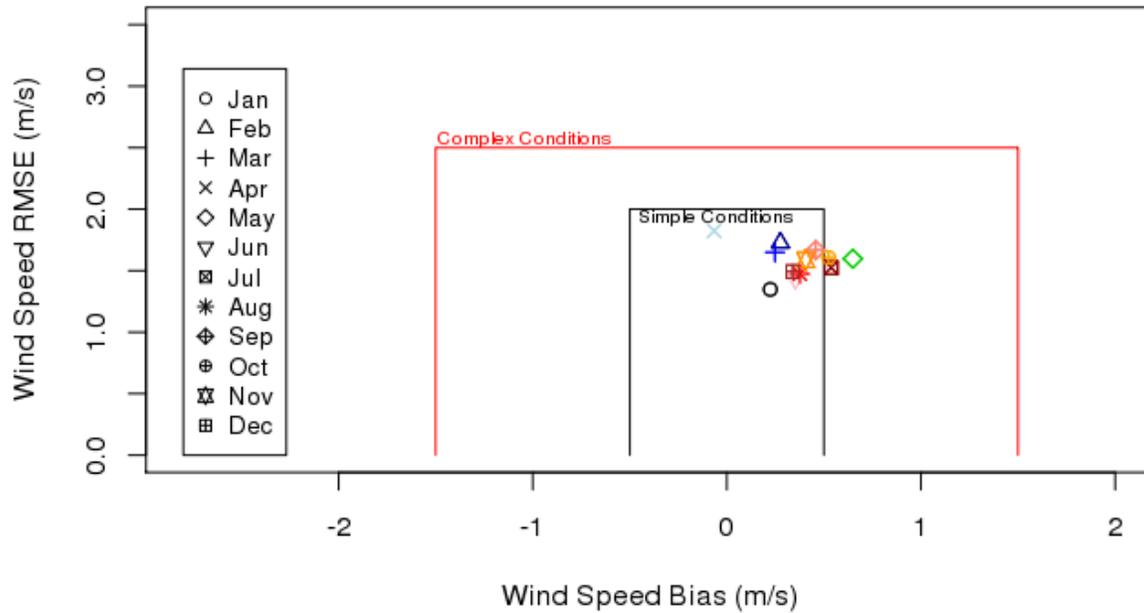


Figure 24. ACHD WRF Allegheny Airport METSTAT d05 Wind Speed Performance

Allegheny County Annual PM2.5 SIP d05 Temperature Performance
FINAL - LIB

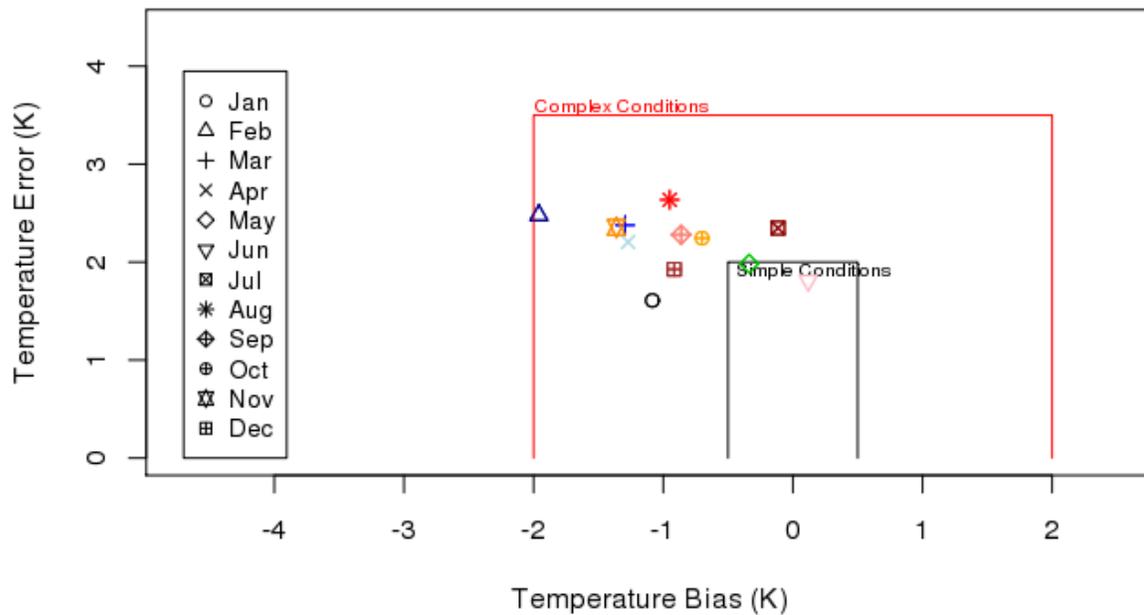


Figure 25. ACHD WRF Liberty Station METSTAT d05 Temperature Performance

Allegheny County Annual PM2.5 SIP d05 Temperature Performance
FINAL - all

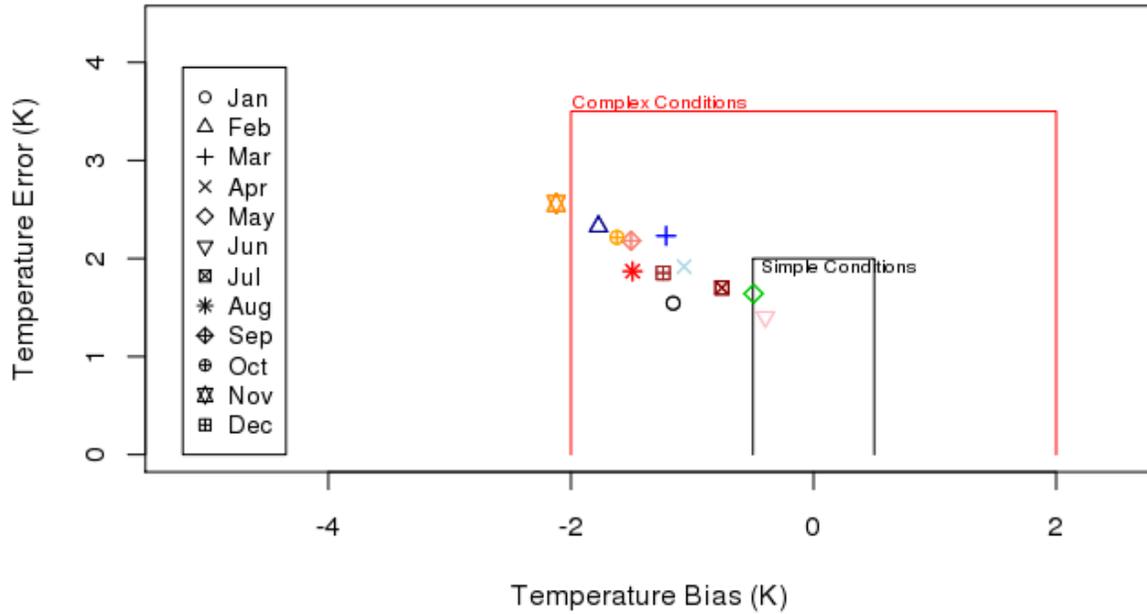


Figure 26. ACHD WRF Allegheny Airport METSTAT d05 Temperature Performance

Allegheny County Annual PM2.5 SIP d05 Humidity Performance
FINAL - all

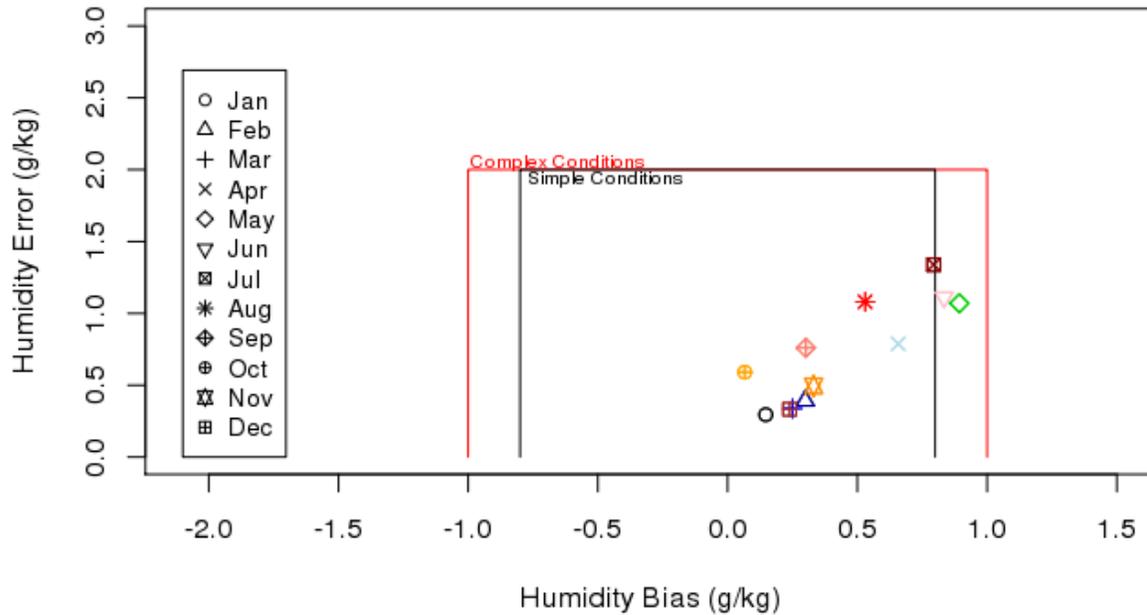


Figure 27. ACHD WRF Allegheny Airport METSTAT d05 Humidity Performance

Allegheny County Annual PM2.5 SIP d03 Temperature Performance

FINAL - all

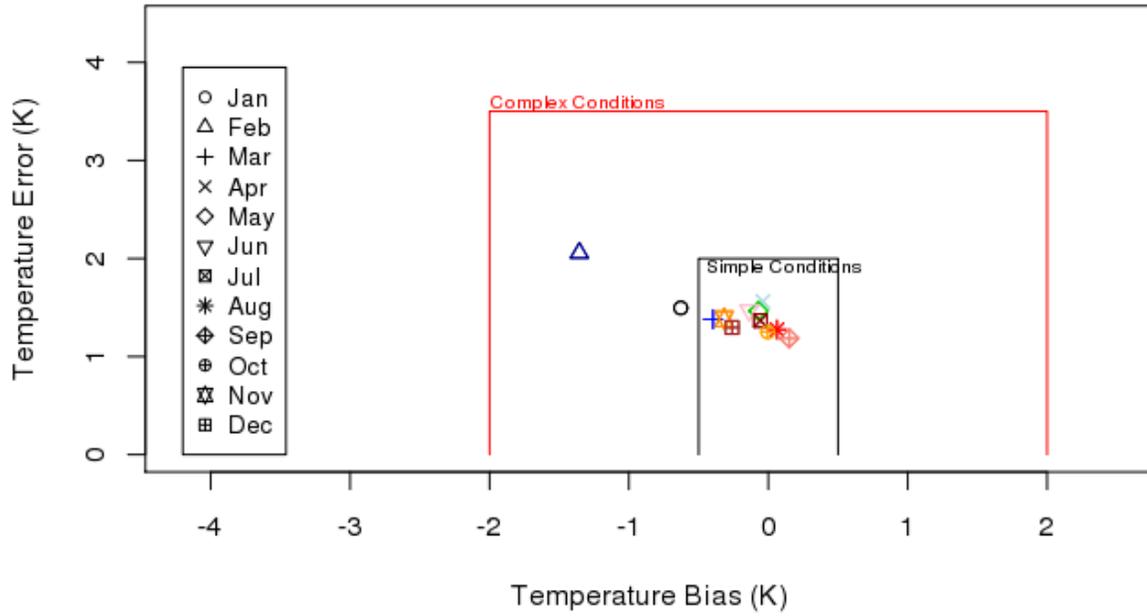


Figure 28. ACHD WRF Allegheny Airport METSTAT d03 Temperature Performance

Appendix B

PRECIPITATION PLOTS

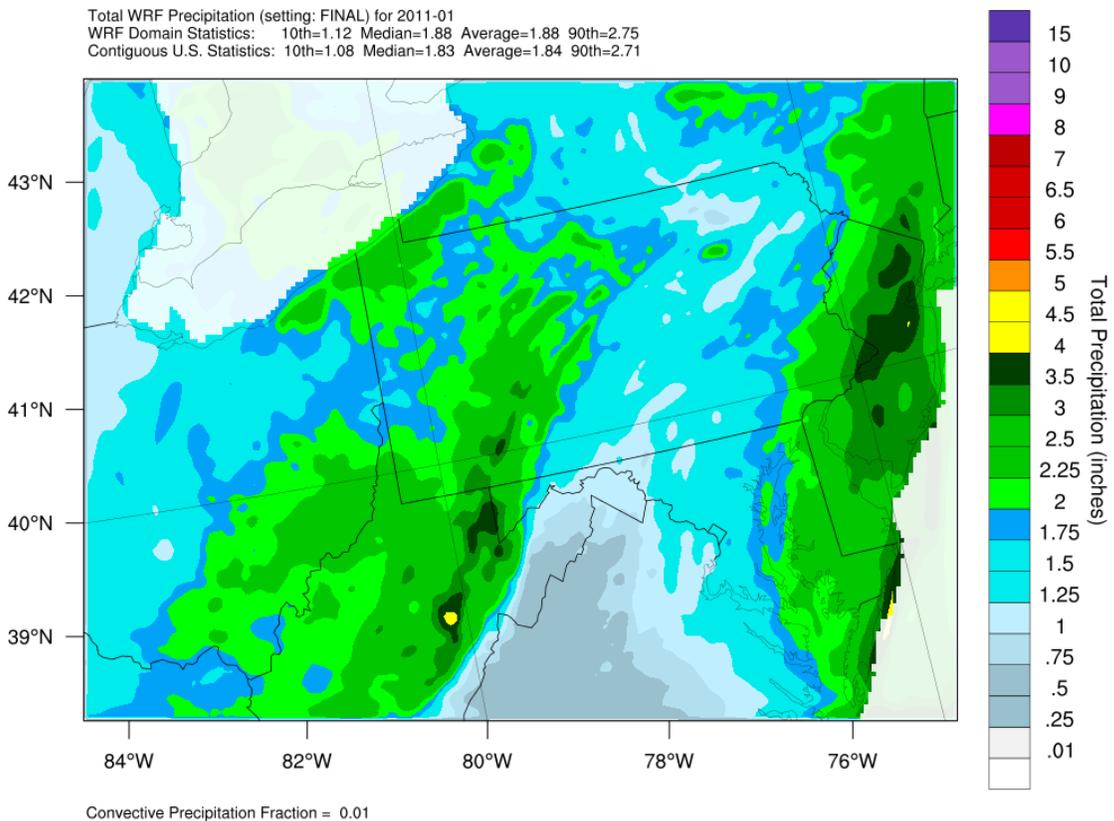
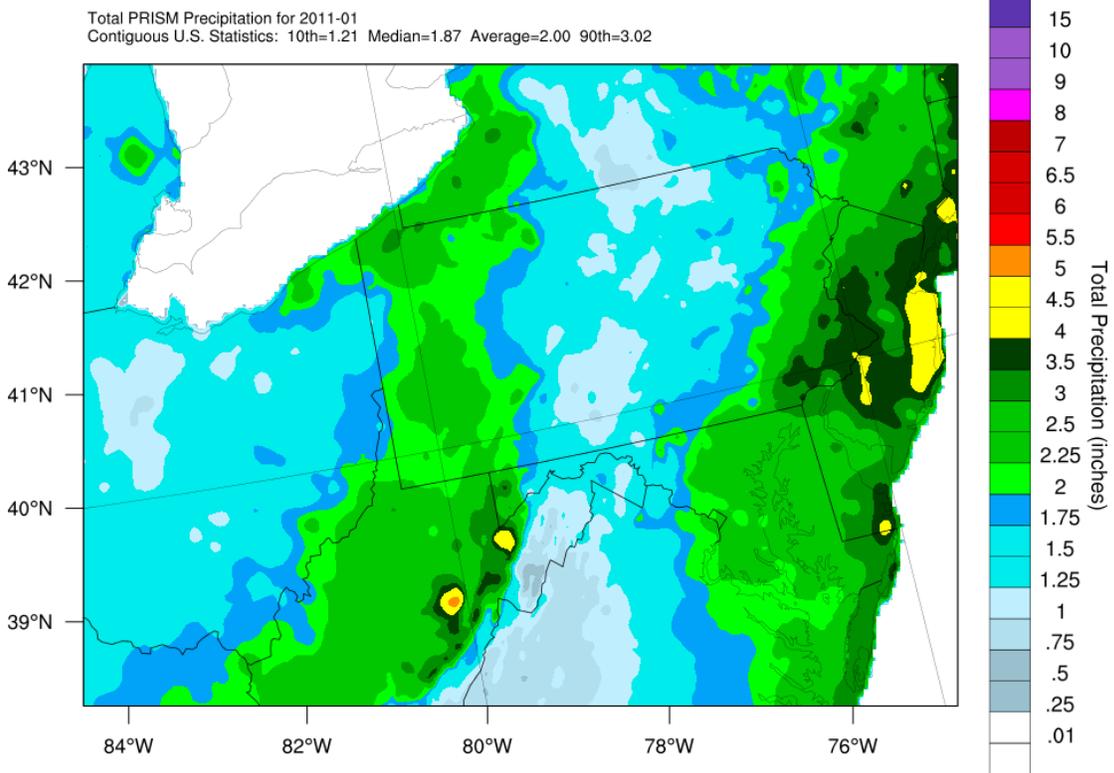


Figure 29. PRISM precipitation total (top) and WRF precipitation (bottom) for January 2011 in the 4 km domain

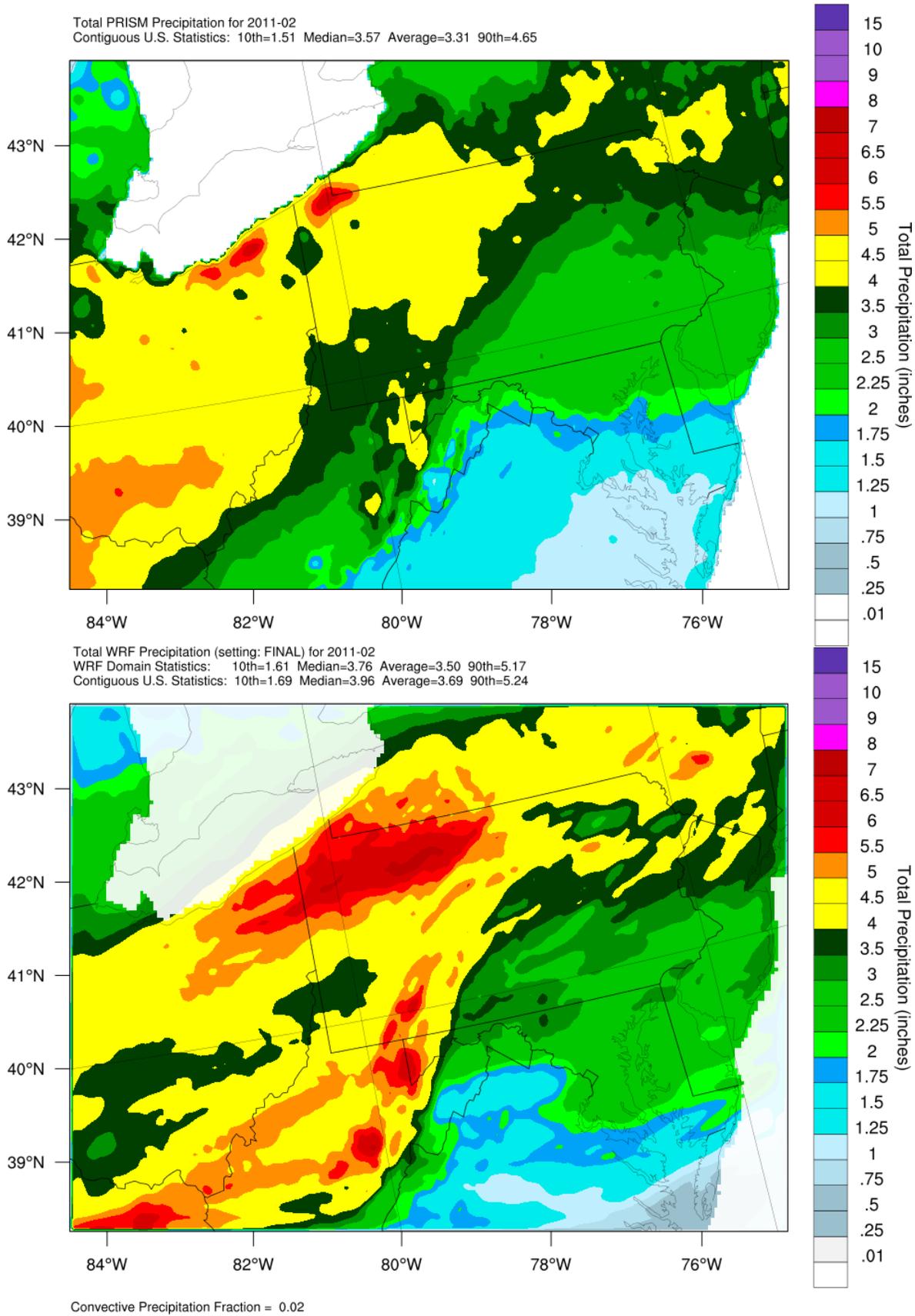


Figure 30. PRISM precipitation total (top) and WRF precipitation total (bottom) for February 2011 in the 4 km domain

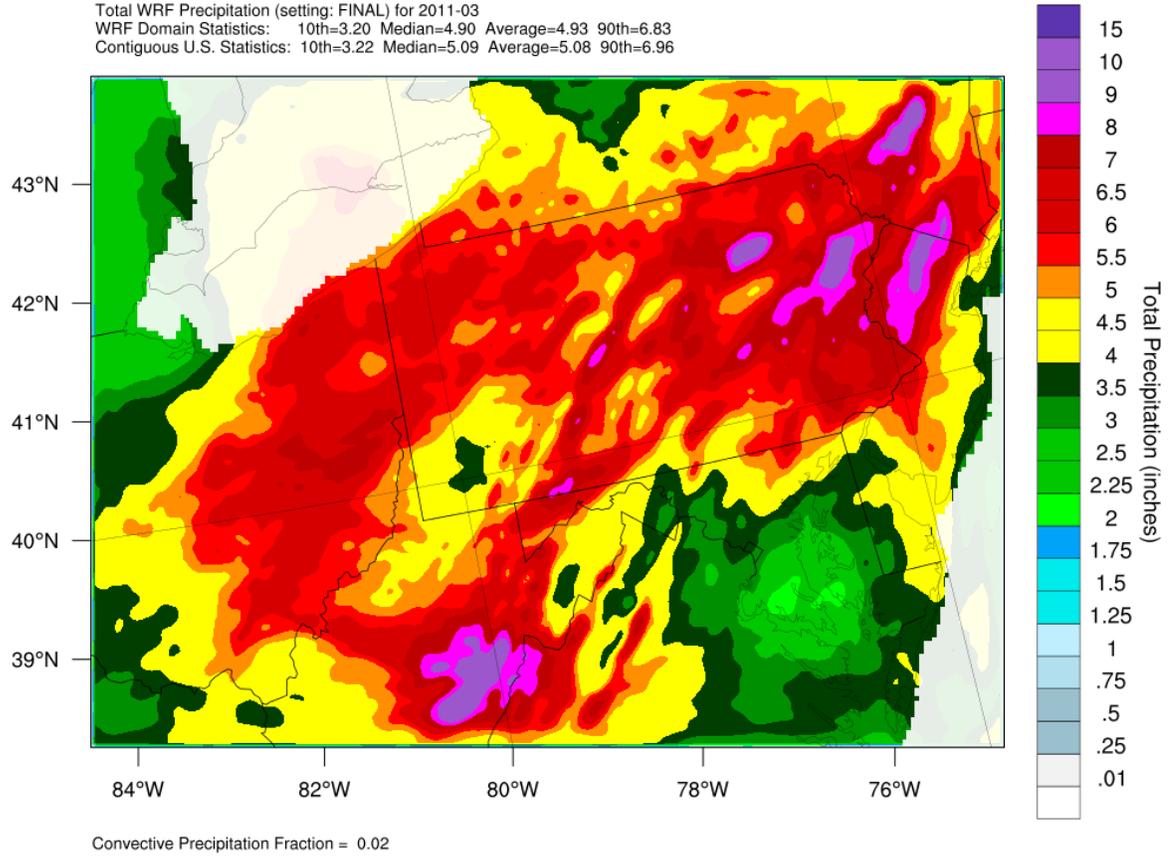
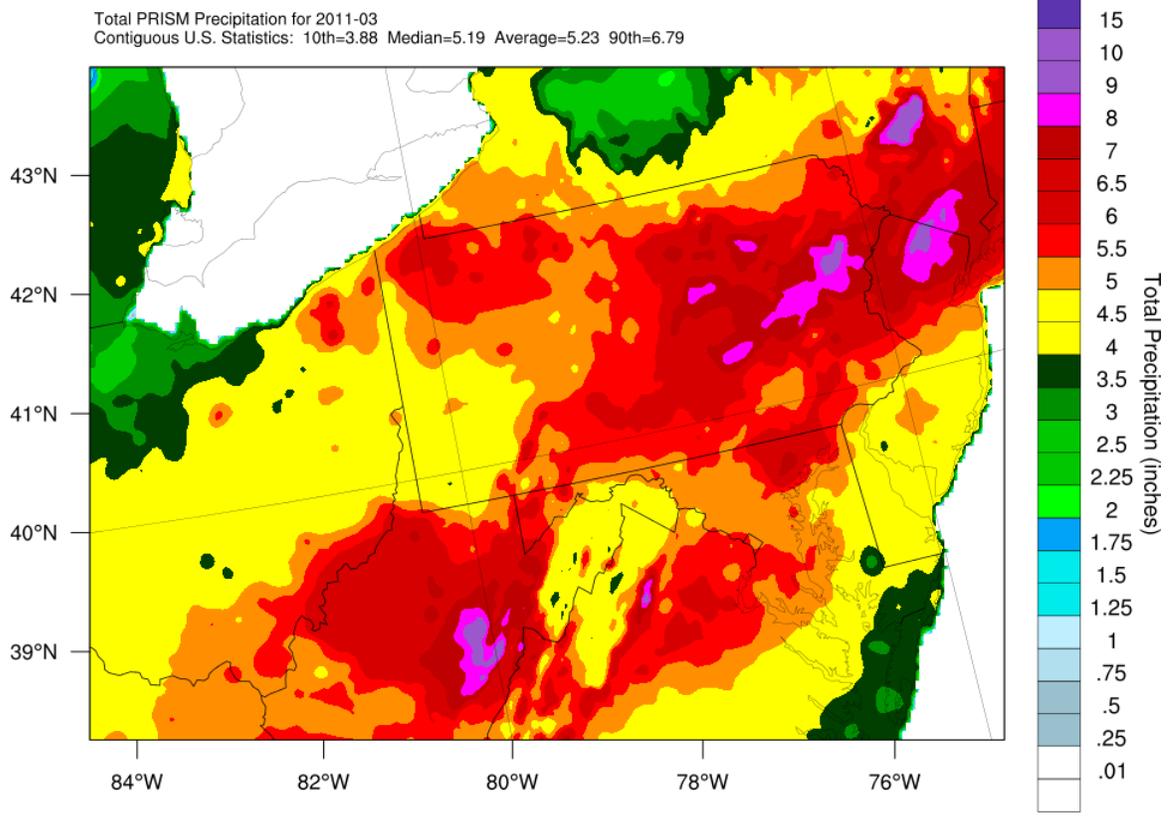
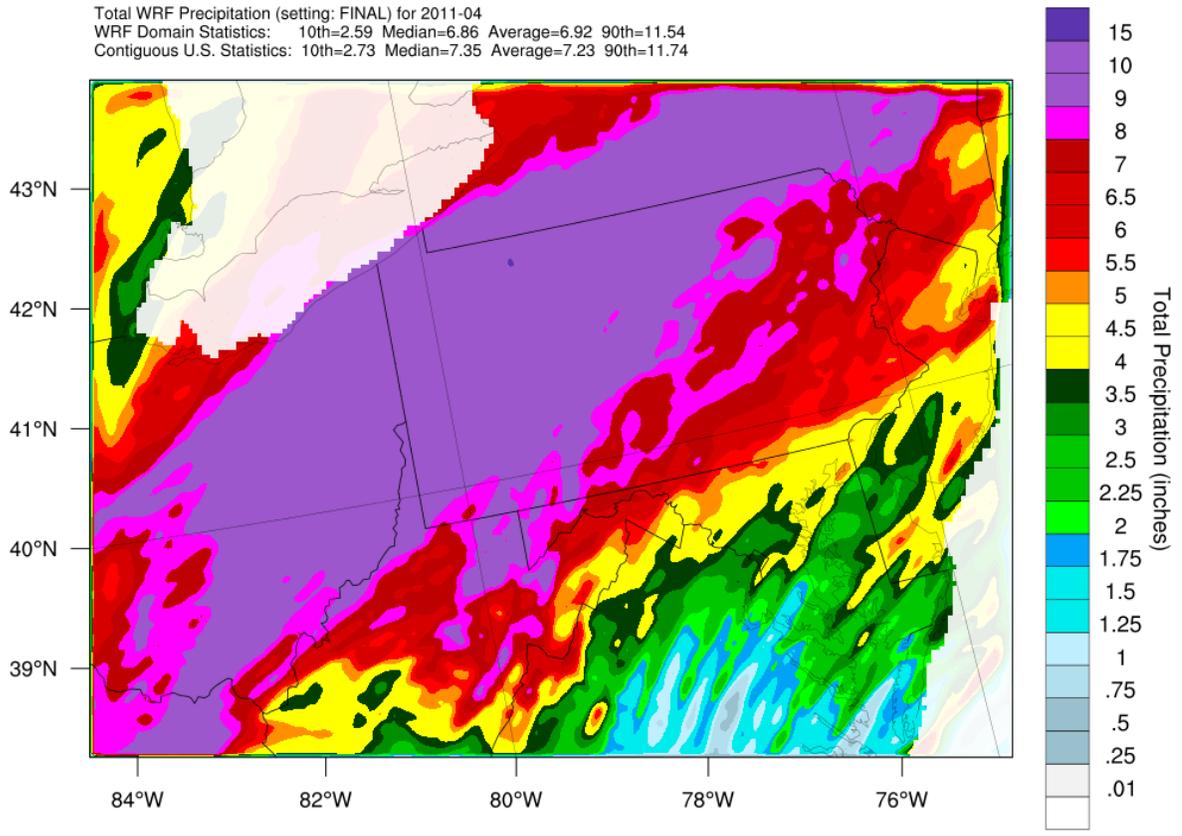
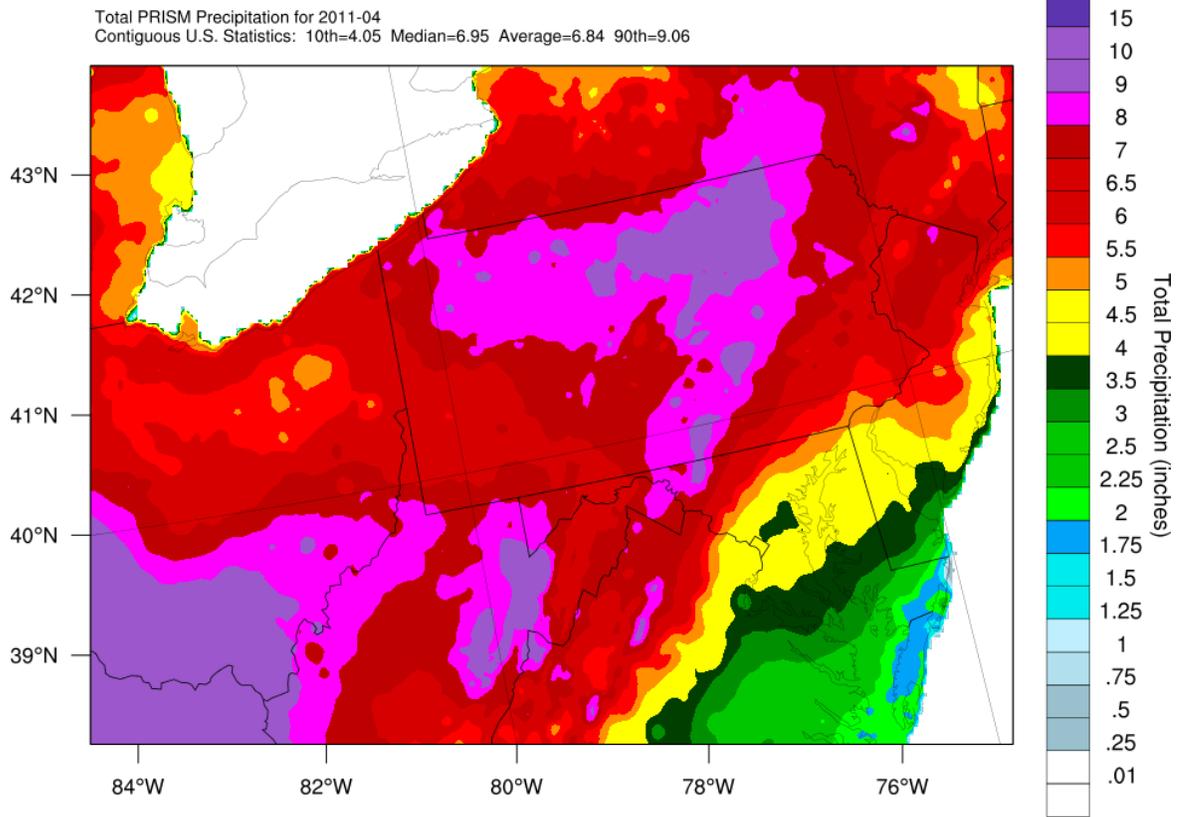


