Environment

Prepared for:
Bureau of Land