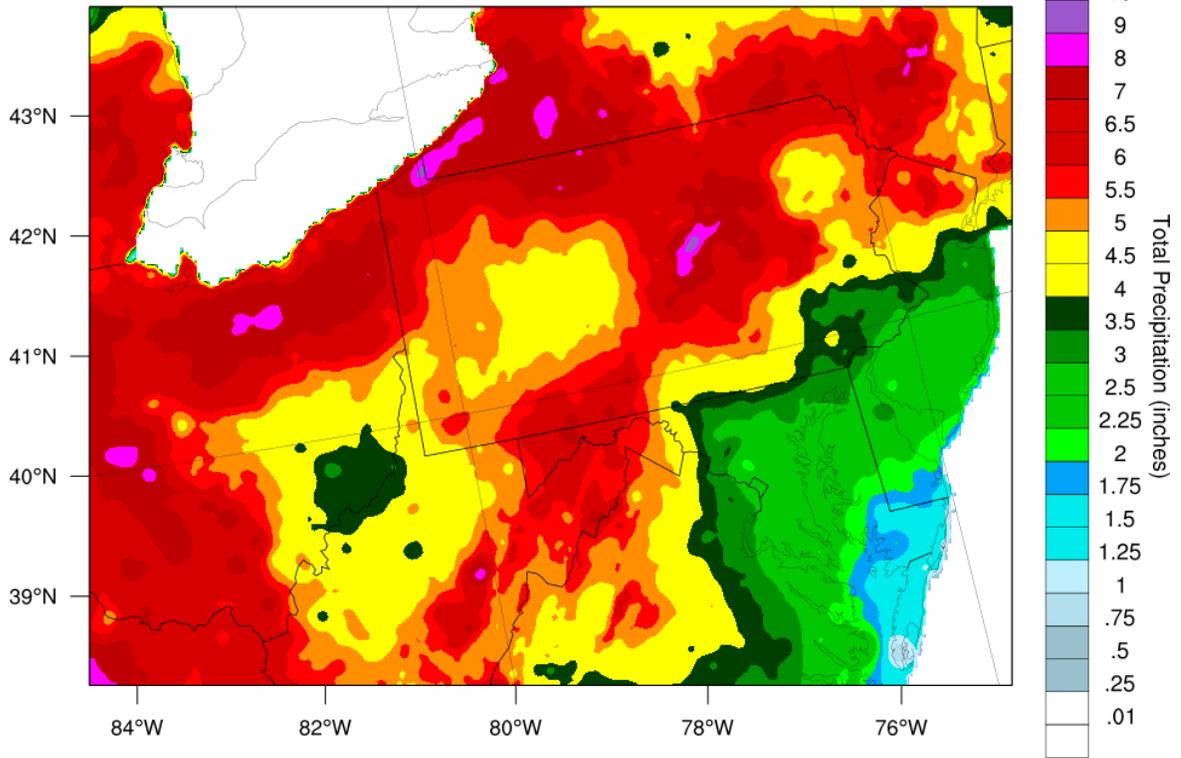
Figure 31. PRISM precipitation total (top) and WRF precipitation total (bottom) for March 2011 in the 4 km domain



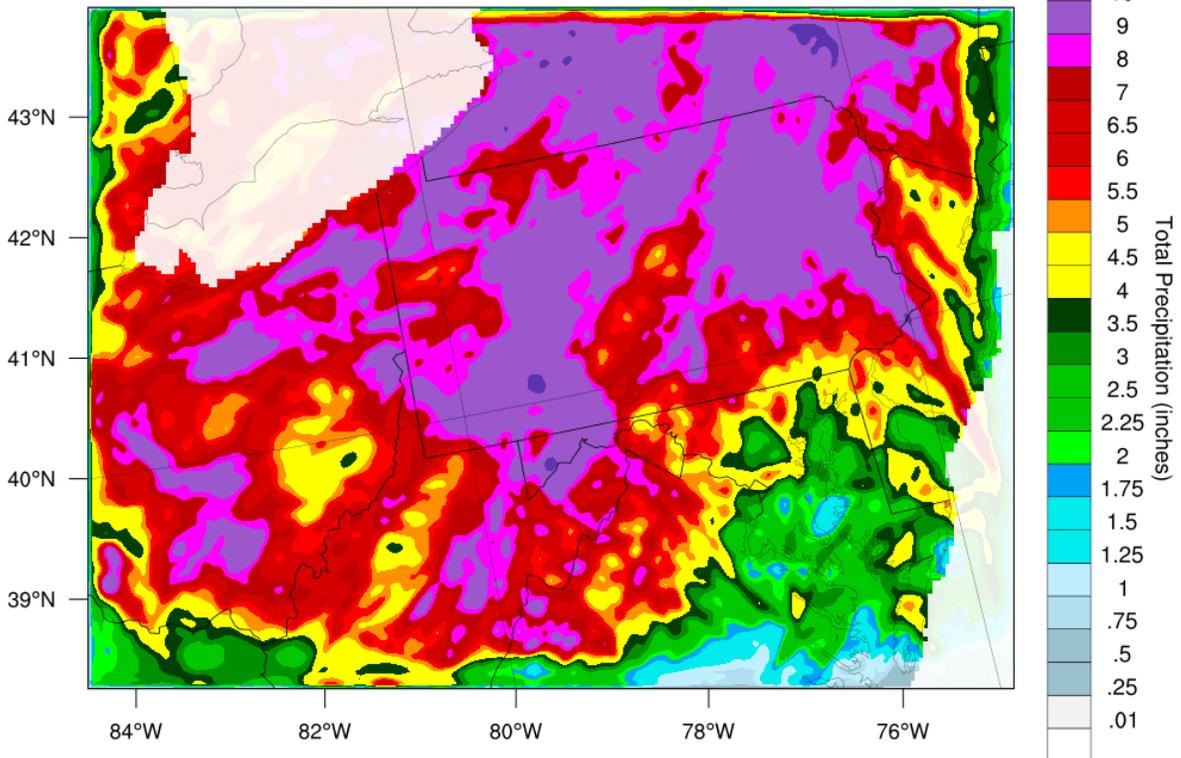
Convective Precipitation Fraction = 0.06

Figure 32. PRISM precipitation total (top) and WRF precipitation total (bottom) for April 2011 in the 12 km domain

Total PRISM Precipitation for 2011-05
 Contiguous U.S. Statistics: 10th=2.74 Median=5.37 Average=5.23 90th=7.14



Total WRF Precipitation (setting: FINAL) for 2011-05
 WRF Domain Statistics: 10th=2.56 Median=6.45 Average=6.45 90th=10.14
 Contiguous U.S. Statistics: 10th=2.72 Median=6.73 Average=6.67 90th=10.36



Convective Precipitation Fraction = 0.10

Figure 33. PRISM precipitation total (top) and WRF precipitation total (bottom) for May 2011 in the 4 km domain

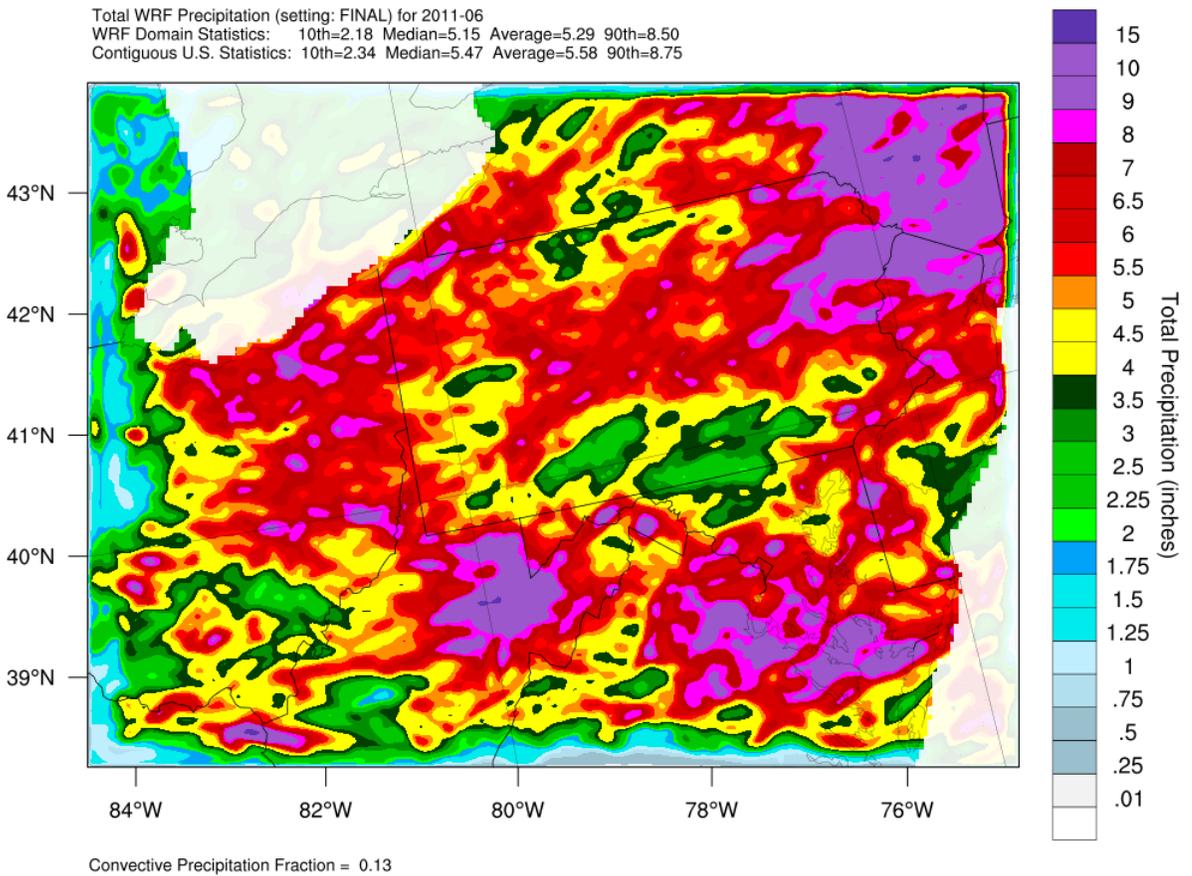
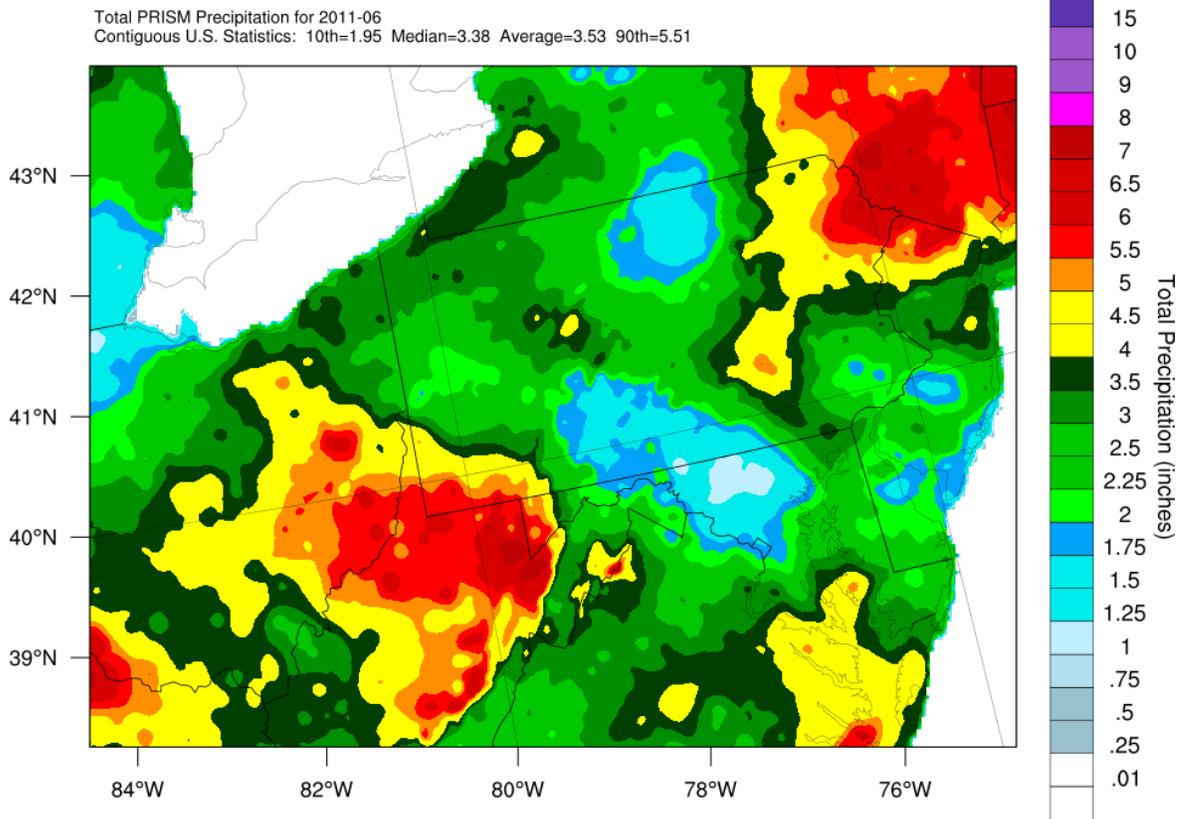


Figure 34. PRISM precipitation total (top) and WRF precipitation total (bottom) for June 2011 in the 4 km domain

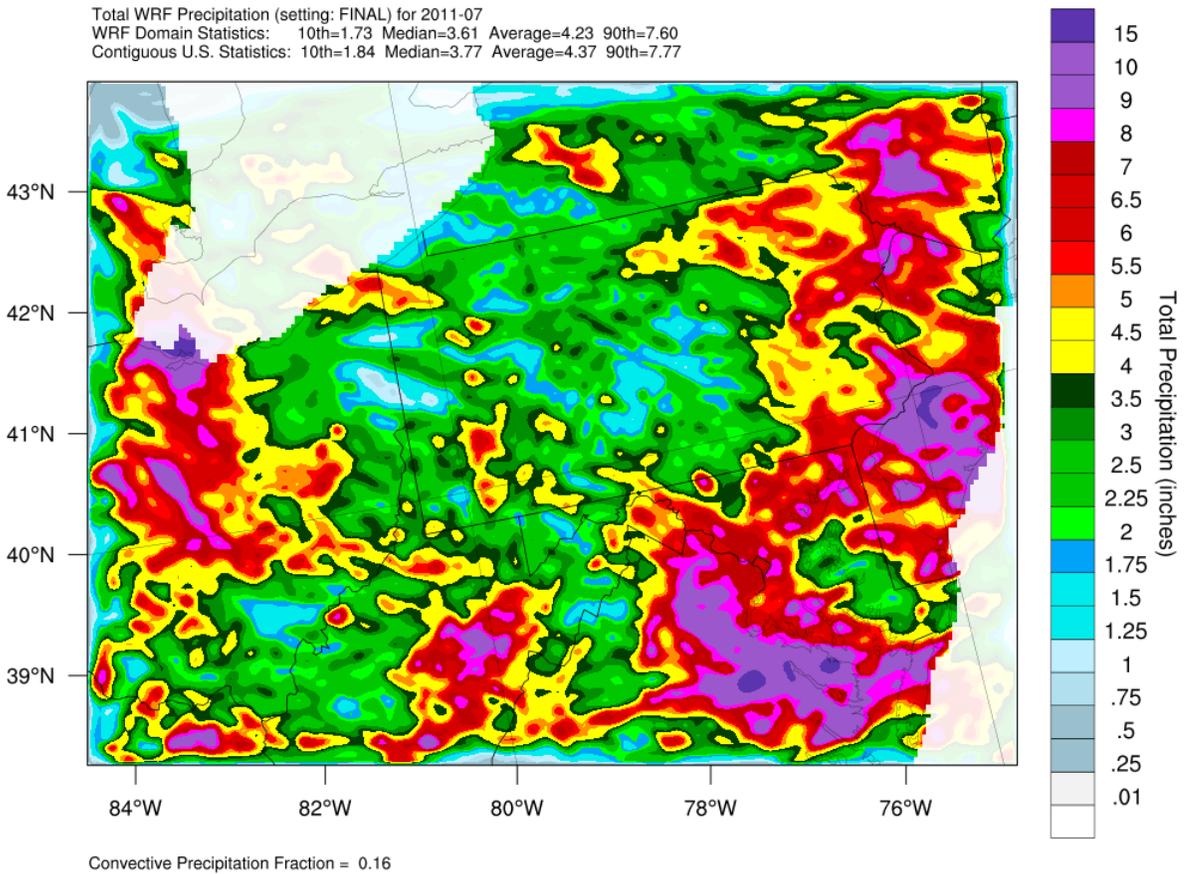
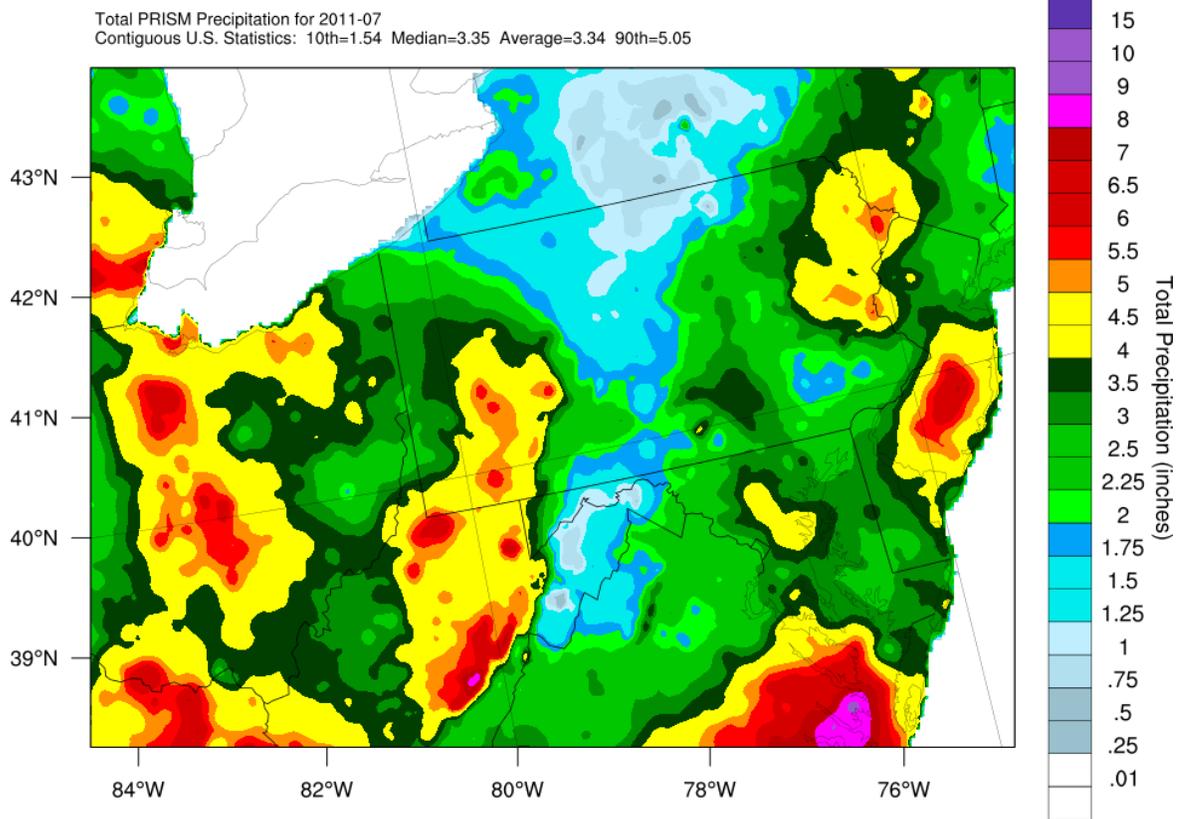
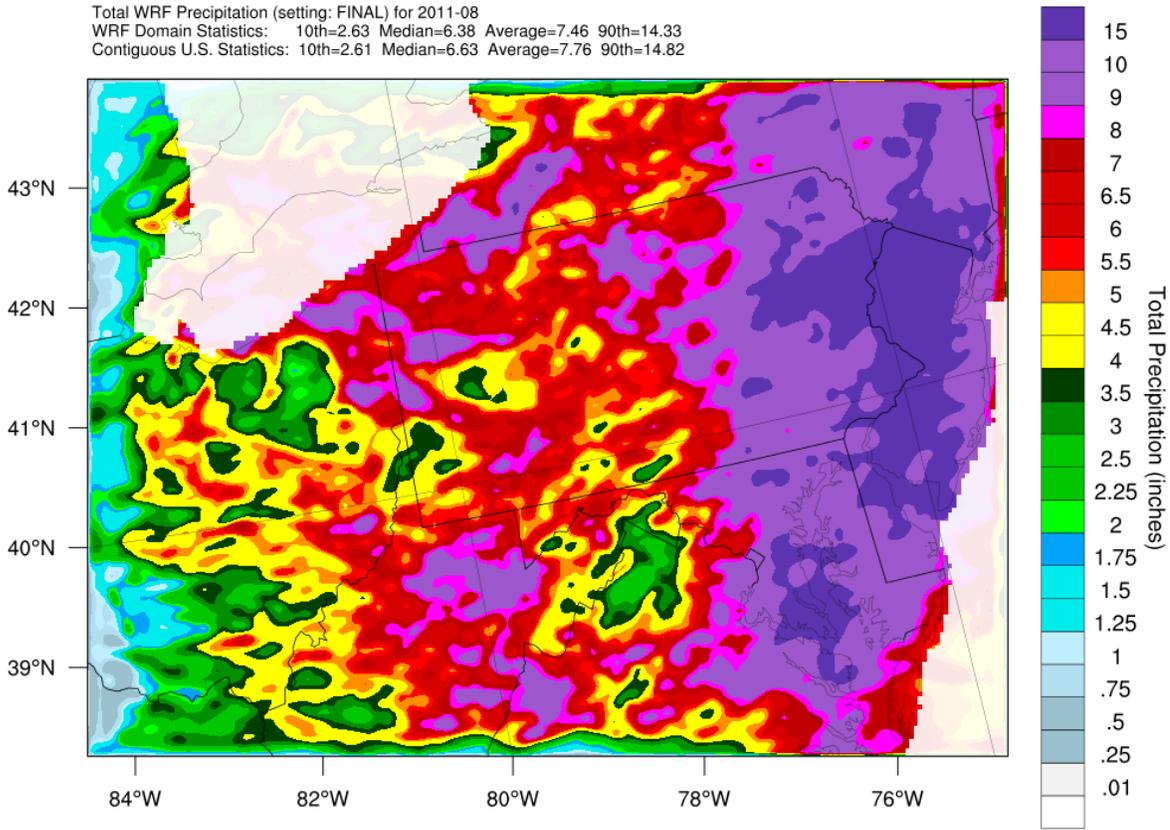
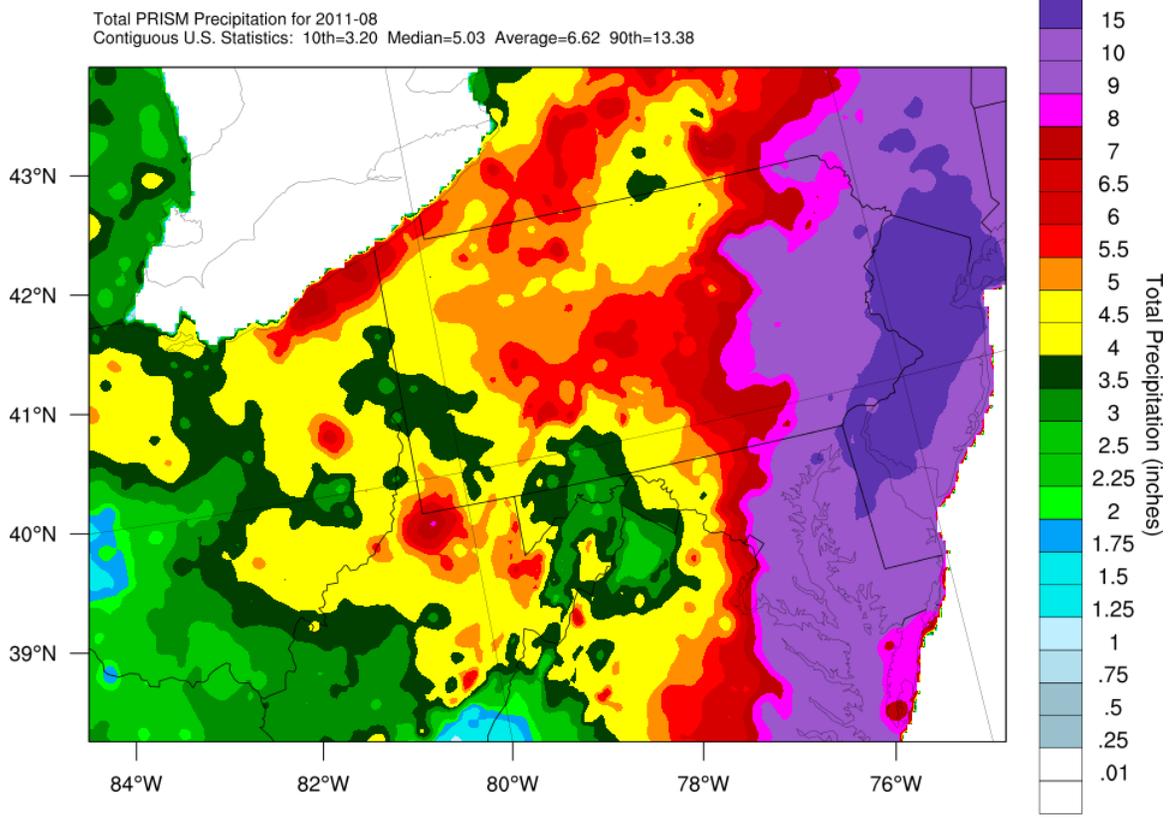


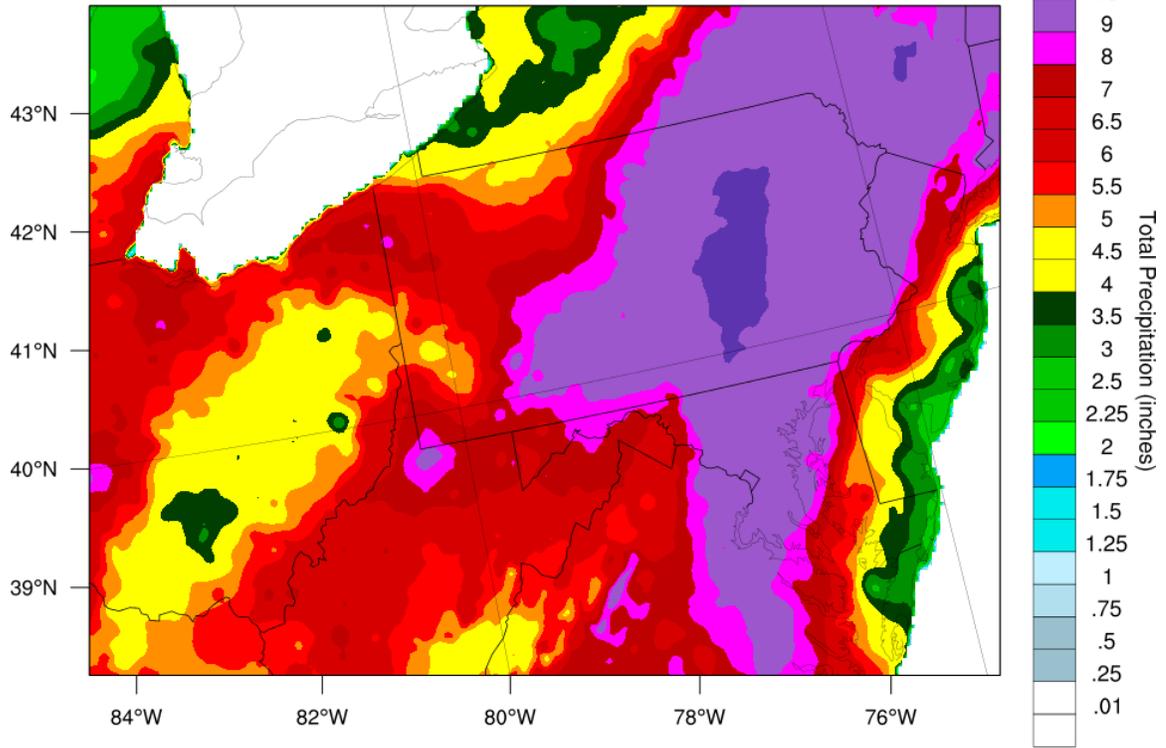
Figure 35. PRISM precipitation total (top) and WRF precipitation total (bottom) for July 2011 in the 4 km domain



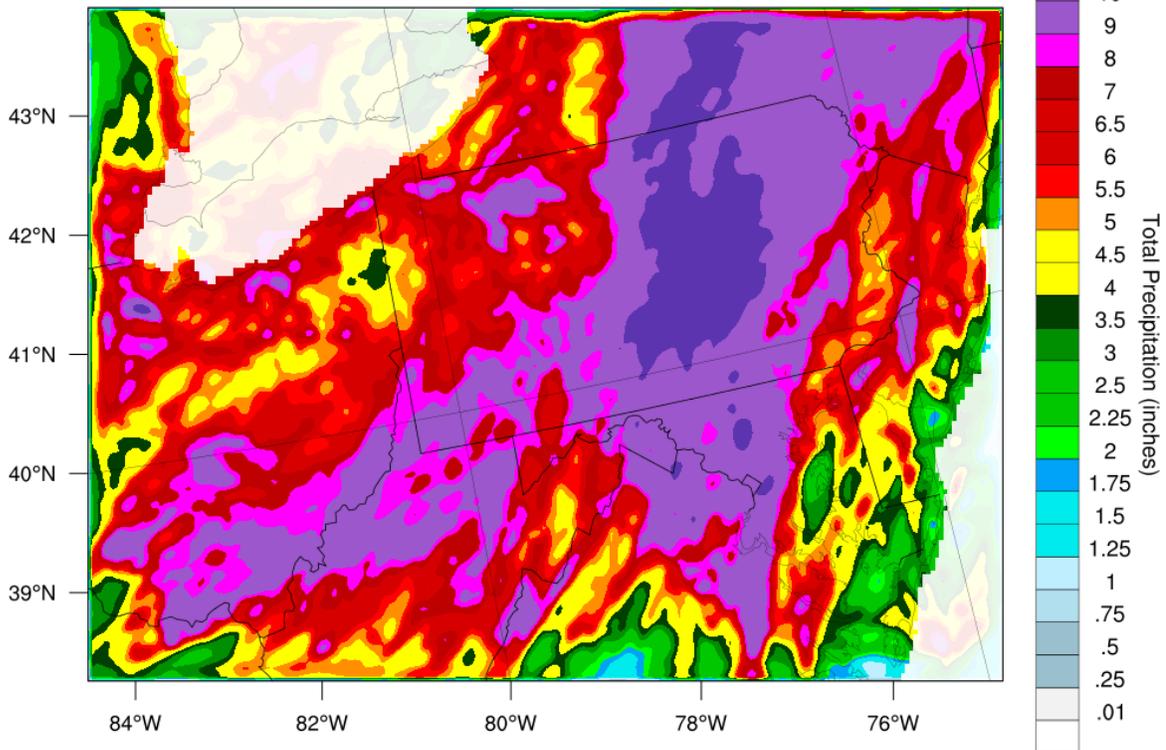
Convective Precipitation Fraction = 0.12

Figure 36. PRISM precipitation total (top) and WRF precipitation total (bottom) for August 2011 in the 4 km domain

Total PRISM Precipitation for 2011-09
 Contiguous U.S. Statistics: 10th=4.19 Median=6.70 Average=7.48 90th=12.02



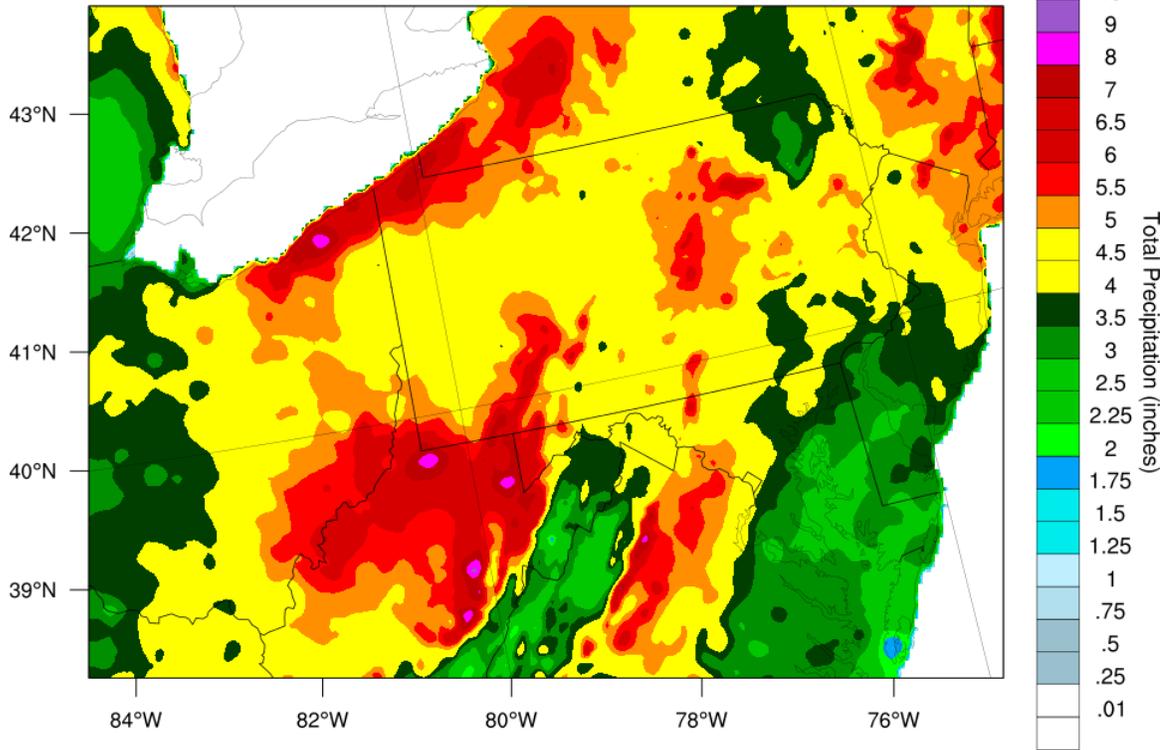
Total WRF Precipitation (setting: FINAL) for 2011-09
 WRF Domain Statistics: 10th=3.36 Median=6.66 Average=7.32 90th=12.21
 Contiguous U.S. Statistics: 10th=3.76 Median=7.09 Average=7.78 90th=12.73



Convective Precipitation Fraction = 0.09

Figure 37. PRISM precipitation total (top) and WRF precipitation total (bottom) for September 2011 in the 4 km domain

Total PRISM Precipitation for 2011-10
 Contiguous U.S. Statistics: 10th=3.21 Median=4.57 Average=4.54 90th=5.81



Total WRF Precipitation (setting: FINAL) for 2011-10
 WRF Domain Statistics: 10th=2.53 Median=4.05 Average=4.01 90th=5.51
 Contiguous U.S. Statistics: 10th=2.59 Median=4.05 Average=4.04 90th=5.57

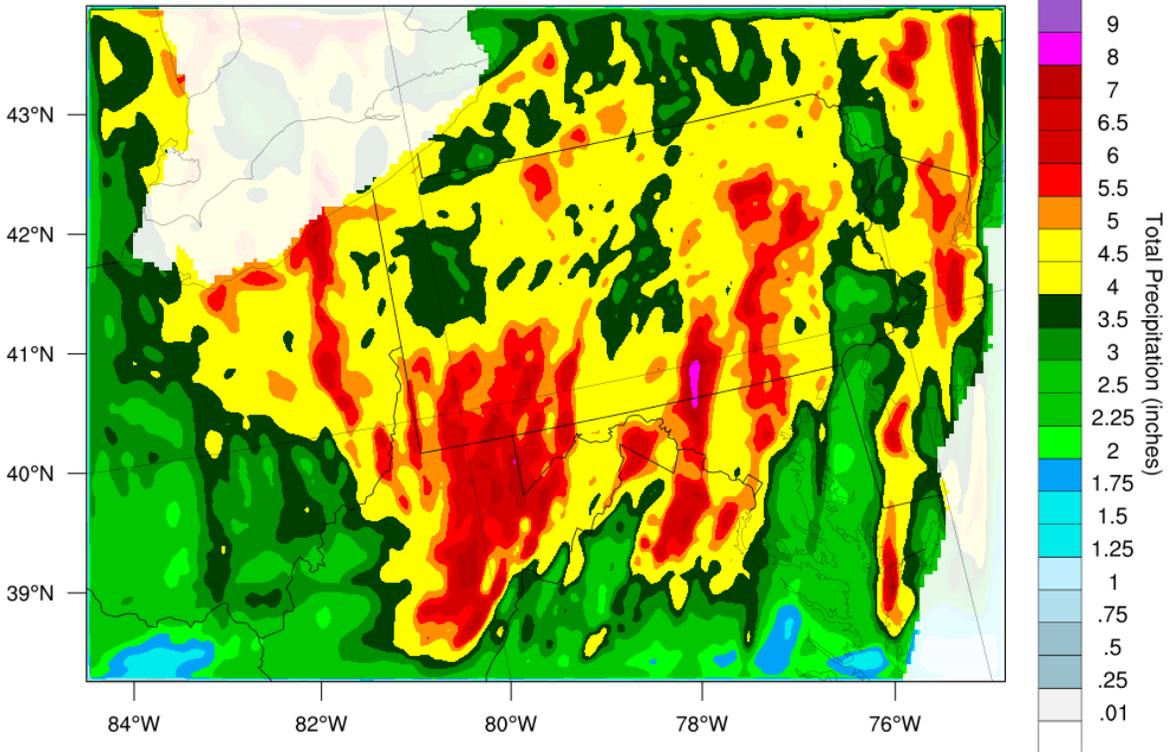


Figure 38. PRISM precipitation total (top) and WRF precipitation total (bottom) for October 2011 in the 4 km domain

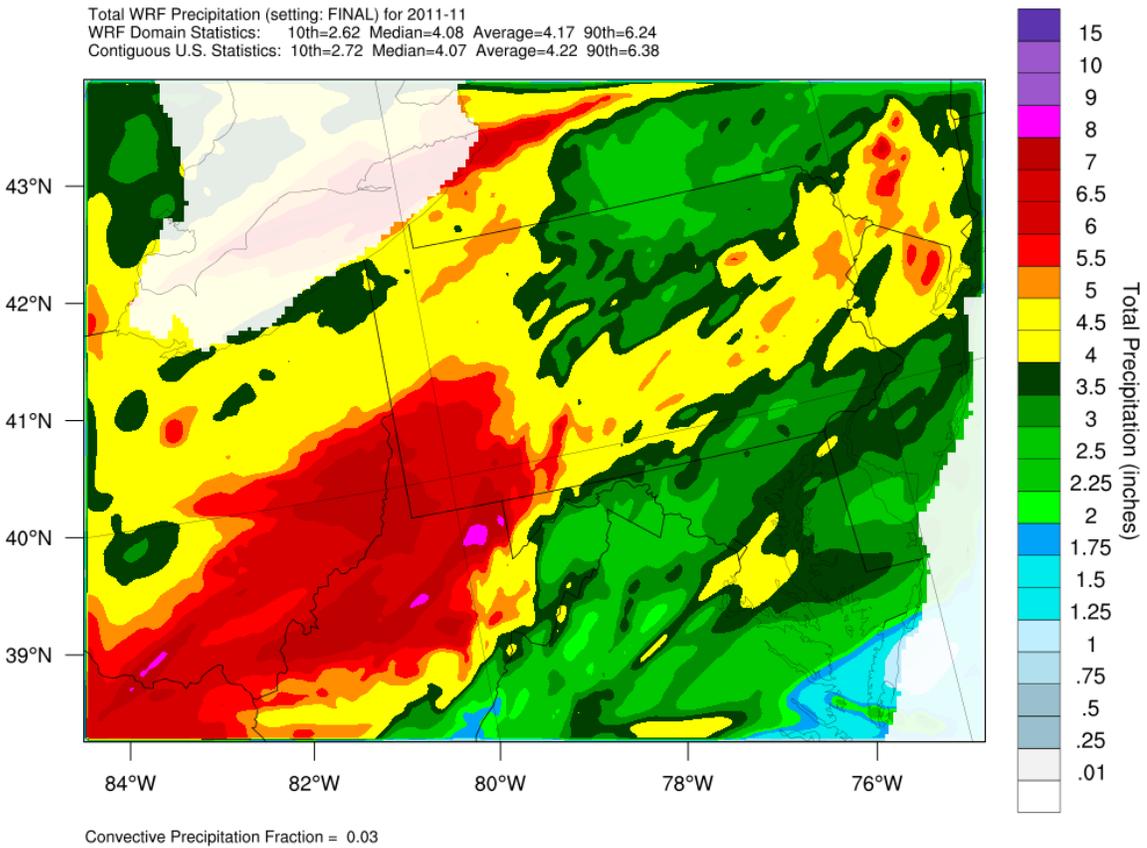
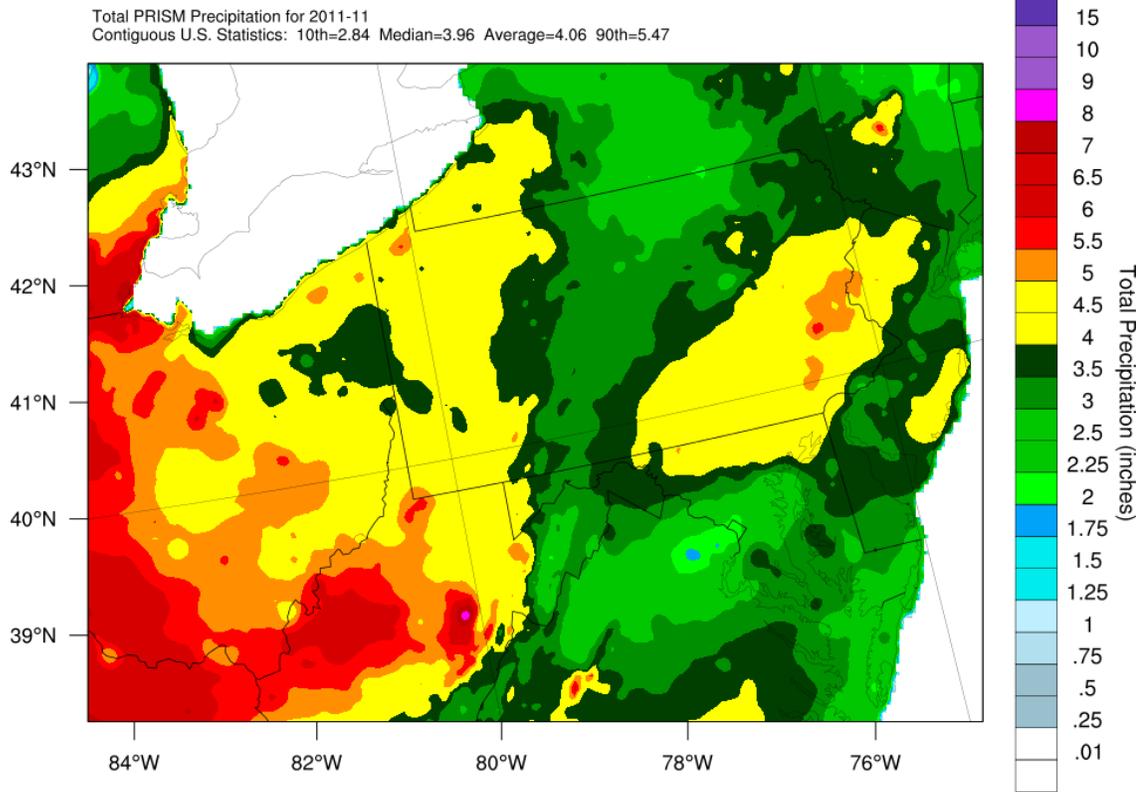
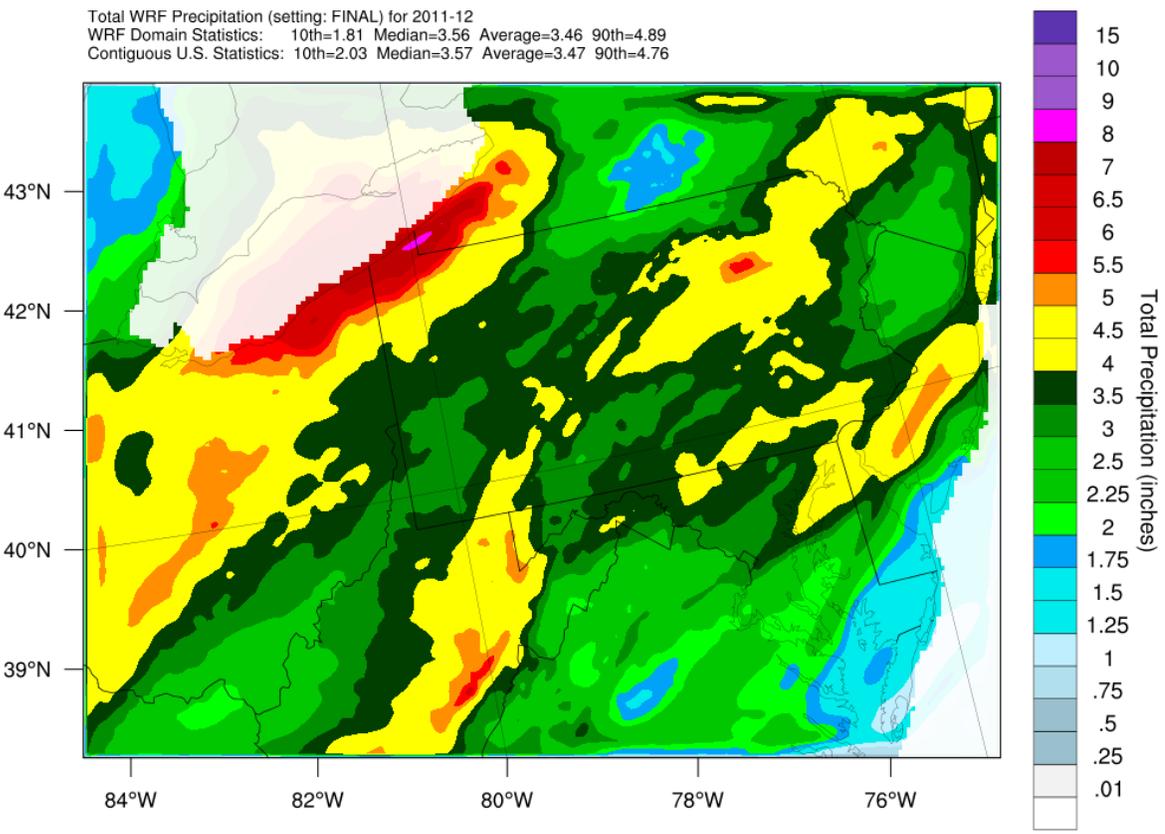
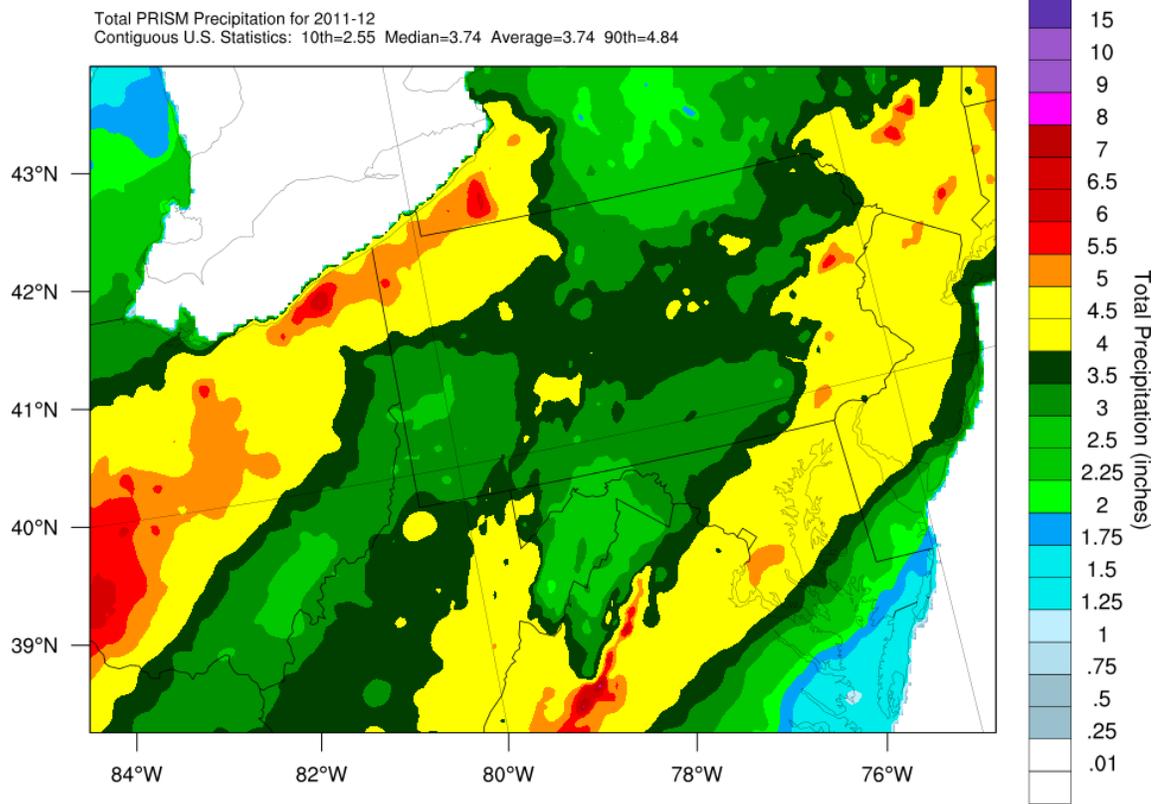


Figure 39. PRISM precipitation total (top) and WRF precipitation total (bottom) for November 2011 in the 4 km domain



Convective Precipitation Fraction = 0.02

Figure 40. PRISM precipitation total (top) and WRF precipitation total (bottom) for December 2011 in the 4 km domain

{This page left blank for printing purposes}

Appendix G.2

CAMx Model Performance Evaluation

{This page left blank for printing purposes}

Prepared for:

Allegheny County Health Department

Prepared by:

Ramboll Environ US Corporation
Lynnwood, Washington
Novato, California

December, 2017 (Rev. 10/18)

Project Number:

06-35892A

**ALLEGHENY COUNTY HEALTH
DEPARTMENT PM_{2.5} STATE
IMPLEMENTATION PLAN
FOR THE 2012 NAAQS
CAM_x BASE CASE MODELING AND MODEL
PERFORMANCE EVALUATION**

CONTENTS

1. Overview	4
2. Modeling Description	5
2.1 Model Selection	5
2.1.1 Meteorological Model Selection	5
2.1.2 Emissions Model Selection	5
2.1.3 Air Quality Model Selection	5
2.1.3.1 CAMx	6
2.1.3.2 AERMOD	8
3. Model Emissions evaluation	10
3.1 Summary of Emission Results	11
4. Model Performance Evaluation (MPE)	12
4.1 Monitoring Networks and Species used in Model Performance Evaluation	12
4.2 Model Performance Statistics and Performance Goals and Criteria	13
4.3 Model Performance Evaluation	16
4.3.1 Spatial Plots of Average PM _{2.5}	17
4.3.2 Total PM _{2.5} Mass	19
4.3.3 Local Source Contributions to Annual Average PM _{2.5}	30
4.3.4 Overview of Speciated PM _{2.5} Model Performance	33
4.3.5 Plume-in-Grid and AERMOD Comparison	37
5. Conclusions	40
6. References	42

TABLES

Table 3-1. Base Year Inventory Data Sources	10
Table 3-2. Summary of Total 2011 Anthropogenic Emissions within the 1.33 km Domain for an Average Day (tons per day).	11
Table 4-1. Ambient Air Quality Monitoring Networks Species used in MPE	13
Table 4-2. Core statistical measures to be used in the Allegheny County air quality model evaluation with ground-level data.	14
Table 4-3. PM Model Performance Goals and Criteria	16
Table 4-4. Comparison of Quarterly FRM, TEOM, and CSN PM _{2.5} Fractional Bias and Error Performance Statistics with PM Performance Goals and Criteria for 1.33 km Domain	20
Table 4-5. Local and Regional Contributions to Annual Average PM _{2.5}	31

FIGURES

Figure 2-1. PiG live puffs sampling grids (blue boxes) that used a 100 m receptor network along with local major sources (red) and FRM PM _{2.5} monitoring sites (yellow) within Allegheny County.	8
Figure 4-1. Quarterly Average PM _{2.5} for 36 km continental U.S. domain	18
Figure 4-2. Quarterly Average PM _{2.5} for 12 km domain.....	19
Figure 4-3. Quarterly PM _{2.5} Fractional Bias and Error Soccer Plots Using Daily FRM observations at 8 Monitoring Sites in Allegheny County	21
Figure 4-4. Quarterly PM _{2.5} Fractional Bias and Error Soccer Plots Using Hourly TEOM observations at Liberty and Lawrenceville.....	22
Figure 4-5. Quarterly Time Series of Daily FRM PM _{2.5} at Liberty.	24
Figure 4-6. Annual Q-Q Plot of FRM and Modeled PM _{2.5} at Liberty	26
Figure 4-7. Annual Q-Q Plot of FRM and Modeled PM _{2.5} at Clairton, Lawrenceville, North Braddock, and South Fayette.	27
Figure 4-8. Quarterly Scatterplots of Modeled PM _{2.5} and FRM Observations from 8 Monitoring Sites in Allegheny County.....	29
Figure 4-9. Quarterly Scatterplots of Modeled PM _{2.5} and TEOM Observations at Liberty and Lawrenceville.....	30
Figure 4-10. Speciated "Excess" PM _{2.5} at Liberty, Observed and Modeled, 2011.	32
Figure 4-11. Quarterly SO ₄ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.....	34
Figure 4-12. Quarterly NO ₃ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.....	35
Figure 4-13. Quarterly NH ₄ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.....	36
Figure 4-14. Quarterly Total Carbon Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.....	37
Figure 4-15. Annual Q-Q plot Comparing 24-hour PM _{2.5} at Liberty in the Base Year Model Runs.....	38

1. OVERVIEW

This document presents the 2011 base case model performance evaluation to support PM_{2.5} attainment demonstration modeling of the Allegheny County, PA nonattainment area (NAA). The Comprehensive Air Quality Model with extensions (CAMx¹) photochemical grid model (PGM) was used to simulate calendar year 2011 using a 36/12/4/1.33 km resolution nested grid structure with the 1.33 km domain focused on the Allegheny County NAA and the 4 km domain covering all of Pennsylvania and parts of surrounding states. CAMx includes a Plume-in-Grid (PiG) option that was used to simulate near-source dispersion and chemical transformation for emissions from local sources and a Particulate Source Apportionment Technology (PSAT) that was used to track contributions of local sources within the CAMx PGM.

This report presents a model performance evaluation of the CAMx 2011 base case simulation. More details about the project background and the CAMx modeling configuration, including meteorological and emissions model inputs, can be found in the accompanying CAMx Modeling Protocol (Ramboll Environ, 2017). This modeling framework was used to evaluate whether the emissions control strategy proposed by Allegheny County will lead to attainment of the 2012 annual National Ambient Air Quality Standard (NAAQS) for PM_{2.5} by December, 2021. This report evaluates CAMx model performance for the base year (2011) scenario to aid in determining if CAMx can be relied upon to make future year PM_{2.5} projections in Allegheny County. Modeled total PM_{2.5} mass and speciated PM_{2.5} are in general agreement with observations in Allegheny County, suggesting that CAMx is suitable to aid in development of the Allegheny County PM_{2.5} State Implementation Plan (SIP).

⁽¹⁾ <http://www.camx.com>

2. MODELING DESCRIPTION

2.1 Model Selection

Three types of models are used in the Allegheny County PM_{2.5} attainment demonstration modeling to simulate emissions, meteorology, and air quality. The 2011 calendar year was selected for the baseline modeling and model performance evaluation. Discussion on the rationale and justification for their selection, as well as detailed information about each model, are provided in the CAMx Modeling Protocol (Ramboll Environ, 2017).

2.1.1 Meteorological Model Selection

The Weather Research and Forecasting (WRF) model² is used to represent meteorological conditions in this study area. The Advanced Research WRF (ARW) version of WRF was selected for the Allegheny County PM_{2.5} attainment demonstration modeling and is further described in a separate WRF Modeling Protocol (Ramboll Environ, 2016).

2.1.2 Emissions Model Selection

A suite of emission models are used to generate air quality model-ready emissions for various source categories. The air quality model requires hourly, gridded, and speciated emissions inputs. The Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system³ is an emissions processing system that generates hourly gridded speciated emission inputs (Coats, 1995; Houyoux et al., 2000). SMOKE was used to process anthropogenic emissions that were available through the EPA 2001 National Emissions Inventory (NEI), MARAMA and ACHD to the hourly gridded speciated inputs needed by the PGM. Biogenic emissions were processed using SMOKE-BEIS3.

2.1.3 Air Quality Model Selection

The CAMx photochemical grid model was selected for the Allegheny County PM_{2.5} attainment demonstration modeling. In addition, the impacts of primary PM emissions from local sources were also simulated using the AERMOD Gaussian plume model.

(2) <https://www.mmm.ucar.edu/weather-research-and-forecasting-model>

(3) <https://www.cmascenter.org/smoke>

2.1.3.1 CAMx

The Comprehensive Air Quality Model with extensions version 6.30 (CAMx⁴; Ramboll Environ, 2016) photochemical grid model was used for this analysis. The Particulate Source Appointment Technology (PSAT) feature was used to track source groups, including PM_{2.5} contributions from local source components, and the sub-grid scale Plume-in-Grid (PiG) feature was used to simulate dispersion of major sources in Allegheny County at very fine resolution.

CAMx was applied on the 36/12 km domains using two-way grid nesting and the results were post-processed to generate boundary condition (BC) inputs for the 4 km Pennsylvania domain (i.e., one-way grid nesting between the 12 and 4 km domains). CAMx was then applied on the 4/1.33 km domains using two-way grid nesting. The proposed modeling domains were devised to include all the major area and point sources of NO_x, SO₂ and PM_{2.5} emissions in Allegheny County.

The following list summarizes the CAMx inputs used for the 2011 base case model run.

- Meteorological Inputs: The WRF-derived meteorological fields were prepared for CAMx using the latest version of WRFCAMx. The CMAQ-Kv method along with the KVPATCH adjustment was used to generate vertical diffusivity for CAMx.
- Initial/Boundary Conditions: The boundary conditions (BCs) for the 36 km CONUS domain simulation were extracted from a global chemistry model, GEOS-Chem. Existing programs were used to interpolate from the GEOS-Chem horizontal and vertical coordinate system to the CAMx LCC coordinate system and vertical layer structure and to map the GEOS-Chem chemical species to the CB6r2 chemical mechanism.
- Photolysis Rates: Photolysis inputs as well as albedo/haze/ozone/snow inputs for CAMx were based on the Total Ozone Mapping Spectrometer (TOMS) data. The TUV processor was used to generate clear-sky photolysis rates which will be adjusted for the presence of clouds and aerosols by CAMx.
- Landuse: Landuse fields were generated based on USGS GIRAS data.
- Spin-Up Initialization: Ten days of model spin up were used to eliminate the influences of the initial concentrations for the 36/12 km configuration and 5 days of spin-up were used for the 4/1.33 km simulation.

⁽⁴⁾ <http://www.camx.com>

- **Plume-in-Grid:** The Plume-in-Grid (PiG) subgrid-scale plume module was used to simulate the near-source plume chemistry and dispersion for all identified local major sources in Allegheny County for the 4/1.33 km model run. Sampling receptor grids with a 100-meter grid resolution were used to obtain concentrations due to the PiG puffs in the vicinity of the key PM monitoring sites in Allegheny County, as shown in Figure 2-1.
- **PSAT:** The Particulate Source Apportionment Technology (PSAT) was used for the identified local sources in Allegheny County in the 4/1.33 km model run to separate track their contributions to PM concentrations in the CAMx PGM. For the 2011 base case, PSAT is used to separate track the PM contributions due to the five Source Groups: (1) Local Major Source primary PM emissions; (2) Local Major Source gaseous PM precursor emissions; (3) Local Minor Source primary PM and gaseous emissions; (4) remainder emissions inside Allegheny County; and (5) source outside of Allegheny County but within the 4/1.33 km modeling domains.

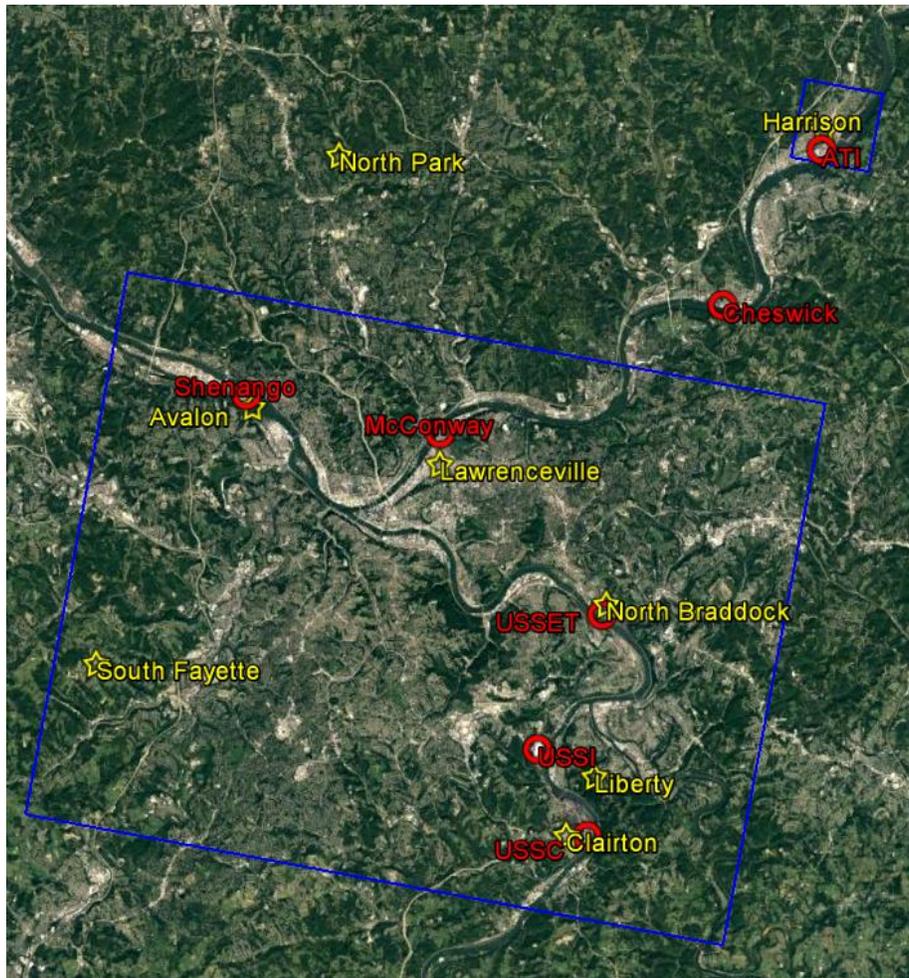


Figure 2-1. PiG live puffs sampling grids (blue boxes) that used a 100 m receptor network along with local major sources (red) and FRM PM_{2.5} monitoring sites (yellow) within Allegheny County.

Figure 2-1 shows the two PiG sampling grids used in the 4/1.33 km CAMx modeling run. Within each sampling grid (blue boxes), a receptor network with 100-m spacing is used to sample live PiG puff concentrations associated with major local sources within Allegheny County. These sources are shown in red in Figure 2-1 and PM_{2.5} monitoring sites within the county are shown in yellow.

2.1.3.2 AERMOD

The AERMOD Gaussian plume model was also used to simulate the effects of primary PM_{2.5} emissions from local sources on PM_{2.5} concentrations at the Liberty monitoring site. The unique, sub-grid cell terrain features surrounding Liberty and

its proximity to major PM_{2.5} sources can make it difficult to accurately model PM_{2.5} concentrations in the region and at Liberty in particular.

As discussed in the Modeling Protocol, the AERMOD model can be used to estimate PM_{2.5} contributions associated with primary PM_{2.5} emissions from local major sources. This is an alternative approach to using PiG/PSAT treatment to obtain the contributions of local sources. It is informative to compare the model-estimated local PM_{2.5} from both AERMOD and PiG/PSAT approaches and utilize both methods to calculate PM_{2.5} design values in Allegheny County.

The AERMOD model (version 16216r) was applied for the local sources associated with the 2011 base case emissions scenario. AERMOD meteorological inputs were generated by passing 2011 WRF model output through the Mesoscale Model Interface Program⁵ for both 1.33 km and 0.444 km resolution WRF simulations. Terrain elevations for receptors were prepared using 1/3th arc-second elevation data from the National Elevation Dataset (NED), which is a product of the United States Geological Survey (USGS). The NED is a seamless elevation dataset covering the continental United States, Alaska, and Hawaii. The datasets have a horizontal spatial resolution of approximately 10 meters. For the dispersion model analyses, receptors were placed at the location of every PiG sampling grid cell center (100-m spacing) across both PiG sampling grids. The base elevation and hill height scale for each receptor were determined using the EPA's terrain processor AERMAP (Version 11103), which generated the receptor files that are read by AERMOD. All receptor locations are in the Universal Transverse Mercator (UTM) coordinate system, using the spatial reference of NAD 83, Zone 11.

The AERMOD local source contributions were combined with the CAMx regional source contributions in a mass consistent fashion to provide total concentration estimates (i.e., using AERMOD results to replace the CAMx PiG/PSAT local source contributions).

⁵ https://www3.epa.gov/ttn/scram/models/relat/mmif/MMIFv3.3_Users_Manual.pdf

3. MODEL EMISSIONS EVALUATION

Natural and anthropogenic emissions inventories are converted into CAMx-ready emission files using the Sparse Matrix Operator Kernel Emissions (SMOKE) system and other methods. The emissions inventories and processing techniques are described in more detail in the Modeling Protocol (Ramboll Environ, 2017).

Local sources in Allegheny County are based on actual emissions data reported for 2011 and include stacks, quench towers, ambient-temperature fugitives, and coke oven batteries. The local major point sources in Allegheny County are modeled using the PiG option in CAMx. Local sources are also flagged for treatment using the PSAT source apportionment technology.

Table 3-1 summarizes inventory data sources by source category and region for the 2011 base year.

Table 3-1. Base Year Inventory Data Sources

Source Category	Allegheny Cty (1.3 km Domain)	Mid-Atlantic (4 km Domain)	Eastern U.S. (12 km Domain)	Continental U.S. (36 km Domain)
Area / Nonroad	MARAMA Alpha2 2011	MARAMA Alpha2 2011	EPA 2011 v6.2 MP	EPA 2011 v6.2 MP
Onroad (Mobile)	MARAMA Alpha2 2011	MARAMA Alpha2 2011	EPA 2011 v6.2 MP	EPA 2011 v6.2 MP
Point	ACHD Local + MARAMA Alpha2 2011	ACHD Local + MARAMA Alpha2 2011	EPA 2011 v6.2 MP	EPA 2011 v6.2 MP
EGU Point	EPA 2011 v6.2 MP 2011 w/CAMD CEMS	EPA 2011 v6.2 MP 2011 w/CAMD CEMS	EPA 2011 v6.2 MP 2011 w/CAMD CEMS	EPA 2011 v6.2 MP 2011 w/CAMD CEMS
Fires	EPA 2011 v6.2 FIRES	EPA 2011 v6.2 FIRES	EPA 2011 v6.2 FIRES	EPA 2011 v6.2 FIRES
Biogenics	EPA 2011 NEIv2 BEIS	EPA 2011 NEIv2 BEIS	EPA 2011 NEIv2 BEIS	EPA 2011 NEIv2 BEIS
Sea Salt and Lightning	CAMx processors	CAMx processors	CAMx processors	CAMx processors

Notes:

1. MARAMA Alpha2 and EPA v6.2 MP are developed from 2011 NEI V2
2. Point sources include non-EGUs and small EGU.
3. EGU emissions include SO₂/NO_x CAMD CEMS data for temporal profile; EPA 2011 (annualized) for other pollutants
4. ACHD Local is corrected MARAMA inventory for emissions, stack parameters, coordinates, etc. 36/12 km domains are used to develop boundary conditions for 4/1.33 km domains

3.1 Summary of Emission Results

Table 3-2 summarizes the anthropogenic emissions within the 1.33 km domain by major source category for an average day in 2011. The largest source of anthropogenic CO emissions in the 1.33 km domain is the on-road mobile sector (49%), followed by non-road mobile sources (28%). On-road mobile sources also contribute the most to NO_x emissions (43%) in the 1.33 km domain, followed by point sources (27%). Area (47%) and on-road (29%) sources are the largest two contributors to VOC emissions in the 1.33 km domain with non-road mobile sources (17%) contributing most of the remainder. Point sources contribute the majority of SO₂ (82%) in the 1.33 km domain while area sources (46%) and point sources (40%) both contribute heavily to PM_{2.5} emissions. Area sources dominate NH₃ emissions (73%), followed by on-road mobile sources (18%).

Table 3-2. Summary of Total 2011 Anthropogenic Emissions within the 1.33 km Domain for an Average Day (tons per day).

Source Category	CO	NO_x	VOC	SO₂	PM_{2.5}	NH₃
Area Source	67.1	25.6	54.0	8.2	11.5	5.6
Non-Road	199.9	16.6	20.1	0.1	1.5	0.02
Onroad Mobile	343.9	61.3	34.1	0.4	2.0	1.4
Point Source	91.1	38.1	7.8	39.8	9.9	0.6

4. MODEL PERFORMANCE EVALUATION (MPE)

This section presents an evaluation of air quality model performance. Meteorological model performance is described in a separate WRF model performance evaluation report. This evaluation uses a suite of statistical metrics to assess how accurately CAMx predicts observed concentrations of PM_{2.5}, PM_{2.5} precursors, and ozone. The correct simulation of gas-phase oxidant species is needed for PM since correct, unbiased simulation of gas-phase photochemistry is a necessary element of reliable secondary PM predictions. It is important to evaluate total and speciated PM_{2.5} model performance in order to assure that the model is predicting total PM_{2.5} concentrations for the right reasons and because the Software for Model Attainment (Community Edition) (SMAT-CE; Abt Associates, Inc., 2014, EPA, 2014) used to project future year PM_{2.5} Design Values used both total PM_{2.5} mass and speciated PM_{2.5} observations.

This model performance evaluation focuses on Allegheny County for the entire year and also on a quarterly basis to help build confidence that the modeling system is operating correctly. Good agreement between modeled concentrations and observations suggests that CAMx is suitable to estimate realistic future year design value concentrations in Allegheny County.

4.1 Monitoring Networks and Species used in Model Performance Evaluation

The CAMx model was evaluated using ambient observational data from the following 3 monitoring networks:

- EPA's Air Quality System (AQS) database⁶;
- Federal Reference Method (FRM⁷) total PM_{2.5} Mass (which is part of the AQS); and
- Chemical Speciation Network (CSN⁸) speciated PM_{2.5} (which is part of the AQS).

Table 4-1 lists the ambient air quality monitoring network species used in the modeling performance evaluation (MPE) along with their sampling frequency.

⁽⁶⁾ <https://www.epa.gov/outdoor-air-quality-data>

⁽⁷⁾ <http://www.epa.gov/ttnamti1/pmfrm.html>

⁽⁸⁾ <https://www3.epa.gov/ttn/amtic/speciepg.html>

Table 4-1. Ambient Air Quality Monitoring Networks Species used in MPE

Network	Species	Sampling Time
AQS	TEOM PM _{2.5} , SO ₂ , NO ₂ , O ₃	Hourly
FRM	Total PM _{2.5} Mass	Daily
CSN	PM _{2.5} , SO ₄ , NO ₃ , NH ₄ , EC, OC, elements	Every several days

Since each monitoring network employs a unique measurement approach that “measures” a different amount of a given species, especially for PM species, CAMx is evaluated separately for each monitoring network. Additionally, there is often ambiguity in the mapping of modeled PM species to measurements. For example, PM monitors measure only the carbon component of OC, whereas in the model, the entire mass of organics (OA or Organic Aerosol) is simulated, which includes carbon and the other elements attached to the carbon (e.g., hydrogen and oxygen). An OA/OC factor of 1.4 has been used in this analysis; see the Modeling Protocol for more details. In this evaluation, the other PM_{2.5} (OPM_{2.5}) group is calculated as the difference between the total PM_{2.5} mass and the speciated PM_{2.5} (OPM_{2.5} = PM_{2.5} – SO₄ – NO₃ – NH₄ – EC – OA). Since total PM_{2.5} and speciated PM_{2.5} were measured using different techniques, the observed and modeled OPM_{2.5} species can represent different compounds and measurement artifacts can lead to negative OPM_{2.5} values, which are set to zero.

There are eight FRM PM_{2.5} monitoring sites within Allegheny County: Avalon, Lawrenceville, Liberty, South Fayette, North Park, Harrison, North Braddock, and Clairton (see Figure 2-1). Speciated PM_{2.5} CSN monitoring data is available from 2 sites: Liberty and Lawrenceville. Hourly AQS monitoring data for TEOM PM_{2.5} are also available at Liberty and Lawrenceville.

4.2 Model Performance Statistics and Performance Goals and Criteria

The Atmospheric Model Evaluation Tool (AMET⁹) is the primary software tool used to compare observations and modeled values from the 1.33-km domain in Allegheny County. Table 4-2 lists the statistical metrics that are calculated using AMET. AMET can also generate the whole suite of model performance metrics and

⁽⁹⁾ <https://www.cmascenter.org/amet>

graphical displays of model performance. These graphical displays include scatter plots, soccer plots, and time series. Together with the statistical metrics listed in Table 4-2, the graphical procedures are intended to help determine whether the CAMx base case simulation is performing well enough to make reliable future year PM_{2.5} projections.

Quantile-Quantile (Q-Q) plots are also used to visualize the comparison between modeled and measured values. Q-Q plots show a scatter plot of ranked pairings of predicted and observed concentration, where any rank of the predicted concentration is plotted against the same ranking of the observed concentration. Q-Q plots are developed to evaluate a model’s ability to represent the frequency distribution of the observed 24-hour PM_{2.5} concentrations. Q-Q plots are useful for investigating whether predictions are biased high or low with respect to observed concentrations at the upper end of the frequency distribution.

AMET is also used to evaluate the ability of CAMx to characterize the PM_{2.5} from local sources within Allegheny County. As described in the Modeling Protocol, if CAMx is not able to reproduce PM_{2.5} contributions from these sources, the AERMOD dispersion model may be used to simulate the contributions of primary PM emissions from local sources and results from both models can be combined during post-processing.

Table 4-2. Core statistical measures to be used in the Allegheny County air quality model evaluation with ground-level data.

Statistical Measure	Mathematical Expression	Notes
Accuracy of paired peak (A_p)	$\frac{P - O_{peak}}{O_{peak}}$	P _{peak} = paired (in both time and space) peak prediction
Coefficient of determination (R²)	$\frac{\left[\sum_{i=1}^N (P_i - \bar{P})(O_i - \bar{O}) \right]^2}{\sum_{i=1}^N (P_i - \bar{P})^2 \sum_{i=1}^N (O_i - \bar{O})^2}$	P _i = prediction at time and location i; O _i = observation at time and location i; \bar{P} = arithmetic average of P _i , i=1,2,...,N; \bar{O} = arithmetic average of O _i , i=1,2,...,N
- table continued on next page -		

Statistical Measure	Mathematical Expression	Notes
Normalized Mean Error (NME)	$\frac{\sum_{i=1}^N P_i - O_i }{\sum_{i=1}^N O_i}$	Reported as %
Root Mean Square Error (RMSE)	$\left[\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2 \right]^{1/2}$	Reported as %
Fractional Gross Error (FE)	$\frac{2}{N} \sum_{i=1}^N \left \frac{P_i - O_i}{P_i + O_i} \right $	Reported as %
Mean Absolute Gross Error (MAGE)	$\frac{1}{N} \sum_{i=1}^N P_i - O_i $	Reported as concentration (e.g., $\mu\text{g}/\text{m}^3$)
Mean Normalized Gross Error (MNGE)	$\frac{1}{N} \sum_{i=1}^N \frac{ P_i - O_i }{O_i}$	Reported as %
Mean Bias (MB)	$\frac{1}{N} \sum_{i=1}^N (P_i - O_i)$	Reported as concentration (e.g., $\mu\text{g}/\text{m}^3$)
Mean Normalized Bias (MNB)	$\frac{1}{N} \sum_{i=1}^N \frac{(P_i - O_i)}{O_i}$	Reported as %
Mean Fractionalized Bias (Fractional Bias, FB)	$\frac{2}{N} \sum_{i=1}^N \left(\frac{P_i - O_i}{P_i + O_i} \right)$	Reported as %
Normalized Mean Bias (NMB)	$\frac{\sum_{i=1}^N (P_i - O_i)}{\sum_{i=1}^N O_i}$	Reported as %
Bias Factor (BF)	$\frac{1}{N} \sum_{i=1}^N \left(\frac{P_i}{O_i} \right)$	Reported as BF:1 or 1:BF or in fractional notation (BF/1 or 1/BF).

Model performance goals and criteria for PM species have been developed as part of the regional haze modeling performed by several Regional Planning Organizations (RPOs) and are provided in Table 4-3. The PM model performance goals focus on Fractional Bias (FB) and Fractional Error (FE) and no observed concentration threshold screenings are applied. Since the FB and FE are divided by the average of

the predicted and observed values, FB is bounded by -200% and +200% and the FE is bounded by 0% to +200%. Note that Emery and Co-workers (2017) have recently published updated recommended model performance goals and criteria for ozone and PM. However, they were published after the MPE was conducted for the Allegheny County CAMx 2011 base case simulation so are not used in this report. For total PM_{2.5} mass and SO₄ and NH₄ components of PM, the Emery et al. (2017) Performance Goals (FB,FE) are more stringent than the RPO (±10%, 35%) and the Performance Criteria is the same as the RPO Performance Goals (±30%, 75%).

Table 4-3. PM Model Performance Goals and Criteria

Fractional Bias (FB)	Fractional Error (FE)	Comment
≤±30%	≤50%	PM model performance Goal, considered good PM performance.
≤±60%	≤75%	PM model performance Criteria, considered average PM performance. Exceeding this level of performance for PM species with significant mass may be cause for concern.

The model performance evaluation using AMET presented in this study focuses on the CAMx model performance within the 1.33 km domain, which is the grid resolution that is most comparable to observations in Allegheny County. For the 36 and 12 km domains, pseudonetcdf and NCL tools were used to generate plots of quarterly average PM_{2.5} concentrations. These spatial plots were qualitatively inspected to ensure that modeled PM_{2.5} values at these grid resolutions are in-line with expected values and that no outliers are present. These plots are presented in the following section.¹⁰

4.3 Model Performance Evaluation

The CAMx 2011 base year modeling results were compared against measured ambient concentrations within Allegheny County. Details of the CAMx 2011 base case model performance evaluation are presented below.

¹⁰ Although this MPE focuses on PM_{2.5}, it can also be useful to analyze modeled ozone concentrations to provide an additional level of quality assurance and to assure that the model is adequately reproducing the reactivity of the atmosphere. Maximum daily 8-hour average ozone concentrations were found to be reasonable for the domains, showing a modeled range of 40-175 ppb, with the highest concentrations in spring and summer and in areas of high anthropogenic activity.

To obtain results presented in Section 4.3.2 through Section 4.3.4, PM_{2.5} concentrations in live PiG puffs are combined with regional/background PM_{2.5} and PM_{2.5} in dumped PiG puffs using bilinear interpolation techniques. During interpolation, model results in each 1.33 km grid cell are interpolated to the resolution of the PiG sampling grid (~100 m spacing).

For calculating design values, PM_{2.5} concentrations from live PiG puffs and AERMOD model results are combined with regional/background PM_{2.5} using the average of all 100 m receptor live puff concentrations within a 1.33 km grid cell in place of bilinear interpolation of the CAMx 1.33 km grid cell average concentration to the 100 m receptors used in the MPE. During averaging, all PiG or AERMOD receptors within each 1.33 km CAMx grid cell are averaged together and added to the PM_{2.5} concentrations in the 1.33 km grid cell. This ensures that the spatial dimensions and resolution of the 1.33 km CAMx grid cell domain are used to calculate PM_{2.5} design values instead of the very fine (~100 m spacing) PiG and AERMOD sampling grid domains. Additionally, PM_{2.5} design values are calculated by using average values within a grid cell array surrounding the monitor site, so averaging PiG and AERMOD values in each 1.33 km grid cell is more appropriate than using bilinear interpolation. Using averaging techniques to incorporate the live PiG puff and AERMOD components during calculation of PM_{2.5} design values is consistent with techniques used in previous Allegheny County SIP analyses.

Section 4.3.5 presents a comparison between observations and base year model results using both PiG (CAMx+PiG) and AERMOD (CAMx+AERMOD) techniques which are used to represent the primary emissions associated with local major sources in Allegheny County. Design value calculation is not described in this report and more details can be found in the accompanying Air Quality Technical Support document (Ramboll Environ, 2018).

4.3.1 Spatial Plots of Average PM_{2.5}

Figure 4-1 and 4-2 show the spatial distribution of modeled surface-level, quarterly average PM_{2.5} for the 2011 base case 36 and 12 km domains. For all quarters, average PM_{2.5} values are higher in the eastern U.S. compared to the west, with quarterly average maximum values of 23, 15, 19 and 22 µg/m³. Agricultural activity combined with industry and traffic are major contributors to PM_{2.5} concentrations in the central and eastern US. Figure 4-2 shows that the highest surface-level PM_{2.5} concentrations in the eastern US occur in big cities during the

cooler months (quarter 1 and 4). During these periods, PM_{2.5} emissions associated with home heating is significant and shallower boundary layers tend to concentrate PM_{2.5} closer to the surface. Overall, the spatial patterns and magnitudes of PM_{2.5} shown in Figure 4-1 and Figure 4-2 are qualitatively in line with expectations given the location of major PM_{2.5} sources across the United States. The following section (4.3.2) will compare modeled and observed PM_{2.5} in a quantitative framework.

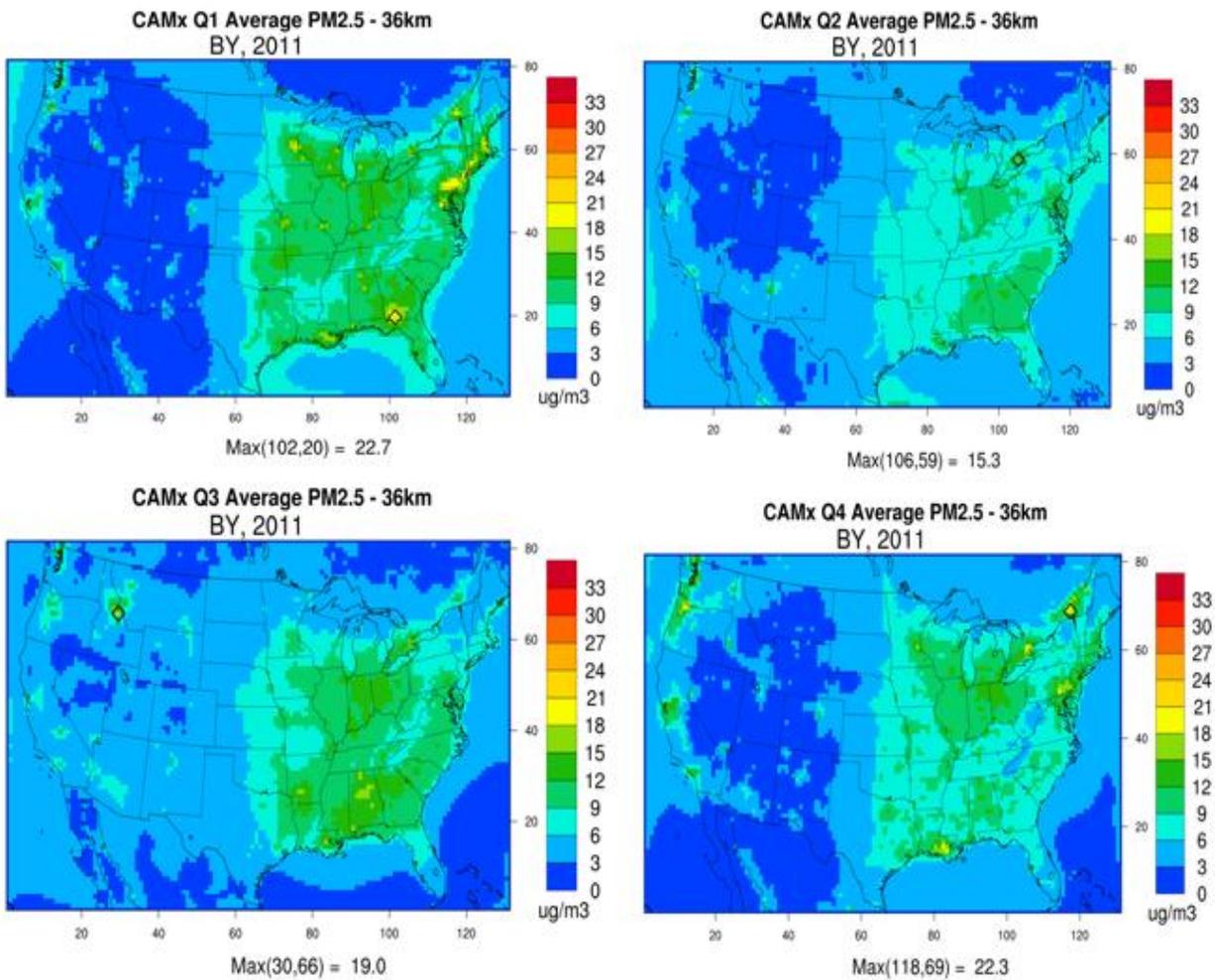


Figure 4-1. Quarterly Average PM_{2.5} for 36 km continental U.S. domain

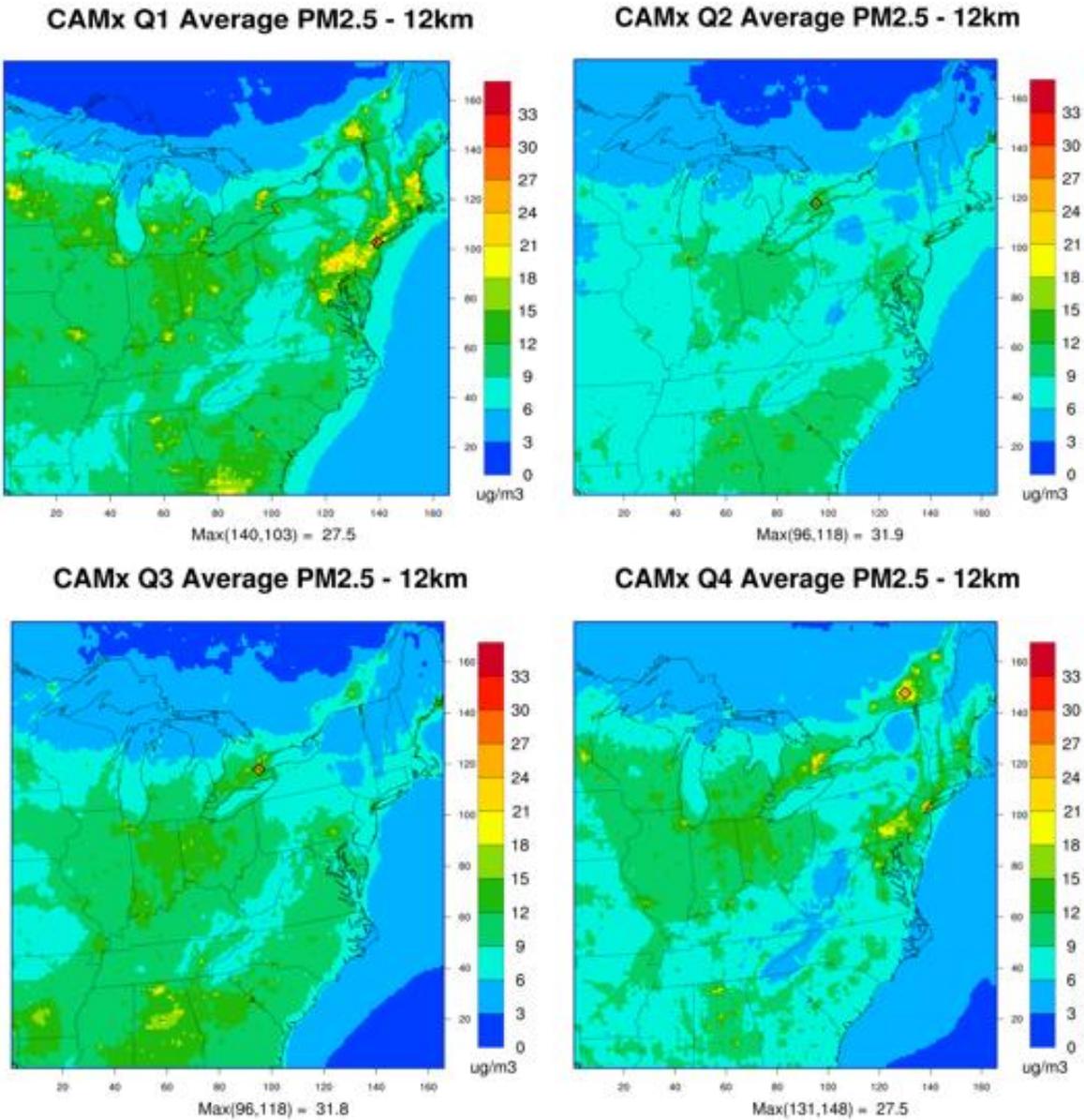


Figure 4-2. Quarterly Average PM_{2.5} for 12 km domain

4.3.2 Total PM_{2.5} Mass

Table 4-4 shows FRM and TEOM total PM_{2.5} Fractional Bias (FB) and Fractional Error (FE) performance statistics for the 1.33 km domain. Results from each monitoring site within a network are averaged together to calculate mean FB and FE values for each quarter. Statistics for each quarter are presented along with the average FB and FE values across all quarters. For both monitoring networks, all FB and FE

values achieve the performance criteria by a wide margin and half of the quarters exceed the PM performance goals. On an annual basis, when the FB and FE values for each quarter are averaged together, the PM performance goals are achieved in all cases.

Table 4-4. Comparison of Quarterly FRM, TEOM, and CSN PM_{2.5} Fractional Bias and Error Performance Statistics with PM Performance Goals and Criteria for 1.33 km Domain

Network	Number Sites	Quarter	Fractional Bias Criteria: $\leq \pm 60\%$ Goal: $\leq \pm 30\%$	Fractional Error Criteria: $\leq 75\%$ Goal: $\leq 50\%$
TEOM	2	1	+35.1%	50.5%
		2	-14.4%	44.3%
		3	+1.6%	46.4%
		4	+33.6	53.6%
		Average	+14.0%	48.7%
FRM	8	1	+45.8%	48.0%
		2	+6.8	35.1%
		3	-2.5	36.8%
		4	+36.4	45.8%
		Average	+21.6%	41.4%

Figure 4-3 presents soccer plots showing quarterly FB and FE when CAMx is compared with daily FRM PM_{2.5} observations and Figure 4-4 presents FB and FE when TEOM observations are used. Soccer plots are scatter plots of FB (X-axis) vs. FE (Y-axis) that include boxes for the PM_{2.5} performance goals and criteria so one can easily interpret when the goals and criteria are achieved. The most inner box in the soccer plot is the ozone (bias, error) performance goal, which is more stringent than those for PM_{2.5} ($\pm 15\%$, 35%). Figure 4-3 and Figure 4-4 show that all FB and FE values fall within the performance criteria and that most sites are within the performance goal for quarter 2 and quarter 3. There is some variation in performance between the monitoring sites in each network, but for most quarters, the FB and FE values are clustered in a similar region of the soccer plot.

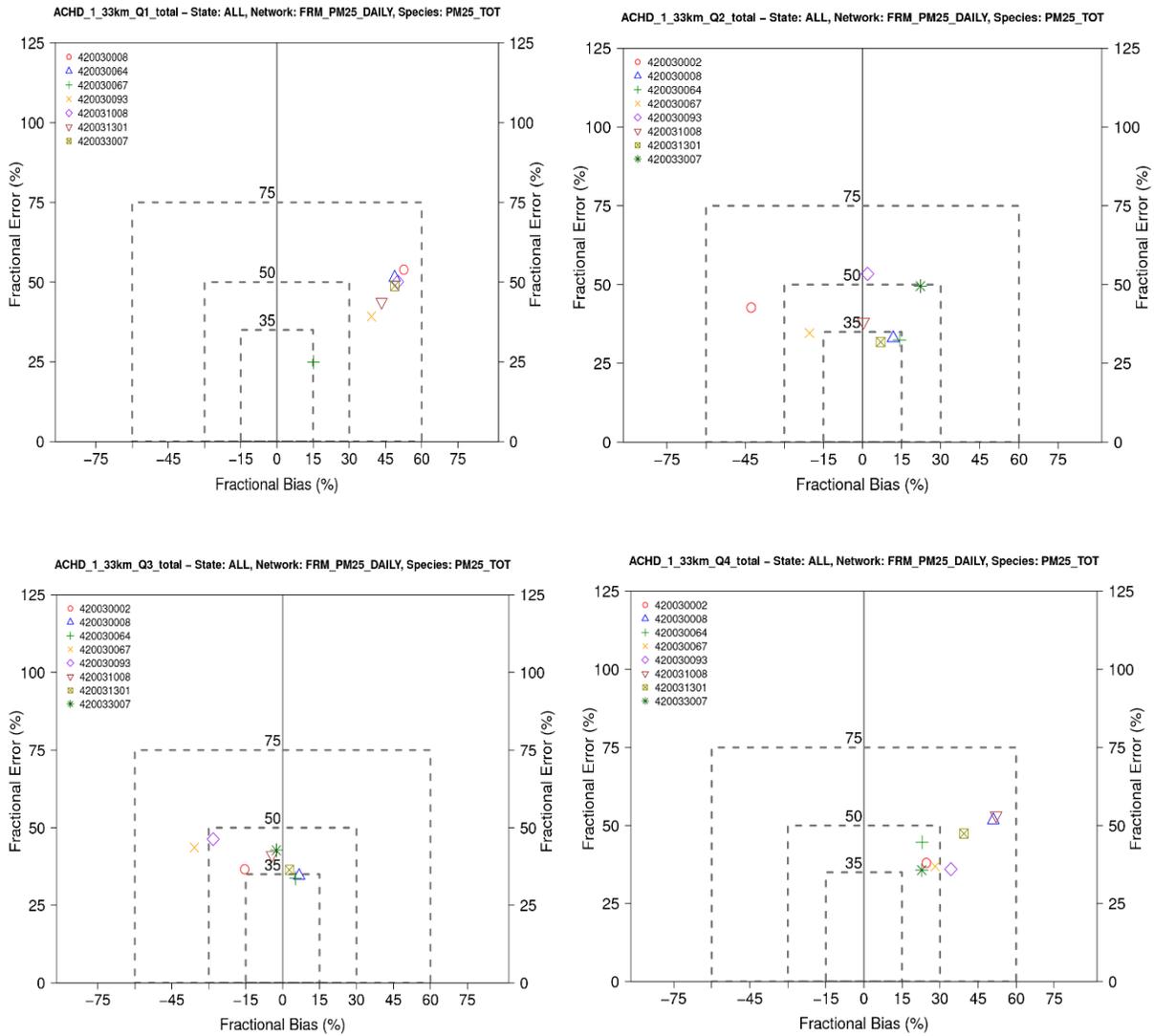


Figure 4-3. Quarterly PM_{2.5} Fractional Bias and Error Soccer Plots Using Daily FRM observations at 8 Monitoring Sites in Allegheny County

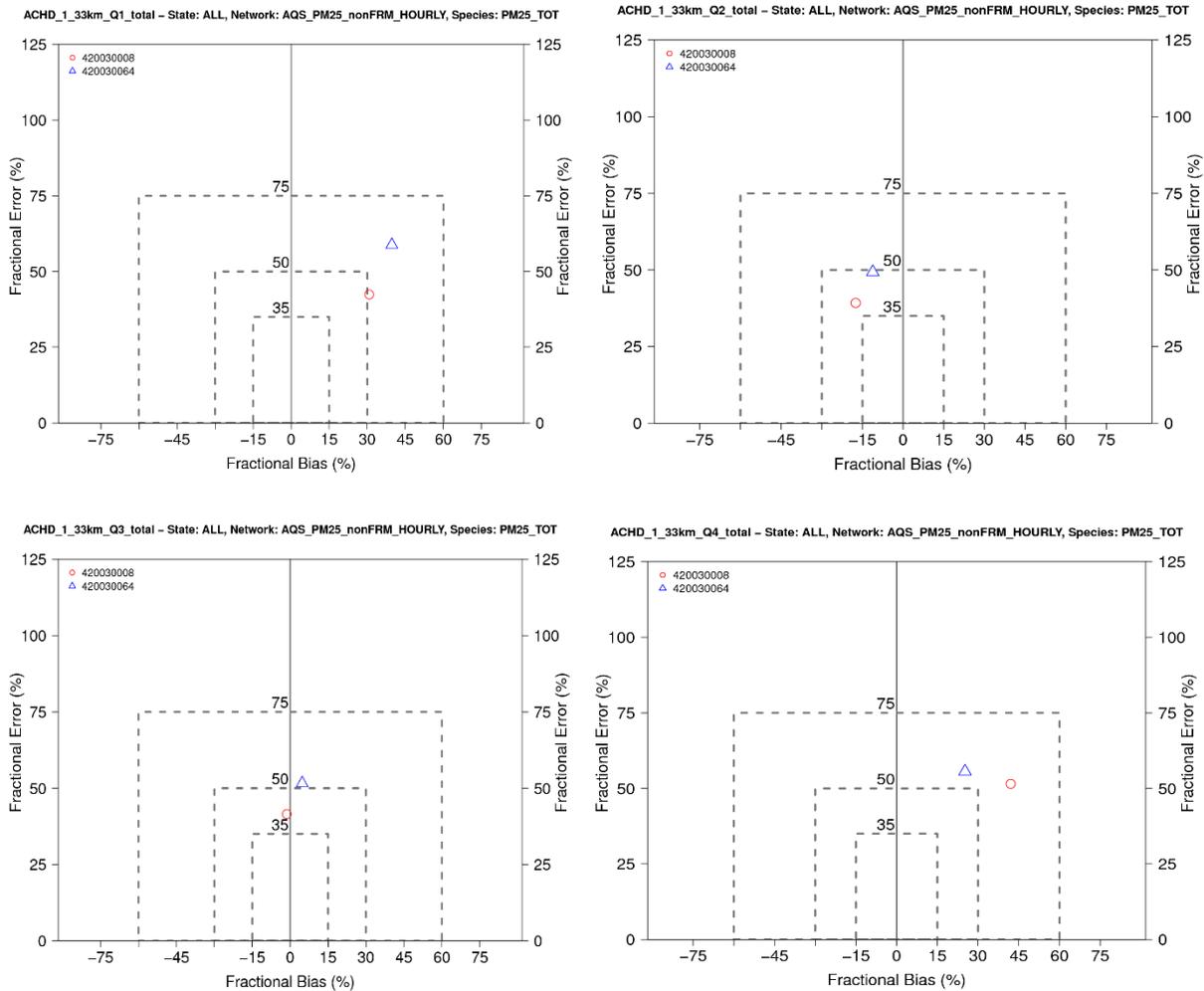


Figure 4-4. Quarterly PM_{2.5} Fractional Bias and Error Soccer Plots Using Hourly TEOM observations at Liberty and Lawrenceville.

Figure 4-5 shows time series of FRM observations compared to CAMx-predicted values at Liberty for each quarter. Observations are presented in black, total (local plus regional) PM_{2.5} is shown in red, and local PM_{2.5} is shown in blue. The CAMx “total” values represent the sum of regional and local daily-average PM_{2.5} concentrations in the CAMx grid cell containing the Liberty monitoring site. In CAMx, local sources were tracked separately using source apportionment techniques and major local sources have also been flagged for PiG treatment in CAMx. Figure 4-5 shows that there is considerable variability in both modeled and observed PM_{2.5} values over the year with no distinct trends. Surface-level PM_{2.5} concentrations at Liberty are influenced by natural and anthropogenic emissions of

PM_{2.5} as well as weather patterns, which influence the amount of vertical mixing in the planetary boundary layer. These weather events include inversions that trap local source pollutants near the surface resulting in high PM concentrations at the Liberty monitoring site. Observed and modeled “total” PM_{2.5} concentrations range from 5 to 55 µg/m³ and PM_{2.5} associated with local sources range from 0 to 30 µg/m³. CAMx does a reasonable job reproducing daily PM_{2.5} observations, especially during the warmer periods (quarter 2 and quarter 3). The model-measurement agreement differs the most in quarter 1, which may partially be due to the challenges involved with capturing sharp inversions in the complex terrain surrounding the Liberty site.

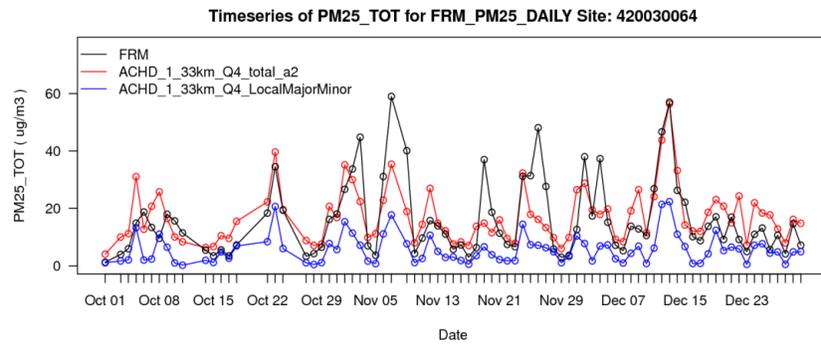
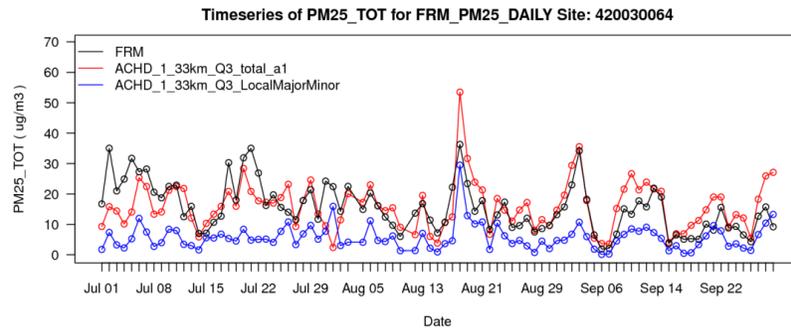
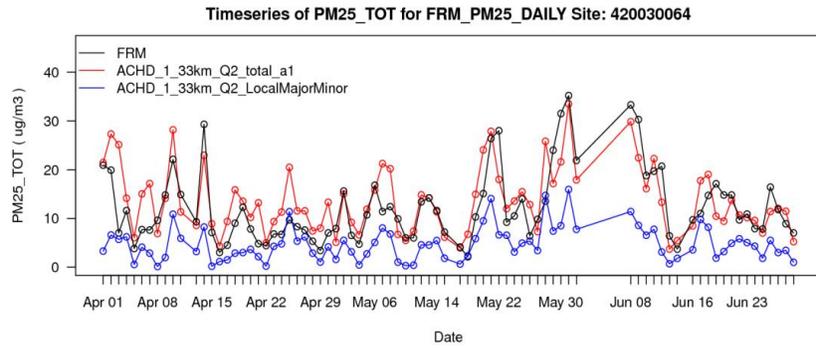
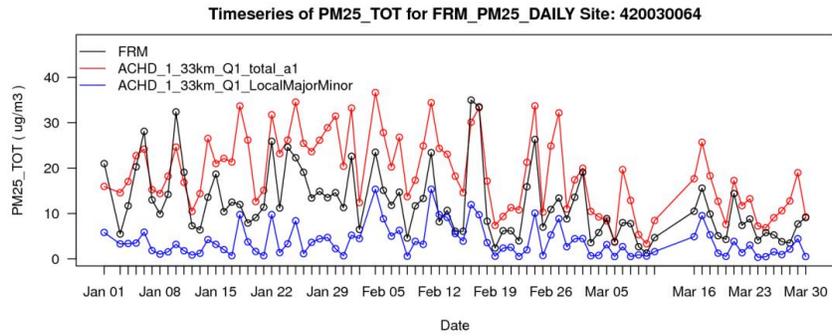


Figure 4-5. Quarterly Time Series of Daily FRM PM_{2.5} at Liberty.

Figure 4-6 is a quantile-quantile (Q-Q) plot that compares daily observed (FRM) and modeled PM_{2.5} concentrations at Liberty for 2011. In the Q-Q plot, daily observed and modeled concentrations are each sorted from highest to lowest and then paired together such that the highest observed and modeled values are directly compared. Exact matches between modeled and observed values are located along the black line that stretches diagonally across the center of the plot. All points that fall between the black centerline and the two dashed lines (red) indicate that modeled and measured values agree within a factor of two. Annual mean observed and modeled values, along with other relevant statistics such as robust highest concentration (RHC), are also shown on the Q-Q plot. Figure 4-6 shows that there is good agreement (within a factor of two) between modeled and measured PM_{2.5} concentrations at Liberty across the vast majority of the frequency distribution. The agreement is best at the upper end of the range and the model slightly overestimates concentrations at the middle and lower end.

Both the observed and modeled annual PM_{2.5} values at Liberty exceed the NAAQS (12 µg/m³), which are 14.1 µg/m³ and 16.4 µg/m³, respectively. Note that days with missing FRM observations were excluded during the calculation of modeled and observed annual mean PM_{2.5} values (335 days were included). The annual mean estimated by CAMx is higher than the observed annual mean, suggesting that CAMx is maintaining a degree of conservatism at Liberty. The Q-Q plot shows that CAMx using the PiG does a very good job in replicating the observed 24-hour PM_{2.5} concentrations above 20 µg/m³ at Liberty that include the occurrences of inversion resulting in high contributions from local sources. The CAMx overestimation of the observed annual PM_{2.5} at Liberty is due to an overestimation of observed lower PM_{2.5} concentrations that is likely due to too high modeled regional PM concentrations. In other studies using the 2011 modeling year, such overestimation of the lower observed PM_{2.5} concentrations have been attributed to overstated GEOS-Chem PM boundary conditions and/or too high PM emissions from residential wood combustion (wood stoves and fireplaces) in and near urban areas.

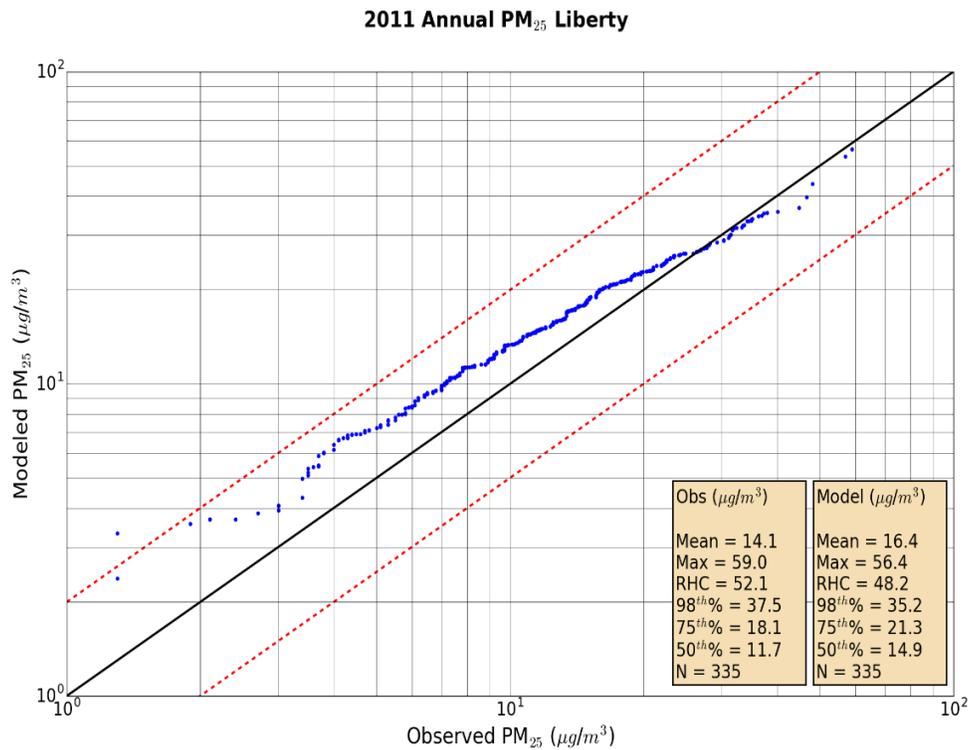


Figure 4-6. Annual Q-Q Plot of FRM and Modeled PM_{2.5} at Liberty

To further explore CAMx model performance within Allegheny County, it is useful to compare modeled and measured PM_{2.5} values at other monitoring sites within the county. Figure 4-7 shows Q-Q plots that compare daily observed (FRM) and modeled PM_{2.5} concentrations at Clairton, Lawrenceville, North Braddock, and South Fayette. Similar to the Liberty Q-Q plot (Figure 4-6), nearly all modeled concentrations are within a factor of two compared to observations at all four monitoring sites. At Clairton, Lawrenceville, and North Braddock, CAMx generally slightly overestimates PM_{2.5} concentrations across much of the frequency domain. At South Fayette, CAMx underestimates the highest concentrations, overestimates the lowest concentrations, and exhibits excellent agreement in the middle of the range. Observed annual mean PM_{2.5} values are generally below the NAAQS while model-predicted values exceed the NAAQS at all sites except South Fayette. Figure 4-7 suggests that CAMx is able to simulate PM_{2.5} concentrations across Allegheny County reasonable well. Overall, CAMx tends to overestimate PM_{2.5} concentrations, suggesting that the PM_{2.5} estimates predicted by CAMx during the future year simulation will likely be conservative. The overestimation of the annual PM_{2.5} at

these other sites confirms that the overestimation of the annual $PM_{2.5}$ at Liberty is due to an overestimation of the regional component of $PM_{2.5}$.

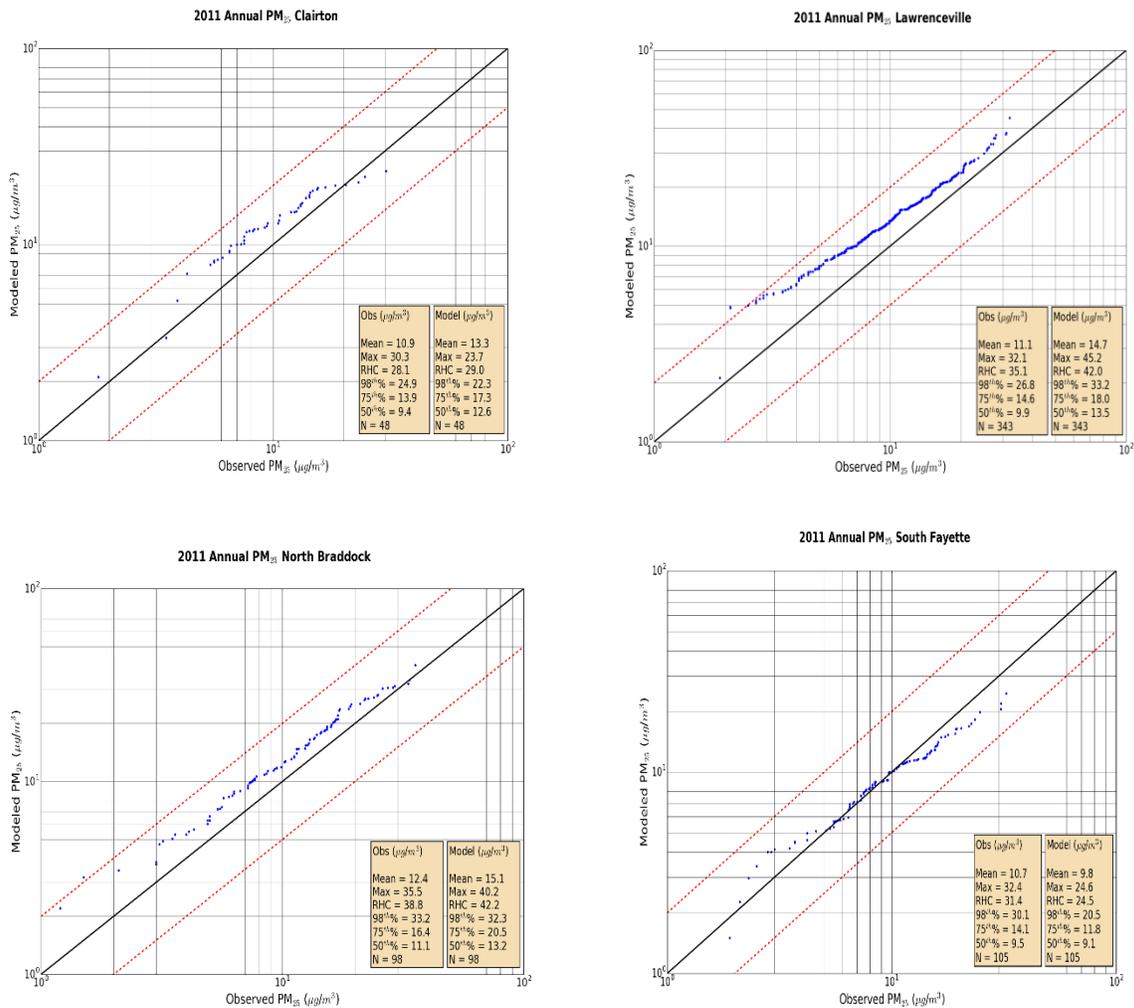


Figure 4-7. Annual Q-Q Plot of FRM and Modeled $PM_{2.5}$ at Clairton, Lawrenceville, North Braddock, and South Fayette.

The Q-Q plots shown in Figure 4-6 and Figure 4-7 clearly indicate whether CAMx is biased high or low compared to observations across the entire range of concentrations, although temporal information is lost during sorting. Scatterplots retain time resolution and provide additional insight into model-measurement agreement.

Figure 4-8 presents scatterplots that compare daily observed and modeled $PM_{2.5}$ concentrations at all 8 FRM monitoring sites within Allegheny County for each quarter. Similar to the Q-Q plot, points along the diagonal black centerline indicate perfect agreement between modeled and observed concentrations, and points located between the two black lines flanking the centerline indicate factor of two agreement between modeled and observed values. During the warmer months (quarter 2 and quarter 3), modeled $PM_{2.5}$ concentrations are within a factor of two compared to observations for the majority of days at all monitoring sites. In quarter 1 and quarter 4, CAMx overestimates $PM_{2.5}$ concentrations during some periods, likely due to overestimation of the regional component of $PM_{2.5}$ discussed previously, although some overestimation of the local component may also be present. Despite the tendency for CAMx to overestimate concentrations during cooler months, CAMx is able to reproduce $PM_{2.5}$ concentrations within a factor of two for the majority of the year at all monitoring sites within Allegheny County. Figure 4-9 shows scatterplots comparing hourly observed and modeled $PM_{2.5}$ at AQS monitoring sites within Allegheny County. Similar to Figure 4-8, the best agreement is seen during the warmer months and CAMx tends toward overestimating during cooler months.

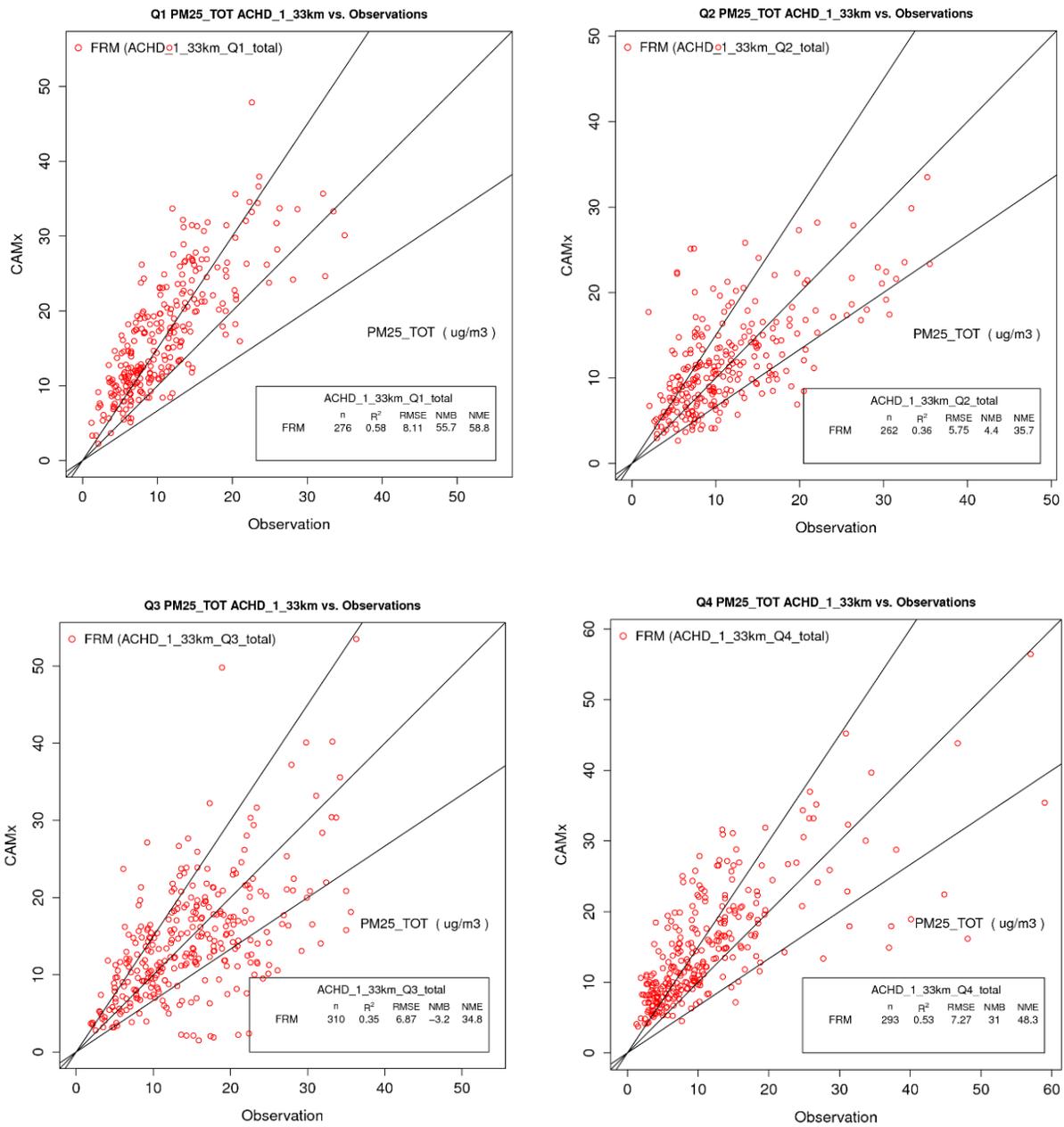


Figure 4-8. Quarterly Scatterplots of Modeled PM_{2.5} and FRM Observations from 8 Monitoring Sites in Allegheny County.

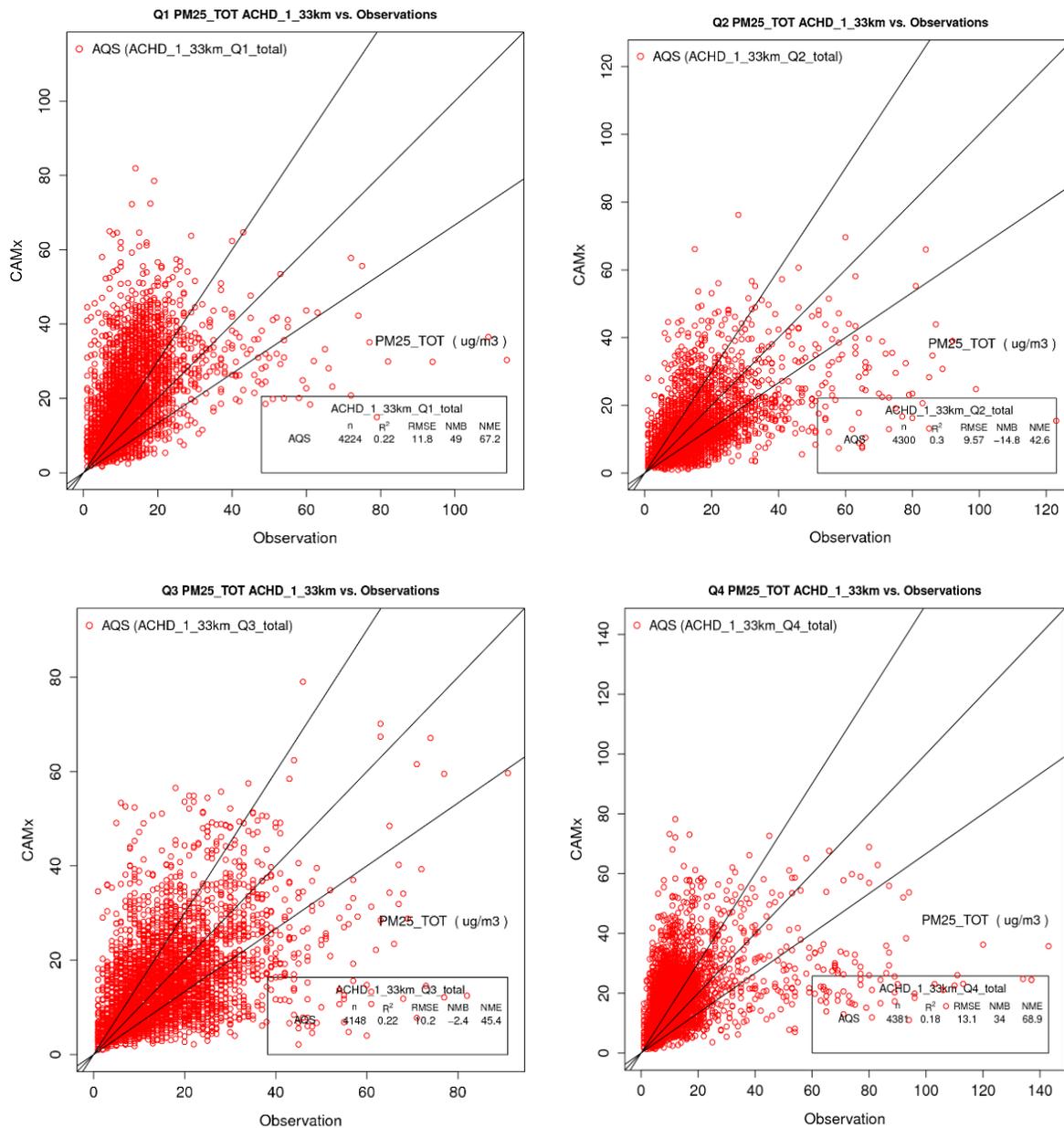


Figure 4-9. Quarterly Scatterplots of Modeled PM_{2.5} and TEOM Observations at Liberty and Lawrenceville.

4.3.3 Local Source Contributions to Annual Average PM_{2.5}

Figure 4-5 shows that local major sources are important contributors to total PM_{2.5} in Allegheny County. Table 4-5 shows the local and regional contributions to annual average PM_{2.5} contributions at several sites within Allegheny County.

Table 4-5. Local and Regional Contributions to Annual Average PM_{2.5}

	PM_{2.5} (µg/m³)			
Site	Observed (FRM)	Modeled		
	Total	Total	Regional	Local
Clairton	10.9	14.1	11.3	2.7
Lawrenceville	11.1	14.5	13.0	1.5
Liberty	14.1	16.2	11.3	4.9
North Braddock	12.4	15.2	10.9	4.3
South Fayette	10.7	9.5	9.2	0.3

At Liberty, model results suggest that local sources contribute 30% to overall annual average PM_{2.5}. There is a wide range of local source contribution at these five monitoring sites. The contribution of local sources to annual average PM_{2.5} is 28% at North Braddock, 19% at Clairton, 10% at Lawrenceville, and 3% at South Fayette. Note that in Table 4-5, all days in 2011 were used to calculate modeled PM_{2.5} values; days with missing FRM observations were not excluded. As a result, there are slight differences in the modeled Liberty annual average total PM_{2.5} values in Table 4-5 and Figure 4-8 (16.2 µg/m³ vs. 16.4 µg/m³).

Figure 4-10 presents pie charts illustrating the chemical speciation associated with the local component of annual average PM_{2.5} at Liberty. The pie chart on the left was generated using 2011 CSN observations within the Pittsburgh metropolitan statistical area (MSA). Speciated CSN measurements across the MSA were used to obtain mean background concentrations, and the local (or "excess") PM_{2.5} contribution at Liberty was determined by taking the difference between background concentrations and CSN measurements at Liberty. The observed annual average PM_{2.5} excess at Liberty is 3.8 µg/m³. The pie chart on the left shows that annual average excess PM_{2.5} at Liberty is dominated by SO₄, EC, and OA, followed by NH₄ and other (e.g., soil, fine particulates).

The pie chart on the right shows the modeled chemical speciation associated with the local source component at Liberty. The modeled local source contribution at Liberty is 4.9 µg/m³, which is slightly higher compared to the observed excess,

although these two methods vary drastically in their approach. The modeled local source component at Liberty is dominated by species such as soil and fine particulates, followed by NH_4 and SO_4 . Both the observed and modeled local component at Liberty show that NO_3 is a minor component of $\text{PM}_{2.5}$ compared to fine particulates, soil, NH_4 , EC, OA, and SO_4 . The differences in speciated $\text{PM}_{2.5}$ distributions is at least partially related to differences in observed and modeled species mapping algorithms (i.e., which individual species are grouped together for each category). The following section will address speciated $\text{PM}_{2.5}$ model performance in more detail.

Although there are uncertainties in the estimated actual local source $\text{PM}_{2.5}$ contribution at Liberty obtained by taking differences in observed concentrations at the monitoring sites, Table 4-5 and Figure 4-10 suggest that some of the CAMx overestimation of the annual observed $\text{PM}_{2.5}$ concentration at Liberty (16.2 vs. $14.1 \mu\text{g}/\text{m}^3$) is due to overestimating the local source contribution (4.9 vs. $3.8 \mu\text{g}/\text{m}^3$). However, when the estimated actual local source contribution is added to the CAMx regional contribution, the total observed $\text{PM}_{2.5}$ at Liberty is still overestimated (15.1 vs. $14.1 \mu\text{g}/\text{m}^3$) indicating that the regional $\text{PM}_{2.5}$ component at Liberty is also slightly overestimated by CAMx.

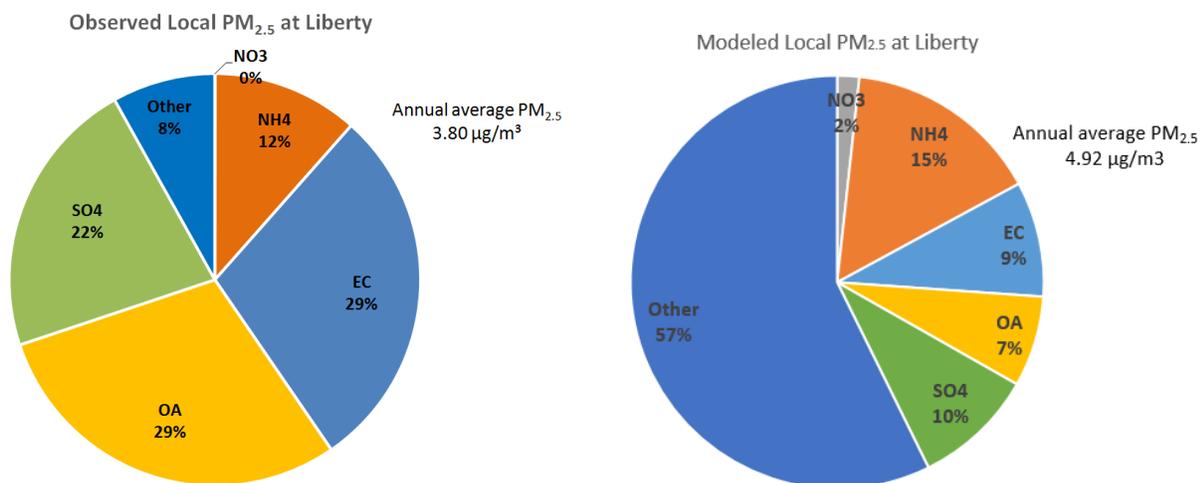


Figure 4-10. Speciated “Excess” $\text{PM}_{2.5}$ at Liberty, Observed and Modeled, 2011.

4.3.4 Overview of Speciated PM_{2.5} Model Performance

This section presents speciated PM_{2.5} soccer plots for each quarter. In these plots, modeled outputs are compared to CSN observations. Figure 4-11 presents soccer plots displaying quarterly FB and FE when CAMx is compared to CSN SO₄ observations at Liberty and Lawrenceville monitoring sites. Fractional bias and error for SO₄ fall within the performance criteria for all quarters. In quarter 1, FB and FE meet the performance goal at Lawrenceville. In quarter 2, the performance goals are met for both sites. In quarter 3, the goal is met for Lawrenceville and nearly met for Liberty. In quarter 4, both sites are just outside of the goal but well inside the performance criteria bounds. In general, model-measurement agreement at Lawrenceville is slightly better for SO₄ compared to Liberty, but overall agreement is good at both sites. Similar to the comparisons between modeled and observed total PM_{2.5}, CAMx performs best during the warmer months (quarter 2 and quarter 3), when surface chemistry is least impacted by strong inversions.

Figure 4-12 shows quarterly NO₃ FB and FE soccer plots for Liberty and Lawrenceville. The FB and FE performance criteria are met in quarters 1 and 4. In quarters 2 and 3, there are notable differences between observed and estimated concentrations at both monitor sites. In these quarters, CAMx under predicts NO₃ concentrations, which is perhaps related to chemical partitioning differences. Chemical partitioning of these nitrogen species is dependent on parameters that can be difficult to model accurately on very small scales, such as atmospheric temperature, humidity and pH. As shown in Figure 4-10, NO₃ contributes the least to observed and modeled PM_{2.5} at Liberty so the disagreements highlighted in Figure 4-12 have a minor impact on overall PM_{2.5} in Allegheny County.

Figure 4-13 presents NH₄ FB and FE soccer plots for each quarter at Liberty and Lawrenceville. The performance criteria are met at both sites for all quarters and performance at Lawrenceville is better compared to Liberty for all quarters except quarter 4. Figure 4-14 shows that the FB and FB performance criteria are met for total carbon (TC) for all but quarter 1, which is just outside of the performance criteria bounds due to an overestimation tendency. This TC overestimation during the cooler months is consistent with the overestimation of residential wood combustion (RWC) emissions in and near urban areas seen in other studies. The spatial surrogate used to spatially allocate county-level RWC emissions (locations of households) fails to account for the fact that rural households are more likely to use wood burning for home heating than urban ones.

Figure 4-11 through Figure 4-14 show that CAMx does a reasonable job estimating SO₄, NH₄, and TC at the two CSN sites in Allegheny County. During quarters 1 and 4, there is good agreement between modeled and observed NO₃. The NO₃ disagreement in warmer quarters may be related to differences in chemical partitioning between the various nitrogen species, although this difference has a minor impact on total PM_{2.5} because NO₃ is a small component of overall PM_{2.5} in Allegheny County. Most of the particulate NO₃ in CAMx is in the form of ammonium nitrate (NH₄NO₃) that is highly volatile so evaporates into ammonia and nitric acid when warm.

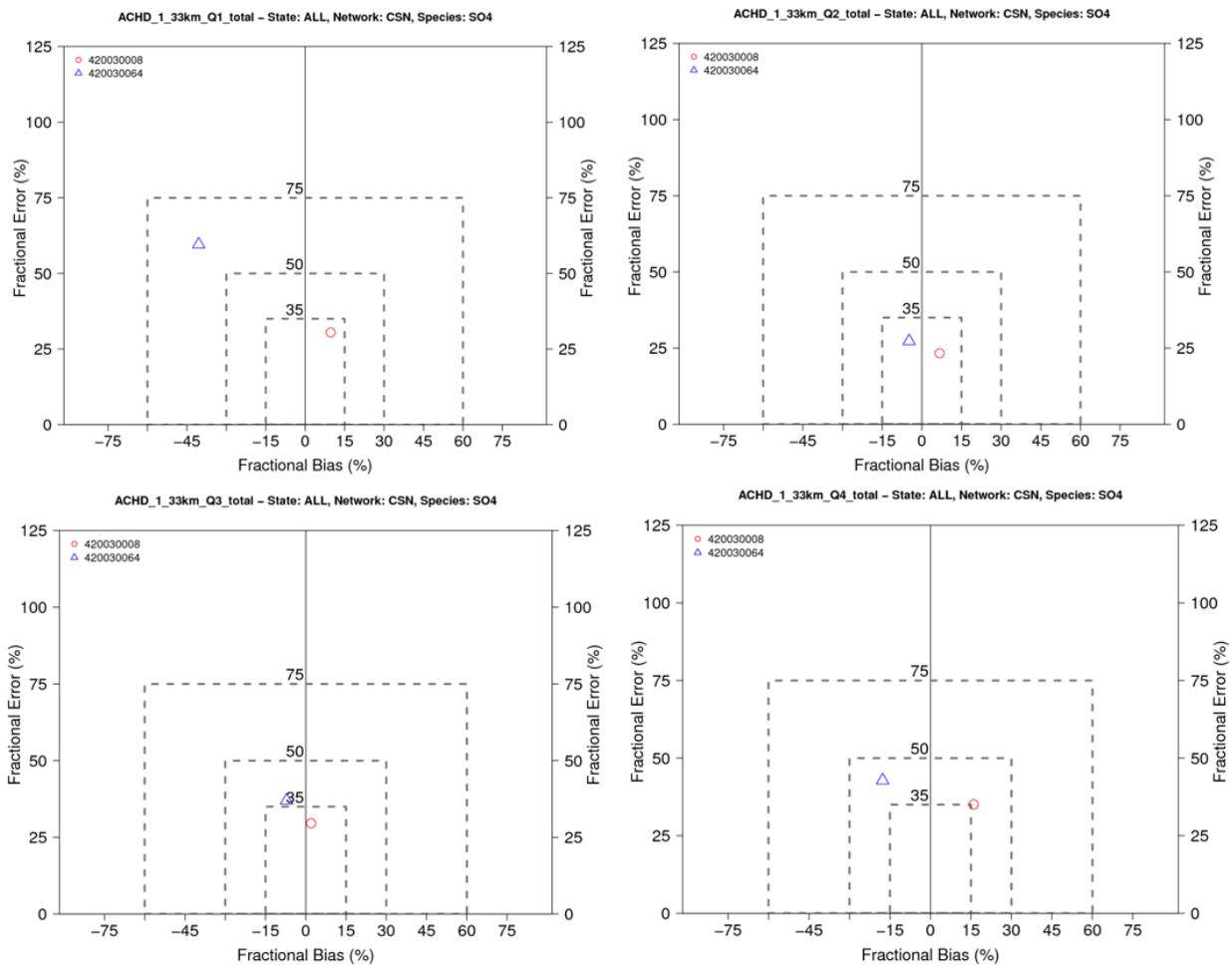


Figure 4-11. Quarterly SO₄ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.

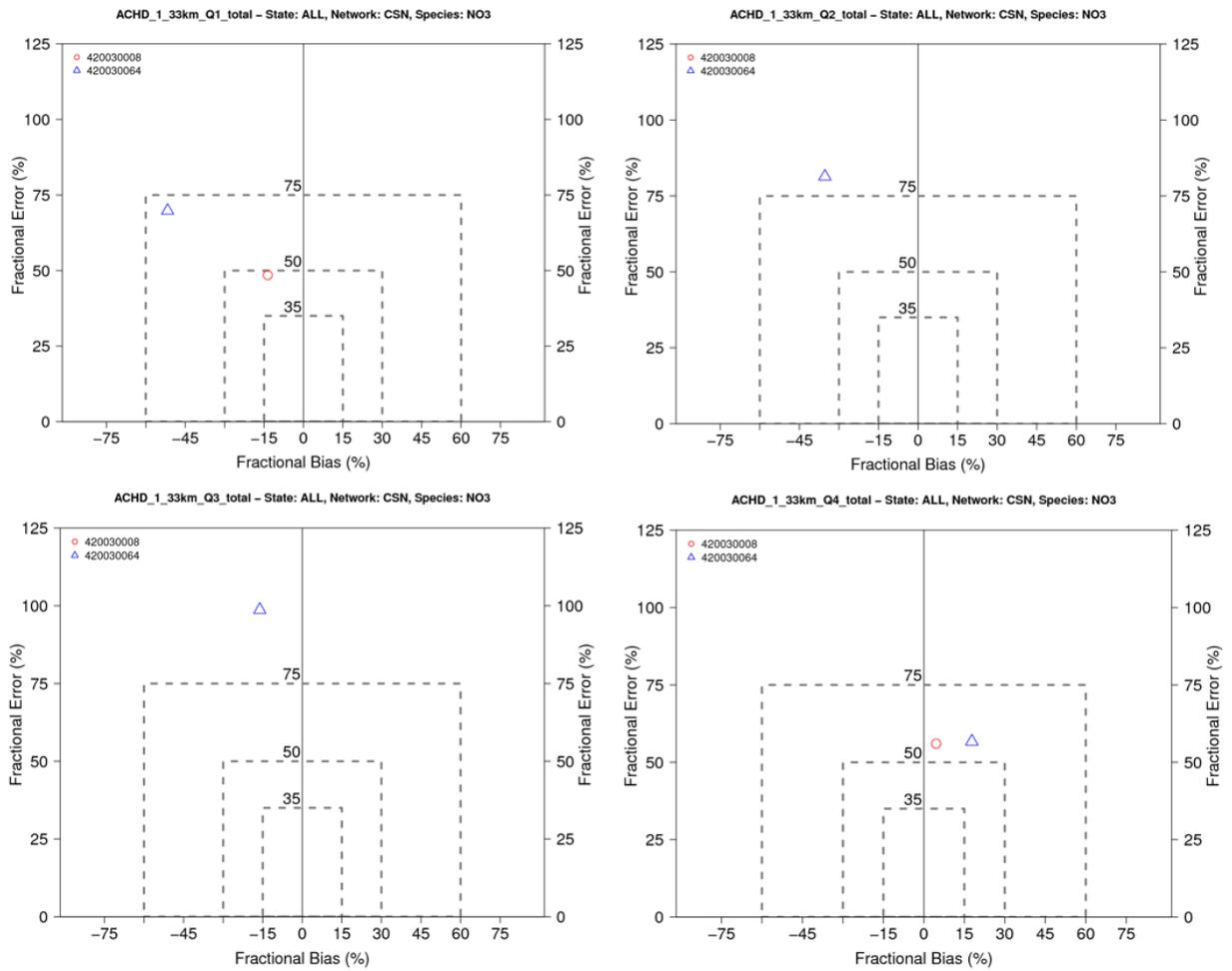


Figure 4-12. Quarterly NO₃ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.¹¹

¹¹ Statistics for Lawrenceville are not shown in the Q2 and Q3 soccer plots because the fractional bias and/or fractional error exceed the axes. For Q2, the fractional bias is -112.1 and the fractional error is 119.2. For Q3, the fractional bias is -143.2 and the fractional error is 154.4.

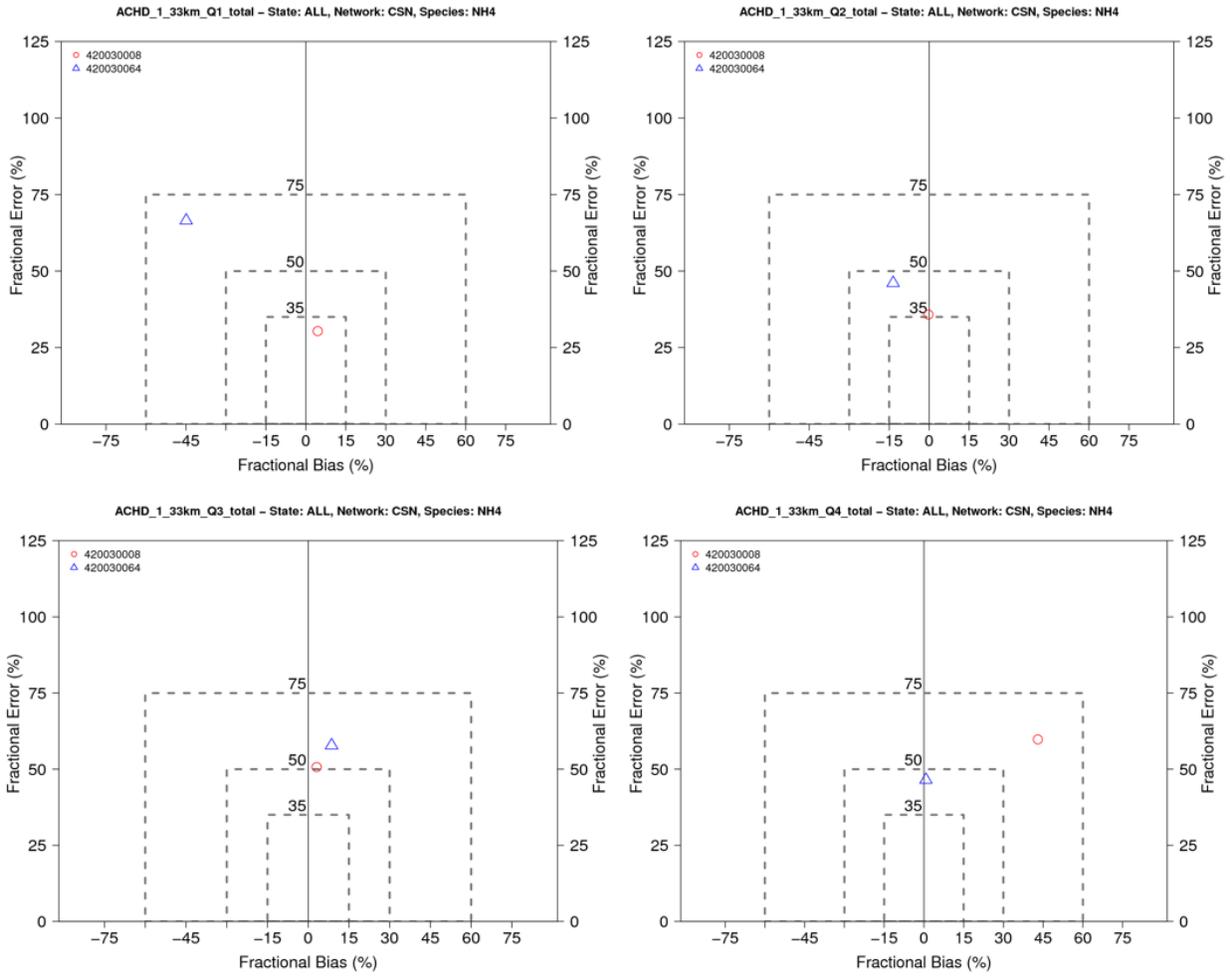


Figure 4-13. Quarterly NH₄ Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.

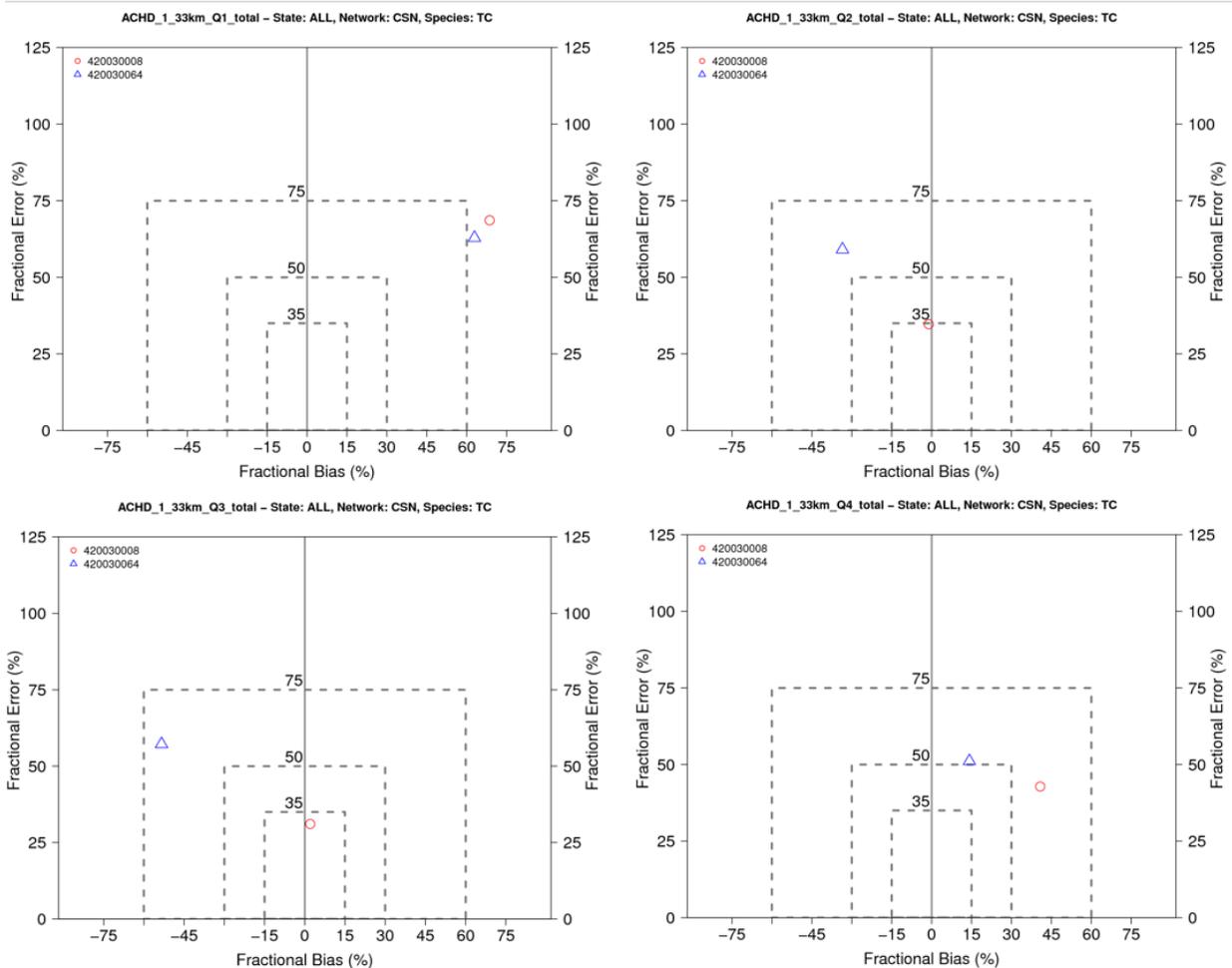


Figure 4-14. Quarterly Total Carbon Fractional Bias and Error Soccer Plots at Liberty and Lawrenceville.

4.3.5 Plume-in-Grid and AERMOD Comparison

Figure 4-15 compares quantile-quantile (Q-Q) cumulative frequency distributions plots of predicted and observed (FRM) 24-hour $PM_{2.5}$ concentrations at Liberty for 2011. In Figure 4-15, FRM observations are compared to the following set of model outputs: CAMx without contributions from primary $PM_{2.5}$ emissions emitted from local major sources, CAMx using PiG/PSAT for local major sources (CAMx+PiG), and the combined CAMx regional plus AERMOD local major source primary $PM_{2.5}$ contributions (CAMx+AERMOD).

Figure 4-15 shows that $PM_{2.5}$ values are higher across the distribution when the contribution from local major sources is included, both using PiG/PSAT or AERMOD

techniques. The CAMx+AERMOD simulation performs better than the CAMx+PiG simulation across much of the frequency distribution, except for the highest observed concentrations ($> \sim 25 \mu\text{g}/\text{m}^3$). This suggests that CAMx+PiG is doing a better job than CAMx+AERMOD in reproducing the high observed $\text{PM}_{2.5}$ at Liberty that occurs during local inversion weather conditions. The annual average $\text{PM}_{2.5}$ in the CAMx+AERMOD simulation is $14.5 \mu\text{g}/\text{m}^3$, which is much closer to the observed annual average ($14.1 \mu\text{g}/\text{m}^3$) compared to the CAMx+PiG simulation ($17.1 \mu\text{g}/\text{m}^3$ when PiG values are averaged in each 1.33 km CAMx grid cell, $16.2 \mu\text{g}/\text{m}^3$ when PiG values are included in the CAMx grid cell using interpolation methods).

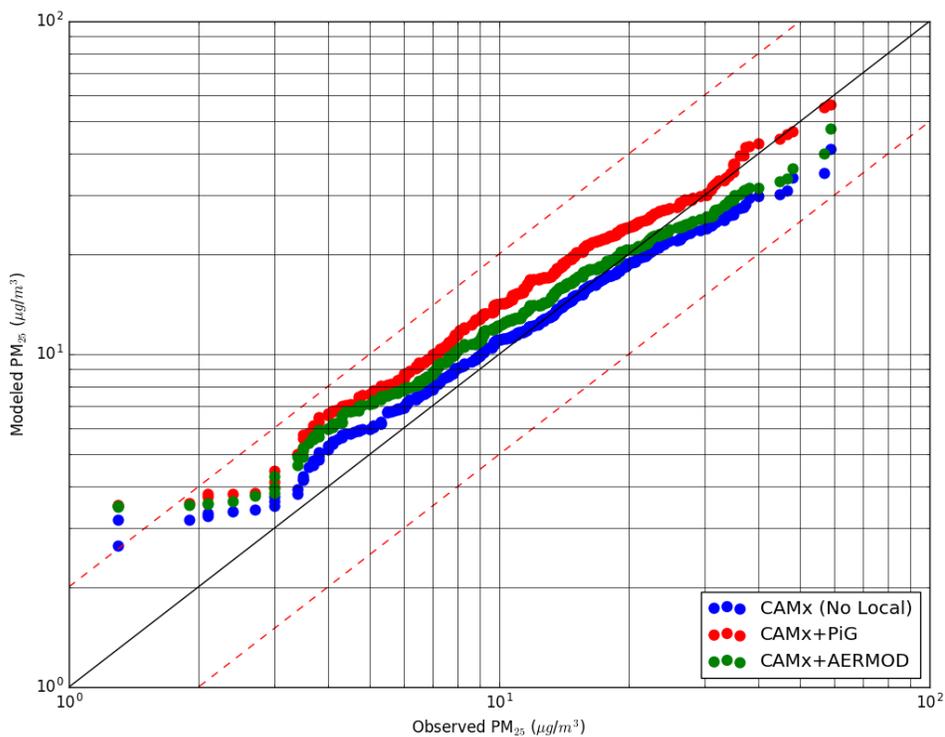


Figure 4-15. Annual Q-Q plot Comparing 24-hour $\text{PM}_{2.5}$ at Liberty in the Base Year Model Runs.

Using PiG/PSAT techniques and the average PiG value in each 1.33 km CAMx grid cell, the base year annual average $\text{PM}_{2.5}$ associated with emissions of primary $\text{PM}_{2.5}$ from local major sources at Liberty is $4.01 \mu\text{g}/\text{m}^3$. This estimate includes the gaseous emissions associated with local major sources in the live PiG puffs. Using

AERMOD techniques and the average AERMOD value in each 1.33 km CAMx grid cell, the annual average PM_{2.5} associated with emissions of primary PM_{2.5} from local major sources at Liberty in the base year is 1.44 µg/m³ for this particular AERMOD modeling configuration. These estimates are considerably lower than those predicted using PiG/PSAT, although this may be partially explained by the lack of local major source gaseous emissions in AERMOD. The AERMOD local source PM_{2.5} contribution at Liberty (1.44 µg/m³) is also considerably lower than the estimated actual local source contribution (3.8 µg/m³) taken from the difference in speciated PM_{2.5} concentrations at Liberty and other monitoring sites in Allegheny County (Figure 4-10).

Although Figure 4-15 suggests that CAMx+AERMOD simulation is in better agreement with PM_{2.5} observations at Liberty compared to the CAMx+PiG simulation for all but the highest contributions, this is likely because CAMx is overestimating the regional PM_{2.5} contribution and AERMOD is underestimating the local source contribution for this particular set of AERMOD modeling inputs and assumptions. The CAMx+PiG simulation does an excellent job reproducing the highest daily PM_{2.5} concentrations, which are relied upon to calculate 24-hr PM_{2.5} design values.

5. CONCLUSIONS

The 2011 base year CAMx results presented in this study indicate that there is good agreement between modeled and observed PM_{2.5} concentrations across Allegheny County. A suite of PM_{2.5} observation networks, including FRM, AQS, and CSN, are compared to CAMx model results at monitoring sites within Allegheny County. Fractional bias and error between modeled and observed hourly and daily total PM_{2.5} are within the performance criteria at all monitoring sites and meet the performance goals at the majority of sites.

Quarterly time series of daily PM_{2.5} at Liberty indicate that CAMx is able to replicate the overall temporal profile at Liberty during 2011. Scatter plots comparing daily PM_{2.5} values at monitoring sites across Allegheny County reveal that modeled PM_{2.5} is within a factor of two compared to observations the majority of the time.

The time series, scatterplots, and soccer plots indicate that there is better agreement during the warmer quarters (quarter 2 and 3). Winter overestimation bias was also found in previous CAMx results for southwestern Pennsylvania.

Q-Q plots showing hourly PM_{2.5} data at several monitor sites show good agreement (within a factor of two) between modeled and observed PM_{2.5} concentrations across the entire range of the frequency distribution. At Liberty, CAMx slightly underestimates the highest concentrations (59.0 vs. 56.4 µg/m³) but overestimates the low and mid-range concentrations, resulting in an annual average PM_{2.5} concentration of 16.4 µg/m³, which is slightly higher than the observed annual average PM_{2.5} (14.1 µg/m³). Modeled annual average PM_{2.5} concentrations at Clairton, Lawrenceville, and North Braddock are roughly 3 µg/m³ higher compared to observations, suggesting that CAMx is overestimating the urban portion of the regional component of PM_{2.5} within Allegheny County. At South Fayette, the most rural site by land use, CAMx underestimates the annual average PM_{2.5} by 1 µg/m³.

Using PiG/PSAT techniques, the model-estimated local component of PM_{2.5} in Allegheny County ranges from 0.3 to 4.9 µg/m³ at several monitoring sites, which is 3 to 30% of the total PM_{2.5} at these sites. Compared to Clairton, Lawrenceville, North Braddock, and South Fayette, the modeled local PM_{2.5} at Liberty is highest (4.9 µg/m³) and contributes most to total PM_{2.5} (30%). The local source contribution estimated using speciated PM_{2.5} observations at Liberty is 1 µg/m³

lower compared to the modeled estimate. There are differences between observed and modeled PM_{2.5} speciation at Liberty, which is at least partially driven by differences in the approaches used to estimate species concentrations and group individual species into larger categories. Nitrate, sulfate, and elemental carbon contribute the most to observed PM_{2.5} at Liberty, while fine particulates and soil ('other' category) contribute the most to modeled PM_{2.5}.

In addition to PiG/PSAT techniques, AERMOD is also used to estimate the PM_{2.5} associated with primary PM_{2.5} emissions from local major sources in Allegheny County. The CAMx+AERMOD model results appear to be in better agreement with FRM observations compared to the CAMx+PiG model results for all but the highest concentrations. However, this is likely because the local source PM_{2.5} underestimation in AERMOD partially offsets the overestimation of regional PM_{2.5} in CAMx. There is excellent agreement between observations and the CAMx+PiG scenario for the highest concentrations at Liberty.

Soccer plots are used to evaluate model performance for several PM_{2.5} species, including sulfate, nitrate, ammonium, and total carbon (elemental and organic). For the vast majority of species, sites, and seasons, the fractional bias and error metrics meet the performance criteria, which indicates reasonable model-measurement agreement. Similar to total PM_{2.5}, the best agreement is seen during the warmer seasons.

The model performance results presented in this study indicate that CAMx is a suitable model to be used in the future year modeling scenario. There is a tendency for CAMx to overestimate the regional PM_{2.5} contributions, which is potentially related to overstated PM_{2.5} boundary conditions and/or too high PM emissions from residential wood combustion. This overestimation suggests that the PM_{2.5} estimates predicted by CAMx during the future year simulation will likely be conservative.

6. REFERENCES

- Abt Associates, Inc. 2014. Modeled Attainment Test Software: User's Manual. Prepared for Office of Air Quality Planning and Standards, U.S. EPA. Research Triangle Park, NC.
- Coats, C.J. 1995. Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System, MCNC Environmental Programs, Research Triangle Park, NC.
- Emery, C., Z. Liu, A. Russell, T. Odman, G. Yarwood and N. Kumar. 2017. Recommendations on statistics and benchmarks to assess photochemical model performance. JAWMA, Vol. 67, Issue 5. (<http://www.tandfonline.com/doi/abs/10.1080/10962247.2016.1265027?journalCode=uawm20>).
- EPA. 2014. Draft Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone, PM_{2.5} and Regional Haze. U.S. Environmental Protection Agency, Research Triangle Park, NC. December. (http://www3.epa.gov/ttn/scram/guidance/guide/Draft-O3-PM-RH-Modeling_Guidance-2014.pdf)
- Houyoux, M.R., J.M. Vukovich, C.J. Coats, Jr. N.J.M. Wheeler, and P. Kasibhatla. 2000. Emission inventory development and processing for the seasonal model for regional air quality. *J. Geophysical Research*, **105**, (D7), 9079-9090.
- Nopmongcol, U., Jung, J., Zatko, M., Yarwood, G. 2016. Updated Boundary Conditions for CAMx Modeling. Final Report prepared for TCEQ, WO:582-16-62241-15.
- Nopmongcol, U., Jung, J., Zatko, M., Yarwood, G. 2017. 2013 Boundary Conditions for CAMx Modeling. Final Report prepared for TCEQ, WO:582-17-71444-21.
- Ramboll Environ. 2016. Allegheny County Health Department PM_{2.5} State Implementation Plan for the 2012 NAAQS – WRF Modeling Protocol. Ramboll Environ US Corporation, Lynnwood, WA and Novato, CA. March.
- Ramboll Environ. 2017. Allegheny County Health Department PM_{2.5} State Implementation Plan for the 2012 NAAQS – CAMx Modeling Protocol. Ramboll Environ US Corporation, Lynnwood, WA and Novato, CA. July.
- Ramboll Environ. 2018. Allegheny County Health Department PM_{2.5} State Implementation Plan for the 2012 NAAQS – Air Quality Technical Support Document. Ramboll Environ US Corporation, Lynnwood, WA and Novato, CA. January.

Sullivan, D. 1996. Review of Meteorology at the Clairton Area: Strengthening Dispersion Modeling for State Implementation Plans. Prepared for Allegheny County Health Department.

{This page left blank for printing purposes}

Appendix G.3

AERMOD Model Performance Evaluation

{This page left blank for printing purposes}



AERMOD Model Performance Evaluation

**Allegheny County, PA PM_{2.5}
Nonattainment Area, 2012 NAAQS**

**Allegheny County Health Department
Air Quality Program**

November 2018

{This page left blank for printing purposes}

TABLE OF CONTENTS

1	OVERVIEW	1
2	MODEL PERFORMANCE RESULTS	2
2.1	Local and Regional Impacts.....	2
2.2	Performance Statistics and Soccer Plots	3
2.3	Scatter and Quantile-Quantile Plots	9
2.4	Time Series Plots	12
2.5	Hourly Diurnal Plots	15
3	CONCLUSIONS.....	17
	APPENDIX.....	18

LIST OF TABLES

Table 2-1.	Local Observed and Modeled Excess ($\mu\text{g}/\text{m}^3$), Quarterly and Yearly, 2011.....	2
Table 2-2.	Regional Observed and Modeled Excess ($\mu\text{g}/\text{m}^3$), Quarterly and Yearly, 2011	3
Table 2-3.	Liberty Daily Statistics, Yearly, 2011	4
Table 2-4.	Liberty Daily Statistics, Quarterly, 2011	5
Table 2-5.	Liberty Daily Modeled Statistics for Benchmarks, Quarterly and Yearly, 2011.....	6

LIST OF FIGURES

Figure 2-1. Liberty Daily Soccer Plot, Mean Fractional Bias and Error, 2011	7
Figure 2-2. Liberty Daily Soccer Plot, Normalized Mean Bias and Error, 2011	8
Figure 2-3. Liberty Daily PM _{2.5} Scatter Plot.....	9
Figure 2-4. Liberty Daily PM _{2.5} Q-Q Plot, 2011	10
Figure 2-5. Liberty Daily PM _{2.5} Q-Q Plots, Log-Scale, Quarterly, 2011.....	11
Figure 2-6. Liberty Daily PM _{2.5} Time Series Plot, 1Q, 2011	12
Figure 2-7. Liberty Daily PM _{2.5} Time Series Plot, 2Q, 2011	13
Figure 2-8. Liberty Daily PM _{2.5} Time Series Plot, 3Q, 2011	13
Figure 2-9. Liberty Daily PM _{2.5} Time Series Plot, 4Q, 2011	14
Figure 2-10. Liberty Hourly PM _{2.5} Averages, 2011	15
Figure 2-11. Liberty Hourly PM _{2.5} Averages, Quarterly, 2011.....	16

Prepared with assistance from:

Ramboll
Environment and Health

Novato, CA
Lynnwood, WA



<http://www.ramboll.com/environment-and-health>

1 OVERVIEW

This document provides a model performance evaluation (MPE) for the modeling used for the Liberty local area analysis (LAA) for the Allegheny County, PA PM_{2.5} State Implementation Plan (SIP) for the 2012 National Ambient Air Quality Standards (NAAQS). Procedures for the modeling have been given in the AERMOD Modeling Protocol (see Appendix F of the SIP). Base case year (2011) is used for the evaluation since monitored results are available for comparison to modeled results.

Based on review of the CAMx¹ modeling results, the Allegheny County Health Department (ACHD) determined that additional local source modeling was needed for a refined analysis for the Liberty monitor site. The results of the attainment tests for the Liberty site will be given in Appendix I of the SIP.

The LAA modeling uses CAMx impacts without the largest local source impacts in combination with AERMOD² local source impacts for a more representative calculation of the projected PM_{2.5} species at Liberty. As part of the AERMOD modeling effort, ACHD also utilized an alternative BLP³/AERMOD hybrid approach for buoyant line sources, approved for use by the U. S. Environmental Protection Agency (EPA) Region 3 on August 16, 2018 (see Appendix H of the SIP).

¹ Comprehensive Air Quality Model with Extensions (<http://www.camx.com/>)

² American Meteorological Society/Environmental Protection Agency Regulatory Model (<https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models>)

³ Buoyant Line and Point Source dispersion model (<https://www3.epa.gov/ttn/scram/userg/regmod/blpug.pdf>)

2 MODEL PERFORMANCE RESULTS

This section provides statistics and figures used to evaluate the performance of the modeling compared to observed (monitored) results for 2011 at the Liberty monitor. Daily observed results were based on the Liberty FRM⁴ monitor, while hourly results were based on the Liberty TEOM⁵ monitor. Daily performance was the focus of this evaluation based on the standards and on model performance benchmarks, but some hourly results are given for a deeper look at the model simulation.

Due to the time difference between CAMx time in UTC⁶ and local time in EST,⁷ there are some missing modeled hours at the end of 2011. The last day (Dec. 31) was excluded from 24-hour averaging, and the last 5 hours of Dec. 30th were also missing from the hourly data comparisons.

Unless indicated, AERMOD local primary PM_{2.5} (LPM) impacts are based on the results for the Liberty expanded scale receptor grid as given in the AERMOD Model Protocol.

2.1 Local and Regional Impacts

Tables 2-1 and 2-2 below show the average observed and modeled concentrations of the Liberty localized excess and regional components, on quarterly⁸ and yearly bases. Observed excess was based on the weighted 2009-2013 differences between Liberty and the surrounding Pittsburgh MSA⁹ non-blank monitored concentrations. Modeled excess was based on the apportioned local and regional concentrations from the AERMOD LPM and CAMx non-LPM modeled concentrations.

Table 2-1. Local Observed and Modeled Excess ($\mu\text{g}/\text{m}^3$), Quarterly and Yearly, 2011

2011 Time Period	Observed Excess ($\mu\text{g}/\text{m}^3$)	Modeled Excess ($\mu\text{g}/\text{m}^3$)
1Q	3.30	3.03
2Q	2.94	3.13
3Q	2.99	3.57
4Q	5.74	5.27
Yearly	3.74	3.75

⁴ Federal Reference Method (used for official comparison to the NAAQS)

⁵ Tapered Element Oscillating Microbalance (used for unofficial continuous surveillance)

⁶ Coordinated Universal Time

⁷ Eastern Standard Time

⁸ Quarterly is based on calendar quarter (not seasonal)

⁹ Metropolitan Statistical Area

Table 2-2. Regional Observed and Modeled Excess ($\mu\text{g}/\text{m}^3$), Quarterly and Yearly, 2011

2011 Time Period	Observed Excess ($\mu\text{g}/\text{m}^3$)	Modeled Excess ($\mu\text{g}/\text{m}^3$)
1Q	9.76	15.89
2Q	9.82	9.85
3Q	12.27	12.20
4Q	8.82	12.27
Yearly	10.17	12.54

Modeled local concentrations showed good agreement with the observed excess at Liberty on both year-round and quarterly bases. Modeled regional concentrations showed the best performance in warmer periods (2nd and 3rd quarters), with some overprediction in cooler periods (1st and 4th quarters). As discussed in the CAMx Model Performance Evaluation (see Appendix G of the SIP), regional overprediction may be attributed to overstated model boundary conditions and/or overestimated PM emissions from residential wood combustion in and near urban areas.

Note that PM_{2.5} modeling is used on a relative basis and for species individually, with the ratio of modeled future case to modeled base case impacts (i.e., relative response factors, or RRFs) used for attainment test projections. Under or overprediction of absolute impacts of any component is not necessarily detrimental to the results.¹⁰

2.2 Performance Statistics and Soccer Plots

Table 2-3 shows the 2011 year-round statistics for the Liberty daily PM_{2.5} observed and modeled concentrations, based on the recommended statistics described in the AERMOD Modeling Protocol. Table 2-4 shows the same performance statistics by each quarter of 2011. Statistics are based on 2011 concentrations, paired in time for observed and modeled. Missing points (27 out of 365) are due to monitor/laboratory malfunctions for observed data or time zone differences for modeled data (described above).

¹⁰ Example: an observed base case high-day concentration is 40 $\mu\text{g}/\text{m}^3$ with 30% (12 $\mu\text{g}/\text{m}^3$) sulfate by composition. A modeled RRF of 0.5 for sulfate would reduce the high-day concentration to 34 $\mu\text{g}/\text{m}^3$ for the future case. The RRF of 0.5 can be derived from modeled concentrations that are equal to observed (future case 6 $\mu\text{g}/\text{m}^3$ divided by base case 12 $\mu\text{g}/\text{m}^3$), underpredicted (e.g., 4 divided by 8), or overpredicted (e.g., 10 divided by 20).

Table 2-3. Liberty Daily Statistics, Yearly, 2011

Liberty PM _{2,5} Daily Statistics, 2011		
METRIC	OBSERVED	MODELED
N (Number of Data Points)	338	338
Arithmetic Mean	14.06	16.57
Mean Bias	--	2.52
Mean Error	--	7.46
Root Mean Square Error	--	5.80
Normalized Mean Bias	--	0.18
Normalized Mean Error	--	0.41
Fractional Bias	--	0.25
Fractional Error	--	0.41
Correlation Coefficient	--	0.68
Factor of Two	--	0.82
Geometric Mean	11.32	14.69
Geometric Correlation Coefficient	--	0.54
Geometric Mean Variance	--	1.32
Robust Highest Concentration (N=26)	52	43

Year-round statistics show some overprediction for modeled data, with higher arithmetic and geometric means and positive bias compared to observed. (Error is reported as absolute values, which are always positive.) However, the robust highest concentration (RHC) is lower for modeled, so the highest peak values are underpredicted. The correlation coefficient (Pearson) and factor-of-two values indicate that there is good correlation between modeled and observed on a time-paired basis.

Table 2-4. Liberty Daily Statistics, Quarterly, 2011

Liberty PM _{2.5} Daily Statistics, 2011 (1Q)			Liberty PM _{2.5} Daily Statistics, 2011 (2Q)		
METRIC	OBSERVED	MODELED	METRIC	OBSERVED	MODELED
N (Number of Data Points)	84	84	N (Number of Data Points)	82	82
Arithmetic Mean	12.18	19.23	Arithmetic Mean	12.41	13.18
Mean Bias	--	7.06	Mean Bias	--	0.77
Mean Error	--	8.92	Mean Error	--	5.54
Root Mean Square Error	--	7.58	Root Mean Square Error	--	4.28
Normalized Mean Bias	--	0.58	Normalized Mean Bias	--	0.06
Normalized Mean Error	--	0.62	Normalized Mean Error	--	0.35
Fractional Bias	--	0.51	Fractional Bias	--	0.12
Fractional Error	--	0.53	Fractional Error	--	0.34
Correlation Coefficient	--	0.76	Correlation Coefficient	--	0.69
Factor of Two	--	0.69	Factor of Two	--	0.91
Geometric Mean	10.05	17.17	Geometric Mean	10.45	11.87
Geometric Correlation Coefficient	--	0.69	Geometric Correlation Coefficient	--	0.49
Geometric Mean Variance	--	1.52	Geometric Mean Variance	--	1.22
Robust Highest Concentration (N=26)	39	40	Robust Highest Concentration (N=26)	38	31
Liberty PM _{2.5} Daily Statistics, 2011 (3Q)			Liberty PM _{2.5} Daily Statistics, 2011 (4Q)		
METRIC	OBSERVED	MODELED	METRIC	OBSERVED	MODELED
N (Number of Data Points)	90	90	N (Number of Data Points)	82	82
Arithmetic Mean	15.45	15.90	Arithmetic Mean	16.12	17.98
Mean Bias	--	0.45	Mean Bias	--	1.87
Mean Error	--	6.18	Mean Error	--	8.67
Root Mean Square Error	--	4.58	Root Mean Square Error	--	6.85
Normalized Mean Bias	--	0.03	Normalized Mean Bias	--	0.12
Normalized Mean Error	--	0.30	Normalized Mean Error	--	0.42
Fractional Bias	--	0.07	Fractional Bias	--	0.29
Fractional Error	--	0.31	Fractional Error	--	0.46
Correlation Coefficient	--	0.67	Correlation Coefficient	--	0.78
Factor of Two	--	0.90	Factor of Two	--	0.77
Geometric Mean	13.16	14.22	Geometric Mean	11.77	16.05
Geometric Correlation Coefficient	--	0.56	Geometric Correlation Coefficient	--	0.71
Geometric Mean Variance	--	1.19	Geometric Mean Variance	--	1.38
Robust Highest Concentration (N=26)	43	37	Robust Highest Concentration (N=26)	68	38

Quarterly statistics show better model performance for the 2nd and 3rd quarters than the 1st and 4th quarter. Based on the results given in Table 2-2, the overprediction is mostly due to the regional component.

Key statistics used for comparison to model benchmarks for PM_{2.5} are shown in Table 2-5 by quarter and by year. The “goal” benchmarks are considered to be the best performance that a model can achieve, while the “criteria” benchmarks are considered to be average or reasonable performance.

Table 2-5. Liberty Daily Modeled Statistics for Benchmarks, Quarterly and Yearly, 2011

Metric	Goal	Criteria	1Q	2Q	3Q	4Q	Yearly
Fractional Bias (FB)	<±30%	<±60%	50.8%	12.0%	7.5%	29.2%	24.6%
Fractional Error (FE)	<50%	<75%	53.3%	34.3%	31.2%	45.7%	41.0%
Normalized Mean Bias (NMB)	<±10%	<±30%	58.0%	6.2%	2.9%	11.6%	17.9%
Normalized Mean Error (NME)	<35%	<50%	62.3%	34.5%	29.7%	42.5%	41.3%
Correlation Coefficient (r)	>0.70	>0.40	0.76	0.69	0.67	0.78	0.68

For year-round concentrations, modeled fractional bias and error fall within the goal range, while normalized mean and bias fall within the criteria range. On a quarterly basis, 2nd and 3rd quarters show the best performance for bias and error, with all data falling within the goal range. For 1st and 4th quarter, bias and error falls within the criteria range except for 1st quarter normalized bias and error.

Correlation coefficient falls just below goal on a year-round basis but well above criteria for all quarters. In contrast to bias and error, correlation is best for the cooler months, indicating that modeled peaks are occurring at the correct times during these periods.

Figures 2-1 and 2-2 visually show “soccer plots” of the fractional and normalized results, respectively. The soccer plots are scatter plots for bias and error, with the goal and criteria ranges shown as boxes (or “soccer goals”) for the quarterly and annual values shown in Table 2-5.

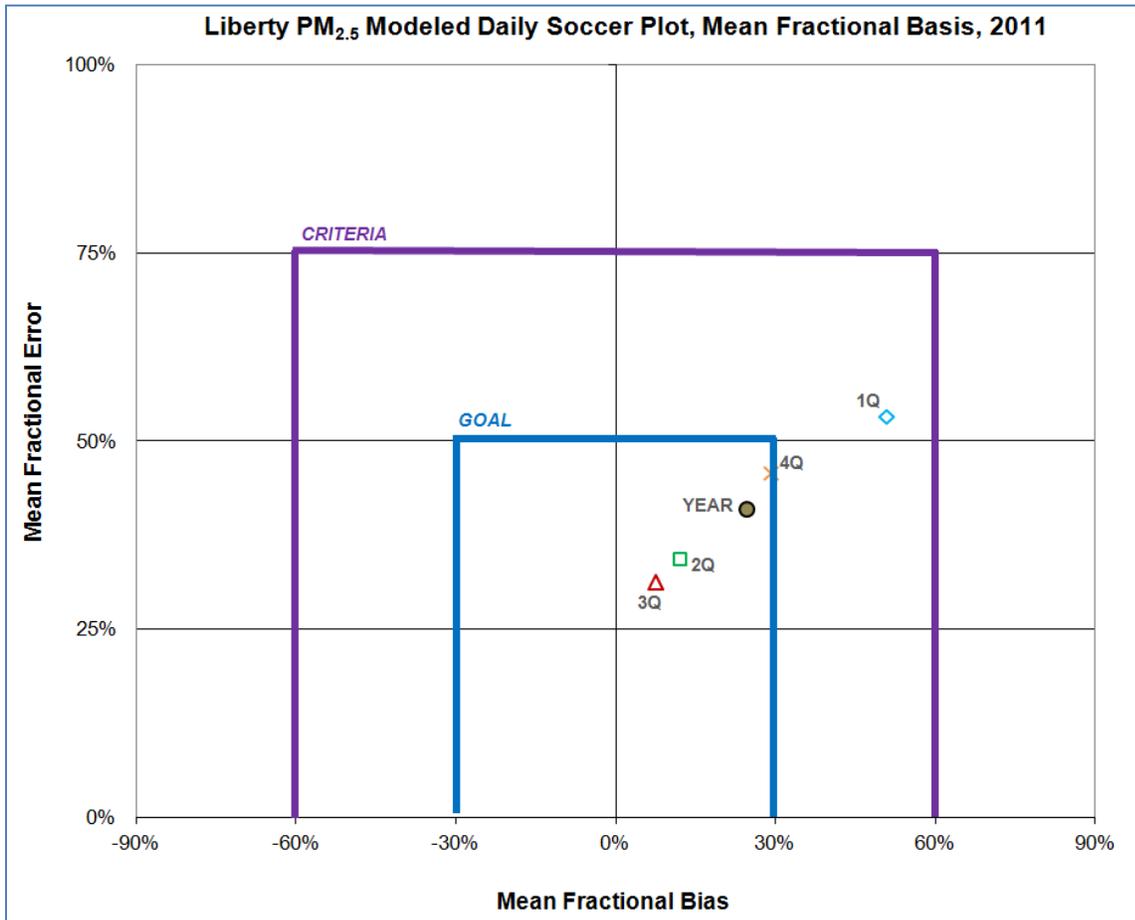


Figure 2-1. Liberty Daily Soccer Plot, Mean Fractional Bias and Error, 2011

The mean fractional soccer plot shows good performance year-round for the modeling, with 3rd quarter showing the best performance by quarter. The 1st quarter shows reasonable performance, falling in the criteria range.

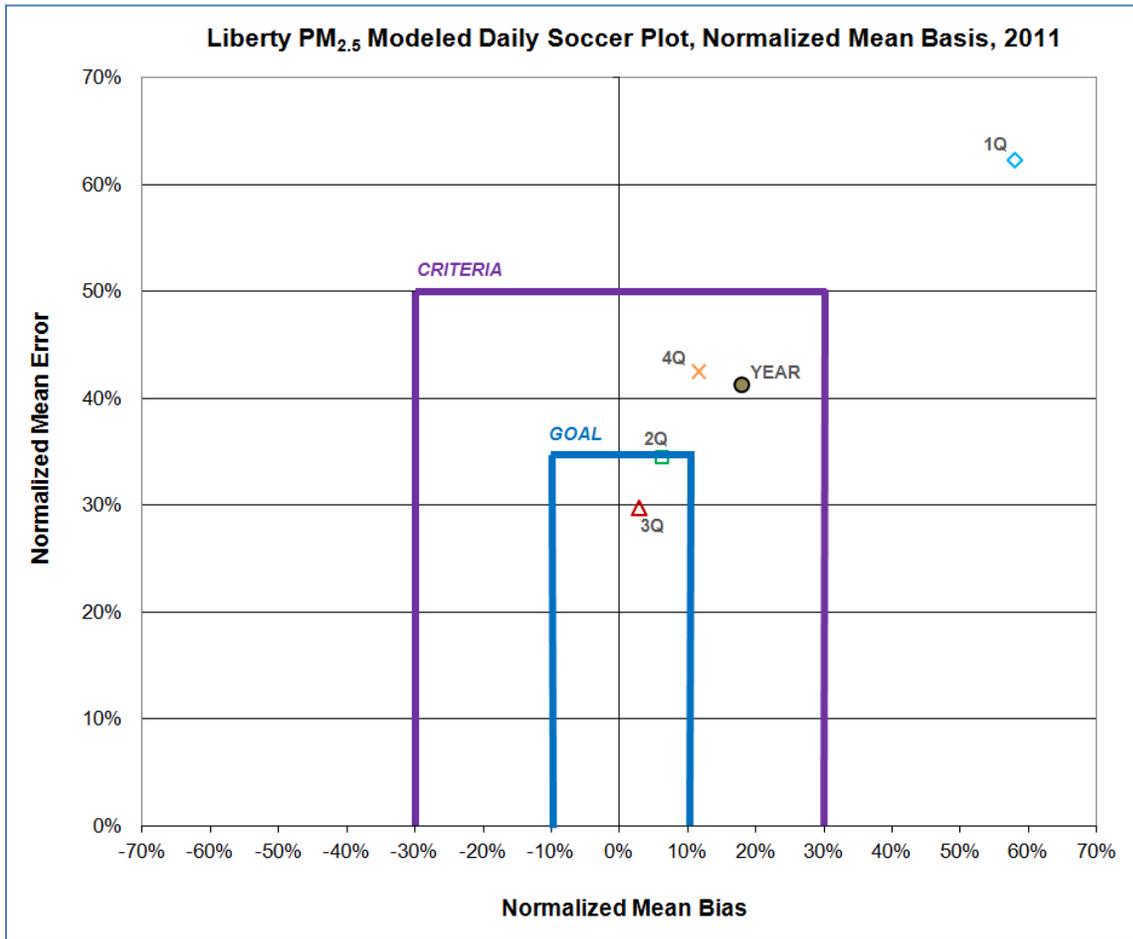


Figure 2-2. Liberty Daily Soccer Plot, Normalized Mean Bias and Error, 2011

On a normalized mean basis, with more strict benchmarks, the modeling produced reasonable results year-round, with the 2nd and 3rd quarters falling within the goal range and only the 1st quarter falling out of the criteria range.

2.3 Scatter and Quantile-Quantile Plots

Figure 2-3 shows the scatter plot for the Liberty 24-hour observed and modeled concentrations, paired in time for all valid days in 2011. The 1:1 line shown in orange indicates a perfect correlation between observed and modeled.

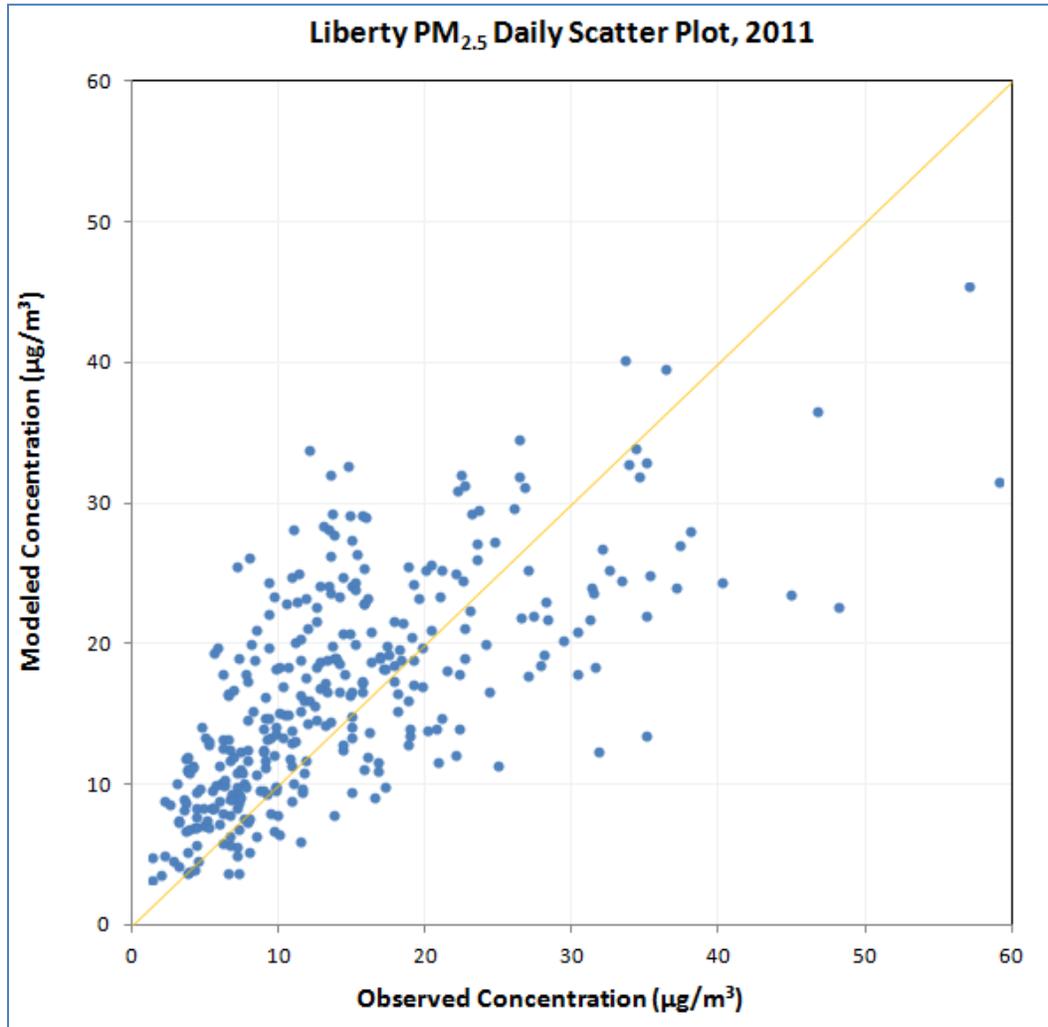


Figure 2-3. Liberty Daily PM_{2.5} Scatter Plot

There is reasonable yearly correlation between observed and modeled, with the correlation coefficient ($r=0.68$) falling just short of the goal benchmark ($r=0.70$). The plotted points are more dense above the 1:1 at concentrations below 30 µg/m³, indicating that the modeling is overpredicting at low to mid-range concentrations below the 24-hour NAAQS. The largest discrepancies are for the highest observed values, indicating that the modeling is underpredicting on days of the highest magnitude of concentrations.

Figure 2-4 shows the quantile-quantile (Q-Q) plot for the Liberty daily concentrations. A quantile-quantile plot is a scatter plot, with the values sorted by magnitude instead of on a time-paired basis, to examine performance for the highest and lowest concentrations. The orange 1:1 line indicates perfect correlation, with the gray lines showing the factor-of-two range for reasonable correlation.

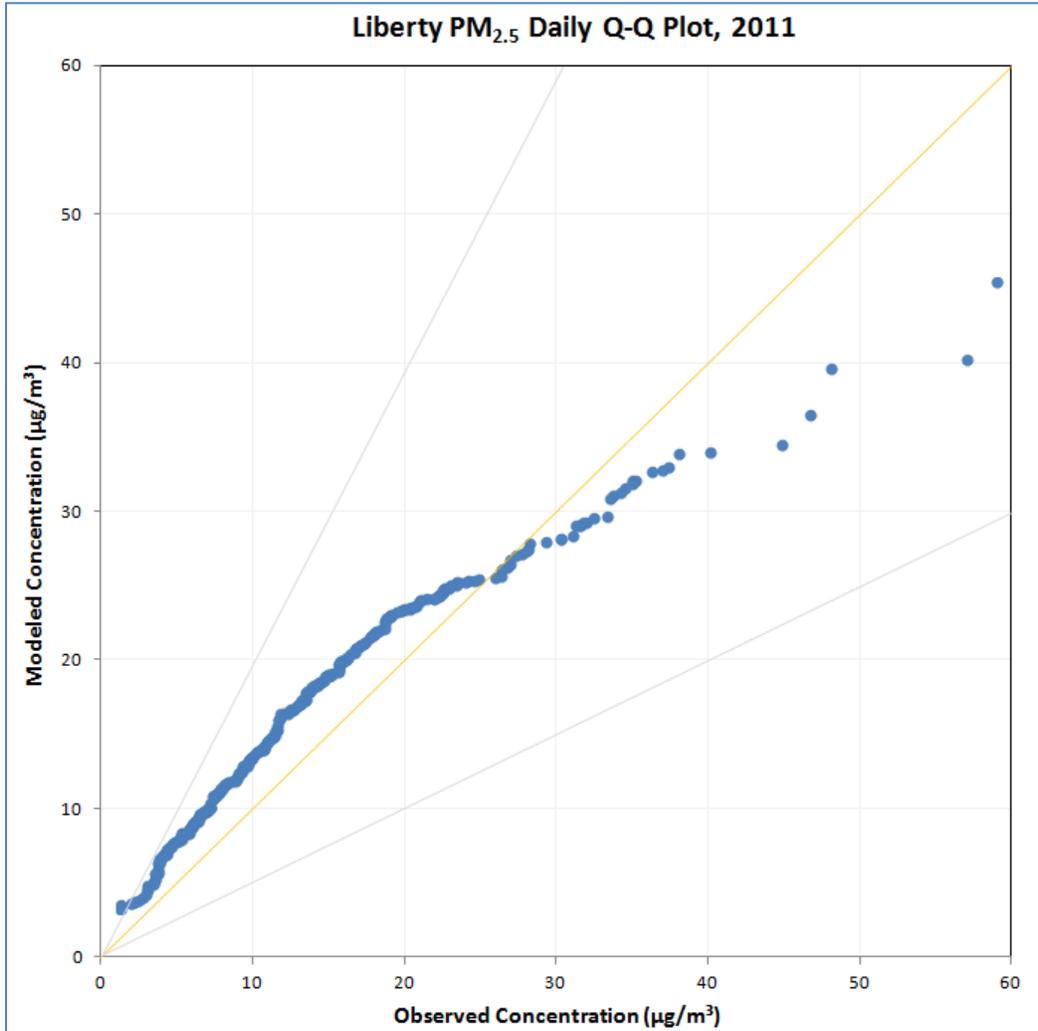


Figure 2-4. Liberty Daily PM_{2.5} Q-Q Plot, 2011

The Q-Q plot shows values that are close to the 1:1 line and with all values within a factor of two except for the lowest overall concentrations. Correlation is best for mid-range values, with some modeled underprediction at the highest concentrations.

Figure 2-5 shows the Liberty 24-hour Q-Q plots by quarter and on a logarithmic scale. A logarithmic scale visually emphasizes the variation from the 1:1 line for all ranges of concentrations.

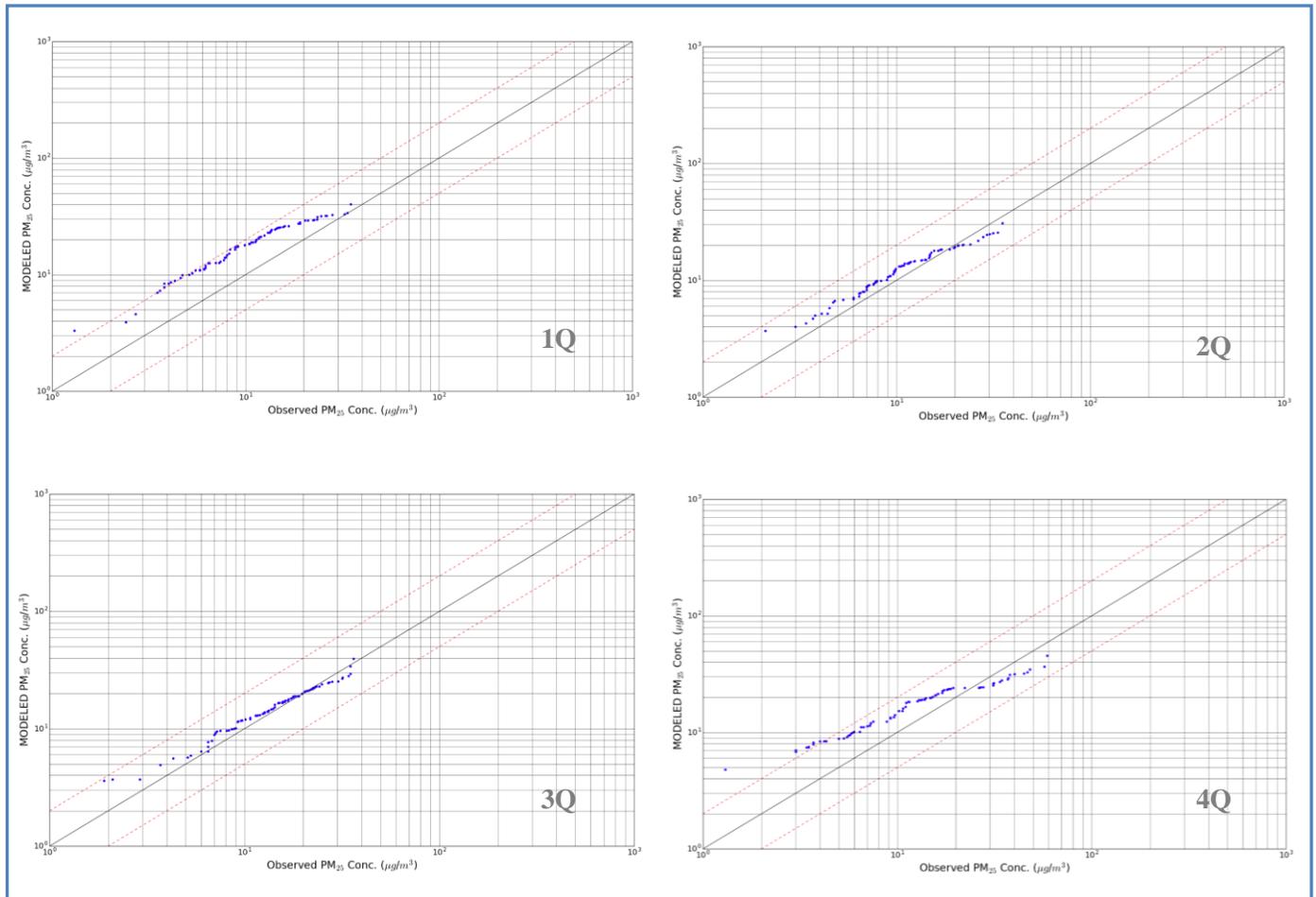


Figure 2-5. Liberty Daily PM_{2.5} Q-Q Plots, Log-Scale, Quarterly, 2011

Quantile-quantile results were best in the 2nd and 3rd quarters, with overprediction evident for several days in the 1st and 4th quarters.

2.4 Time Series Plots

Daily time series plots are shown for 2011 by quarter in Figures 2-6 through 2-9. Time series plots can reveal additional details on the model performance on individual days/seasons.

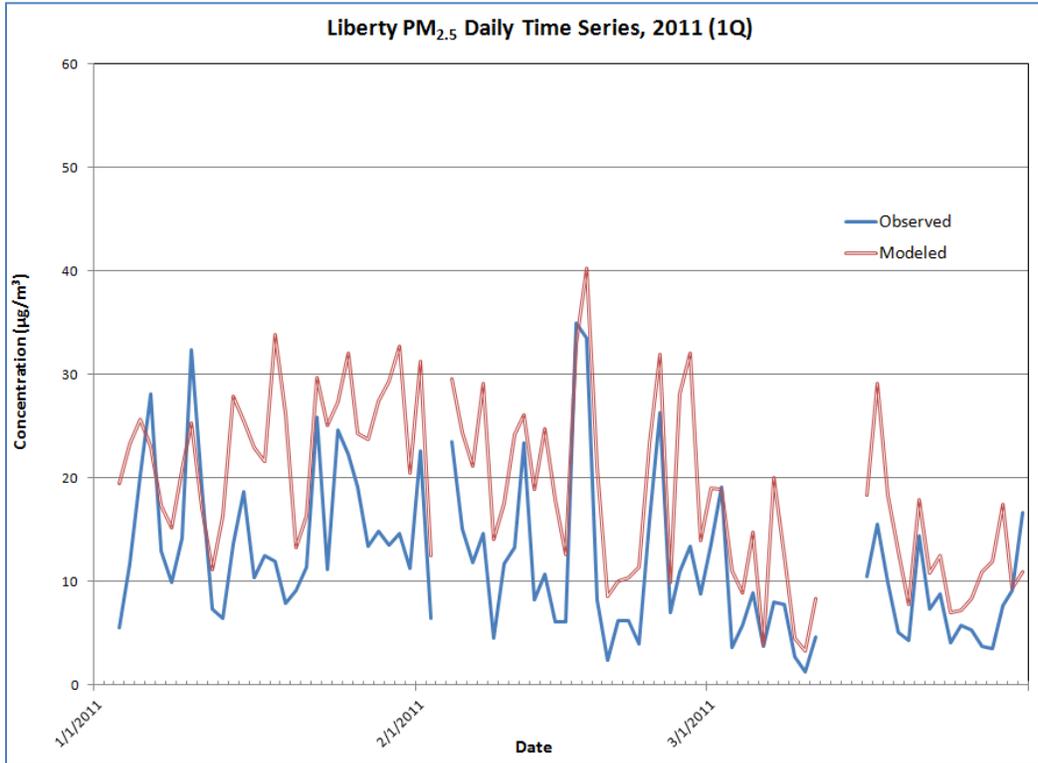


Figure 2-6. Liberty Daily PM_{2.5} Time Series Plot, 1Q, 2011

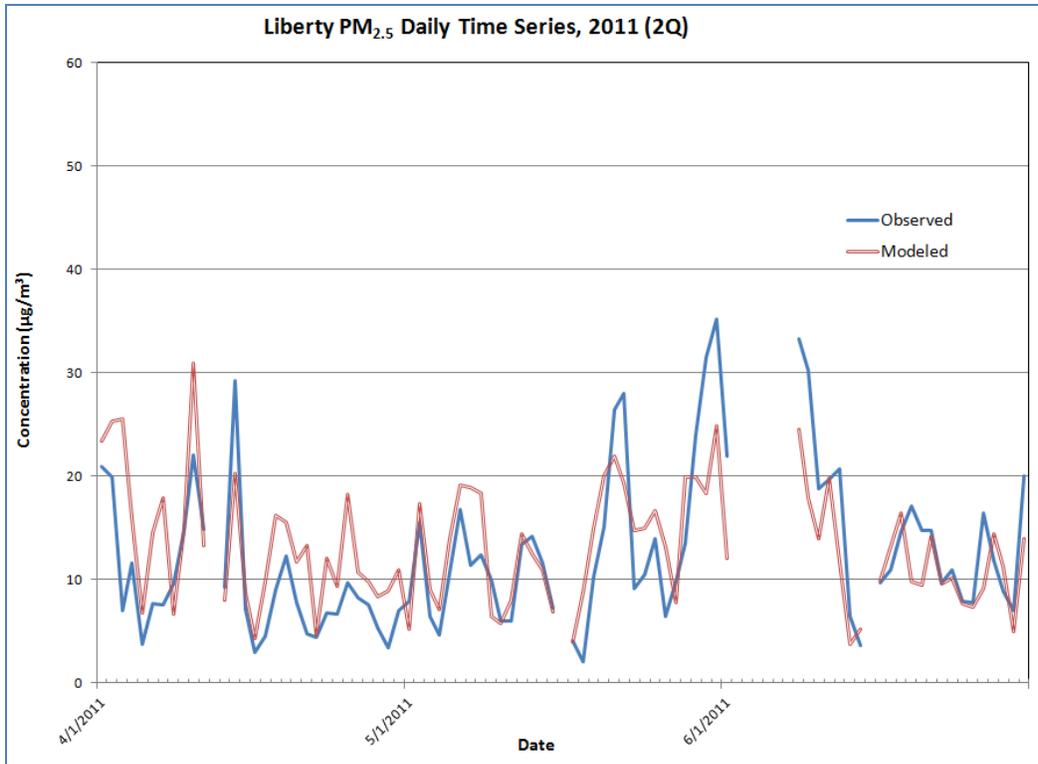


Figure 2-7. Liberty Daily PM_{2.5} Time Series Plot, 2Q, 2011

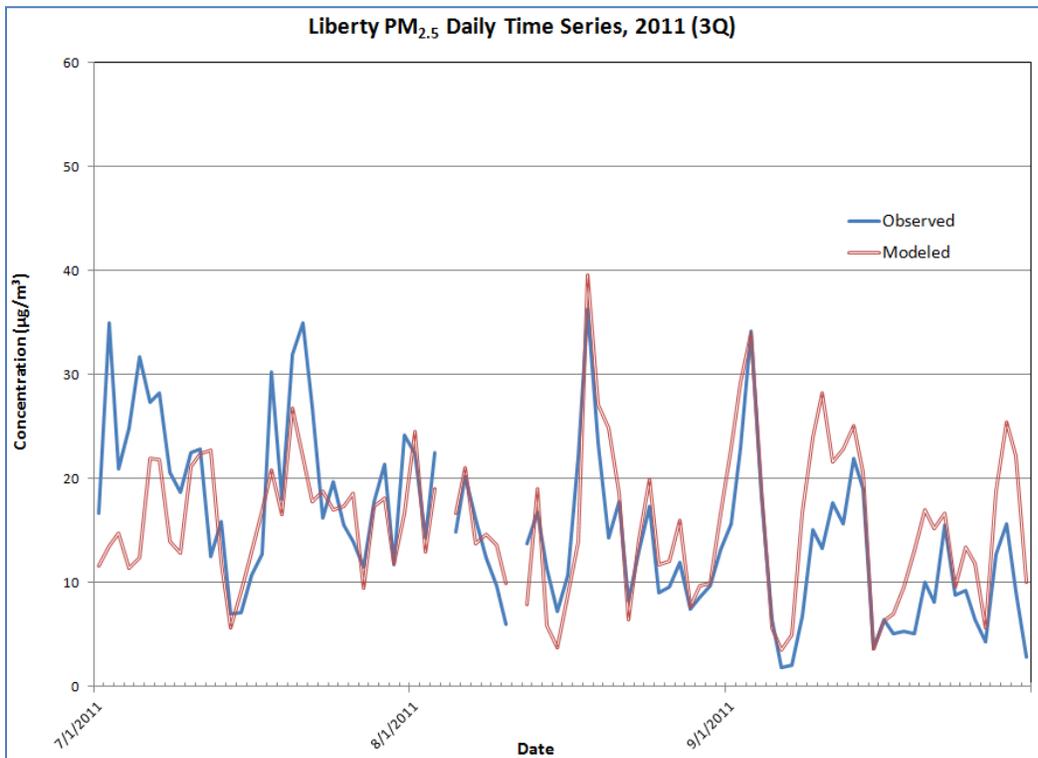


Figure 2-8. Liberty Daily PM_{2.5} Time Series Plot, 3Q, 2011

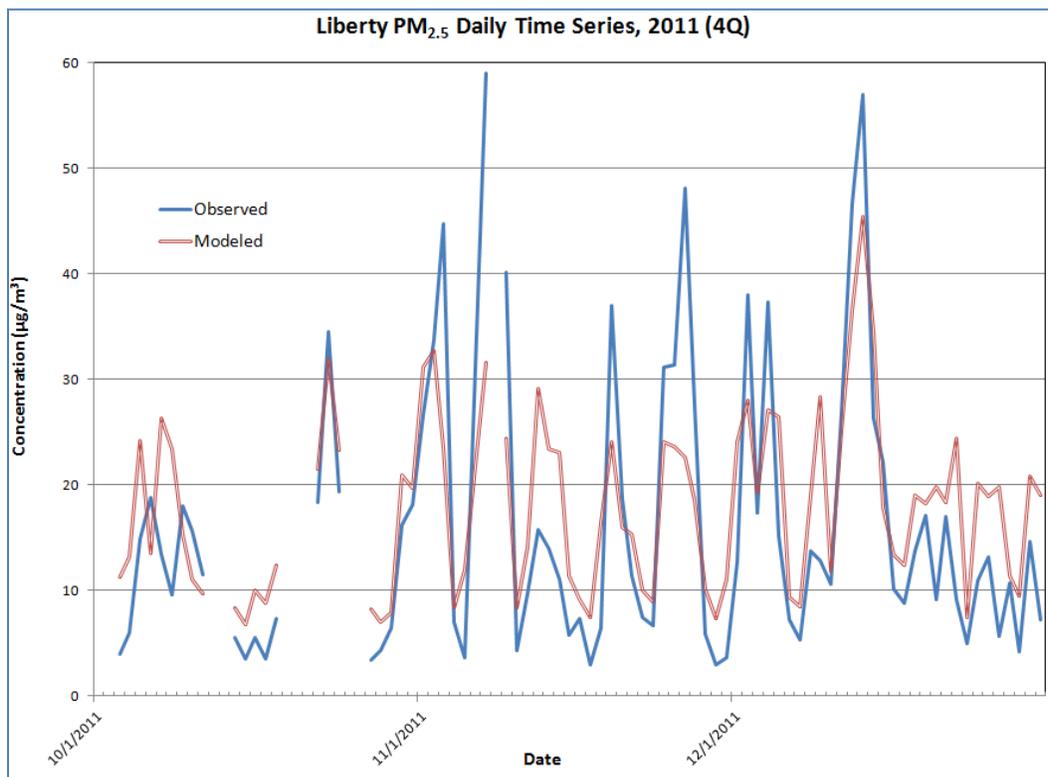


Figure 2-9. Liberty Daily PM_{2.5} Time Series Plot, 4Q, 2011

There is a fairly consistent overprediction in the 1st quarter due to the overestimated regional component in the coldest months (discussed earlier). Modeled results are closer to observed in warmer months in the 2nd and 3rd quarters, with some underprediction on mid-range days in these periods. Two high-days (8/18 and 9/3) were modeled nearly perfectly in the 3rd quarter. The 4th quarter, with the highest individual observed days, showed modeled underprediction on peak days. Thus, even after the refinements with AERMOD, the localized observed impacts during the worst case conditions can be difficult to simulate with steady-state sub-hourly modeling.

The observed and modeled concentrations for each day in 2011 are given in Table A-1 of the Appendix of this document.

2.5 Hourly Diurnal Plots

Hourly average plots are shown on 2011 yearly and quarterly bases in Figures 2-10 and 2-11, respectively. Hourly average plots can reveal specific hourly and diurnal patterns in the data.

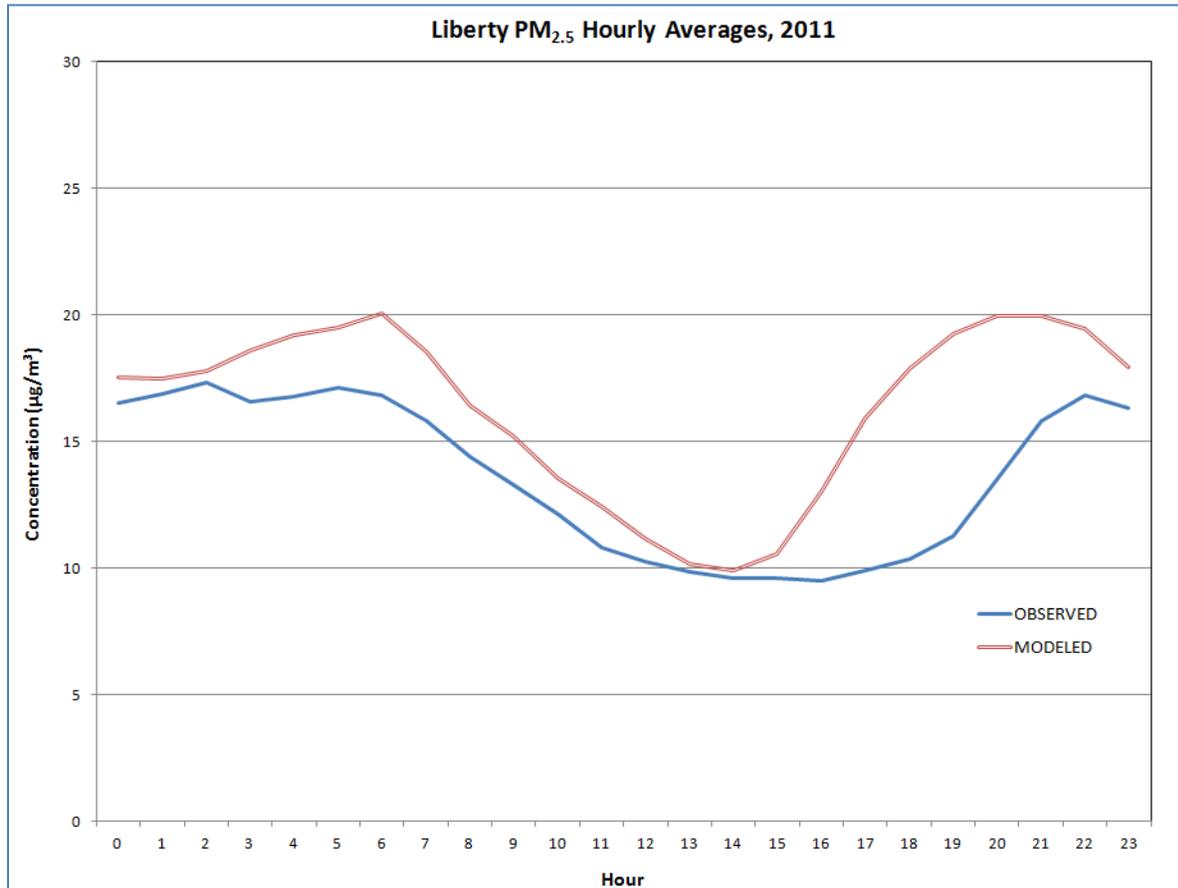


Figure 2-10. Liberty Hourly PM_{2.5} Averages, 2011

Year-round hourly averages show that the modeling is following the same diurnal behavior as the observed results, with the highest concentrations occurring during nighttime inversions. The midnight and midday concentrations are nearly equal for modeled and observed, with modeled overprediction occurring mainly during the day/night transitions.

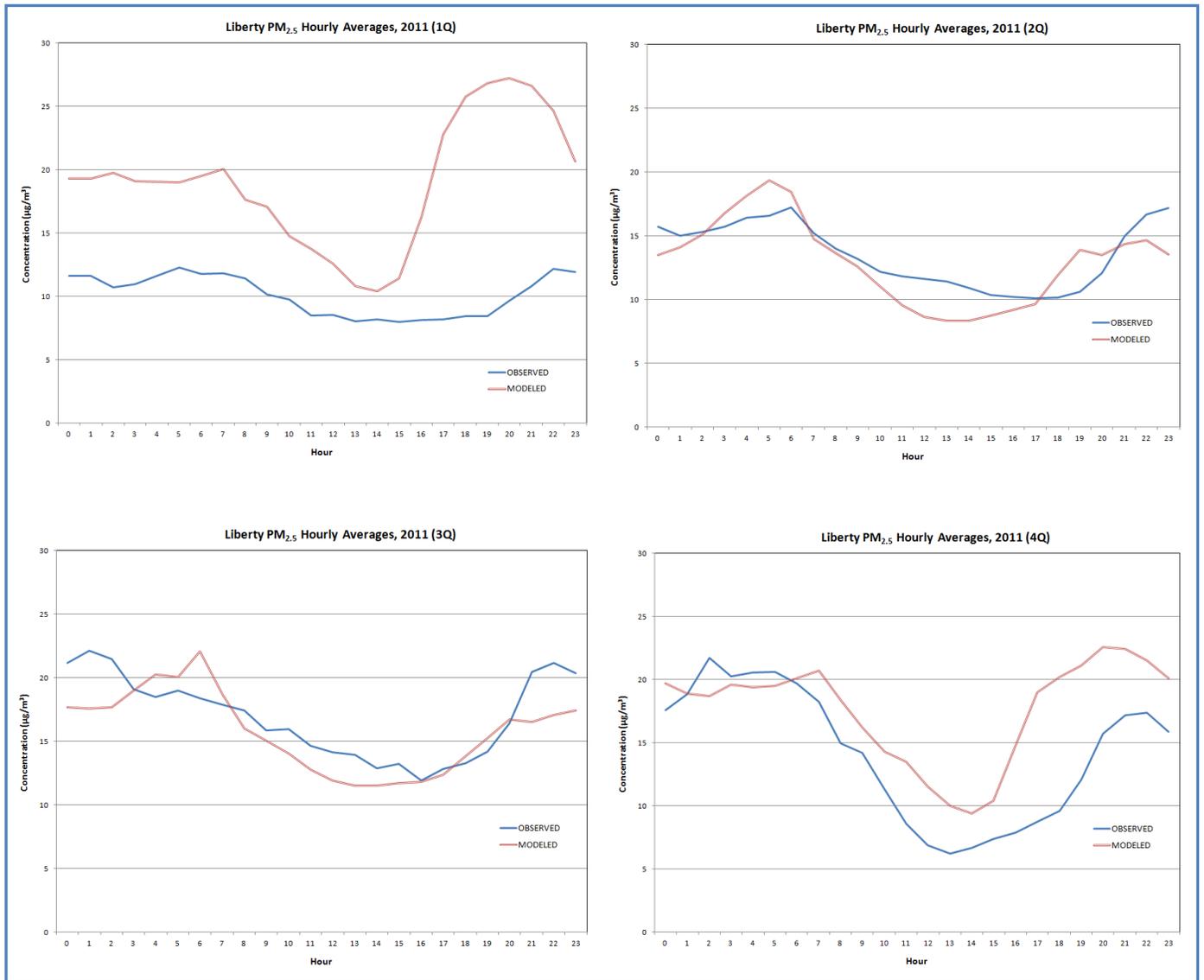


Figure 2-11. Liberty Hourly PM_{2.5} Averages, Quarterly, 2011

Quarterly hourly averages show that the 1st quarter shows the greatest differences between modeled and observed, notably during nighttime hours. The 2nd and 3rd quarters, similar to previous figures, show the overall best performance. In the 4th quarter, while modeling underpredicts the highest individual peaks in the daily time series plots, most nighttime hours are actually overpredicted on an average basis.

3 CONCLUSIONS

Overall, the modeling used for the Liberty LAA showed good performance for base case 2011 and is deemed to be suitable for the future case scenario. Localized excess was well-simulated by the model, with quarterly and year-round modeled averages that were very close to observed averages (with year-round modeled excess almost identical to observed). The regional component, based on the non-LPM CAMx impacts, includes some overprediction in cooler months. However, the localized component is the more important component during strong inversions in these seasons.

Model benchmarks for bias, error, and correlation statistics were achieved for daily total PM_{2.5} concentrations for the year and for all quarters except one. Scatter, quantile-quantile, and time series plots showed reasonable agreement between daily modeled and observed values, and hourly average plots showed similar diurnal behavior for hourly modeled and observed concentrations. Any modeled under or overprediction is compensated for by the use of relative reductions (instead of absolute concentrations) for projected design values.

APPENDIX

Daily observed and modeled concentrations for each quarter of 2011 (from left to right) are given below in Table A-1. Days with invalid observed or modeled concentrations are shown as missing data.

Table A-1. Liberty Daily PM_{2.5} Concentrations, 2011

Date (1Q)	Observed	Modeled	Date (2Q)	Observed	Modeled	Date (3Q)	Observed	Modeled	Date (4Q)	Observed	Modeled
1/1/2011	21.0	25.3	4/1/2011	20.9	23.5	7/1/2011	16.7	11.6	10/1/2011	1.3	4.8
1/2/2011			4/2/2011	19.9	25.3	7/2/2011	35.0	13.5	10/2/2011		
1/3/2011	5.5	19.5	4/3/2011	7.0	25.6	7/3/2011	21.0	14.8	10/3/2011	4.0	11.3
1/4/2011	11.7	23.4	4/4/2011	11.6	16.0	7/4/2011	24.9	11.4	10/4/2011	6.0	13.2
1/5/2011	20.3	25.7	4/5/2011	3.8	6.8	7/5/2011	31.7	12.4	10/5/2011	14.9	24.2
1/6/2011	28.1	23.0	4/6/2011	7.7	14.6	7/6/2011	27.3	22.0	10/6/2011	18.8	13.5
1/7/2011	13.0	17.3	4/7/2011	7.6	17.9	7/7/2011	28.2	21.8	10/7/2011	13.4	26.3
1/8/2011	9.9	15.2	4/8/2011	9.6	6.7	7/8/2011	20.6	14.0	10/8/2011	9.6	23.4
1/9/2011	14.2	20.8	4/9/2011	14.8	14.9	7/9/2011	18.7	12.9	10/9/2011	18.0	15.3
1/10/2011	32.4	25.3	4/10/2011	22.1	30.9	7/10/2011	22.5	21.2	10/10/2011	15.7	11.1
1/11/2011	19.1	17.2	4/11/2011	14.9	13.3	7/11/2011	22.9	22.4	10/11/2011	11.5	9.8
1/12/2011	7.3	11.2	4/12/2011			7/12/2011	12.5	22.7	10/12/2011		
1/13/2011	6.4	16.4	4/13/2011	9.3	8.0	7/13/2011	15.9	12.0	10/13/2011		
1/14/2011	13.6	27.9	4/14/2011	29.3	20.3	7/14/2011	7.0	5.6	10/14/2011	5.5	8.4
1/15/2011	18.7	25.5	4/15/2011	7.1	8.7	7/15/2011	7.1	9.2	10/15/2011	3.5	6.8
1/16/2011	10.4	23.0	4/16/2011	3.0	4.3	7/16/2011	10.7	13.0	10/16/2011	5.6	10.0
1/17/2011	12.5	21.7	4/17/2011	4.5	9.8	7/17/2011	12.7	17.0	10/17/2011	3.5	8.8
1/18/2011	12.0	33.9	4/18/2011	9.0	16.2	7/18/2011	30.3	20.9	10/18/2011	7.3	12.4
1/19/2011	7.9	26.2	4/19/2011	12.3	15.6	7/19/2011	18.0	16.5	10/19/2011		
1/20/2011	9.1	13.3	4/20/2011	7.8	11.8	7/20/2011	31.9	26.8	10/20/2011		
1/21/2011	11.4	16.4	4/21/2011	4.8	13.4	7/21/2011	35.0	22.0	10/21/2011		
1/22/2011	25.9	29.7	4/22/2011	4.4	4.7	7/22/2011	26.9	17.8	10/22/2011	18.4	21.5
1/23/2011	11.2	25.1	4/23/2011	6.8	12.0	7/23/2011	16.2	18.8	10/23/2011	34.5	32.0
1/24/2011	24.6	27.4	4/24/2011	6.7	9.3	7/24/2011	19.7	17.0	10/24/2011	19.4	23.3
1/25/2011	22.3	32.1	4/25/2011	9.7	18.3	7/25/2011	15.6	17.4	10/25/2011		
1/26/2011	19.1	24.3	4/26/2011	8.3	10.7	7/26/2011	14.0	18.6	10/26/2011		
1/27/2011	13.4	23.7	4/27/2011	7.6	9.9	7/27/2011	11.5	9.5	10/27/2011	3.4	8.2
1/28/2011	14.9	27.5	4/28/2011	5.3	8.3	7/28/2011	17.8	17.4	10/28/2011	4.3	7.0
1/29/2011	13.5	29.3	4/29/2011	3.4	9.0	7/29/2011	21.4	18.2	10/29/2011	6.5	7.9
1/30/2011	14.6	32.7	4/30/2011	7.0	10.9	7/30/2011	11.7	11.8	10/30/2011	16.2	20.9
1/31/2011	11.3	20.5	5/1/2011	7.9	5.2	7/31/2011	24.2	16.6	10/31/2011	18.1	19.7
2/1/2011	22.6	31.3	5/2/2011	15.6	17.3	8/1/2011	22.4	24.5	11/1/2011	26.7	31.2
2/2/2011	6.5	12.6	5/3/2011	6.5	9.1	8/2/2011	14.3	12.9	11/2/2011	33.7	32.8
2/3/2011			5/4/2011	4.7	7.1	8/3/2011	22.5	19.0	11/3/2011	44.8	23.6
2/4/2011	23.5	29.6	5/5/2011	10.7	13.9	8/4/2011			11/4/2011	7.0	8.4
2/5/2011	15.1	24.4	5/6/2011	16.8	19.1	8/5/2011	14.9	16.6	11/5/2011	3.7	12.0
2/6/2011	11.8	21.2	5/7/2011	11.4	18.9	8/6/2011	20.3	21.1	11/6/2011	31.1	21.8
2/7/2011	14.7	29.1	5/8/2011	12.4	18.4	8/7/2011	16.1	13.8	11/7/2011	59.0	31.6
2/8/2011	4.6	14.1	5/9/2011	9.9	6.5	8/8/2011	12.4	14.6	11/8/2011		
2/9/2011	11.7	17.6	5/10/2011	6.0	5.8	8/9/2011	9.7	13.7	11/9/2011	40.1	24.4
2/10/2011	13.3	24.2	5/11/2011	6.0	8.0	8/10/2011	6.0	10.0	11/10/2011	4.3	8.4
2/11/2011	23.4	26.1	5/12/2011	13.4	14.5	8/11/2011			11/11/2011	9.7	14.1
2/12/2011	8.2	18.9	5/13/2011	14.2	12.5	8/12/2011	13.7	7.9	11/12/2011	15.8	29.1
2/13/2011	10.7	24.8	5/14/2011	11.6	10.9	8/13/2011	16.8	19.0	11/13/2011	14.0	23.5
2/14/2011	6.1	17.9	5/15/2011	7.2	6.9	8/14/2011	11.4	5.9	11/14/2011	11.1	23.1
2/15/2011	6.1	12.6	5/16/2011			8/15/2011	7.2	3.7	11/15/2011	5.8	11.4
2/16/2011	35.0	33.0	5/17/2011	4.1	4.0	8/16/2011	10.7	8.9	11/16/2011	7.3	9.1

Table A-1. Liberty Daily PM_{2.5} Concentrations, 2011 (continued)

Date (1Q)	Observed	Modeled
2/17/2011	33.5	40.3
2/18/2011	8.3	21.1
2/19/2011	2.4	8.6
2/20/2011	6.2	10.0
2/21/2011	6.2	10.3
2/22/2011	4.0	11.4
2/23/2011	15.9	23.4
2/24/2011	26.3	31.9
2/25/2011	7.0	9.9
2/26/2011	10.9	28.2
2/27/2011	13.4	32.1
2/28/2011	8.8	14.0
3/1/2011	13.6	19.1
3/2/2011	19.1	19.0
3/3/2011	3.6	11.1
3/4/2011	5.8	8.9
3/5/2011	8.9	14.7
3/6/2011	3.8	3.9
3/7/2011	8.0	20.0
3/8/2011	7.8	12.5
3/9/2011	2.7	4.6
3/10/2011	1.3	3.3
3/11/2011	4.7	8.4
3/12/2011		
3/13/2011		
3/14/2011		
3/15/2011		
3/16/2011	10.5	18.4
3/17/2011	15.6	29.1
3/18/2011	9.9	18.4
3/19/2011	5.1	12.9
3/20/2011	4.3	7.8
3/21/2011	14.4	17.9
3/22/2011	7.4	10.9
3/23/2011	8.8	12.5
3/24/2011	4.1	7.0
3/25/2011	5.8	7.3
3/26/2011	5.3	8.4
3/27/2011	3.8	10.9
3/28/2011	3.5	11.9
3/29/2011	7.7	17.4
3/30/2011	9.1	9.4
3/31/2011	16.7	11.0

Date (2Q)	Observed	Modeled
5/18/2011	2.1	8.8
5/19/2011	10.3	15.0
5/20/2011	15.1	20.1
5/21/2011	26.4	21.9
5/22/2011	28.0	19.3
5/23/2011	9.2	14.8
5/24/2011	10.5	15.0
5/25/2011	14.0	16.7
5/26/2011	6.4	13.2
5/27/2011	9.8	7.8
5/28/2011	13.5	19.9
5/29/2011	24.0	20.0
5/30/2011	31.5	18.4
5/31/2011	35.2	24.9
6/1/2011	21.9	12.1
6/2/2011		
6/3/2011		
6/4/2011		
6/5/2011		
6/6/2011		
6/7/2011		
6/8/2011	33.3	24.5
6/9/2011	30.3	17.9
6/10/2011	18.8	14.0
6/11/2011	19.7	19.8
6/12/2011	20.7	11.7
6/13/2011	6.4	3.7
6/14/2011	3.7	5.2
6/15/2011		
6/16/2011	9.7	9.9
6/17/2011	11.0	13.2
6/18/2011	14.7	16.4
6/19/2011	17.1	9.8
6/20/2011	14.8	9.5
6/21/2011	14.8	14.2
6/22/2011	9.7	9.6
6/23/2011	10.9	10.1
6/24/2011	7.9	7.7
6/25/2011	7.8	7.3
6/26/2011	16.4	9.2
6/27/2011	11.8	14.4
6/28/2011	8.9	11.3
6/29/2011	7.0	5.0
6/30/2011	20.0	13.9

Date (3Q)	Observed	Modeled
8/17/2011	22.2	14.0
8/18/2011	36.3	39.6
8/19/2011	23.4	27.2
8/20/2011	14.3	24.8
8/21/2011	17.8	18.5
8/22/2011	8.3	6.4
8/23/2011	13.1	14.3
8/24/2011	17.3	19.9
8/25/2011	9.0	11.7
8/26/2011	9.6	12.1
8/27/2011	12.0	16.0
8/28/2011	7.5	7.7
8/29/2011	8.6	9.7
8/30/2011	9.7	9.9
8/31/2011	13.2	16.7
9/1/2011	15.7	22.9
9/2/2011	23.0	29.3
9/3/2011	34.2	34.0
9/4/2011	18.2	18.9
9/5/2011	6.5	5.7
9/6/2011	1.9	3.6
9/7/2011	2.1	4.9
9/8/2011	6.8	16.8
9/9/2011	15.1	24.0
9/10/2011	13.3	28.2
9/11/2011	17.7	21.6
9/12/2011	15.7	22.9
9/13/2011	21.9	25.1
9/14/2011	19.0	20.6
9/15/2011	3.7	3.7
9/16/2011	6.5	6.4
9/17/2011	5.1	7.0
9/18/2011	5.3	9.6
9/19/2011	5.1	13.1
9/20/2011	10.1	17.0
9/21/2011	8.1	15.3
9/22/2011	15.6	16.7
9/23/2011	8.8	9.6
9/24/2011	9.3	13.4
9/25/2011	6.5	11.8
9/26/2011	4.3	5.7
9/27/2011	12.7	18.8
9/28/2011	15.7	25.4
9/29/2011	9.2	22.2
9/30/2011	2.9	10.1

Date (4Q)	Observed	Modeled
11/17/2011	3.0	7.4
11/18/2011	6.4	16.5
11/19/2011	37.0	24.1
11/20/2011	18.7	16.0
11/21/2011	11.4	15.3
11/22/2011	7.5	10.1
11/23/2011	6.7	8.9
11/24/2011	31.2	24.1
11/25/2011	31.4	23.7
11/26/2011	48.1	22.7
11/27/2011	27.7	18.5
11/28/2011	5.9	10.1
11/29/2011	3.0	7.4
11/30/2011	3.7	11.1
12/1/2011	12.7	24.1
12/2/2011	38.0	28.0
12/3/2011	17.4	19.3
12/4/2011	37.3	27.1
12/5/2011	15.2	26.5
12/6/2011	7.2	9.4
12/7/2011	5.3	8.4
12/8/2011	13.8	19.1
12/9/2011	12.9	28.4
12/10/2011	10.6	11.9
12/11/2011	26.9	25.3
12/12/2011	46.7	36.6
12/13/2011	57.0	45.4
12/14/2011	26.3	34.6
12/15/2011	22.2	17.9
12/16/2011	10.2	13.4
12/17/2011	8.8	12.4
12/18/2011	13.8	19.1
12/19/2011	17.1	18.3
12/20/2011	9.2	19.8
12/21/2011	17.0	18.3
12/22/2011	9.2	24.4
12/23/2011	5.0	7.5
12/24/2011	11.0	20.1
12/25/2011	13.2	18.9
12/26/2011	5.7	19.8
12/27/2011	10.7	11.4
12/28/2011	4.2	9.5
12/29/2011	14.7	20.8
12/30/2011	7.2	19.0

{End of document}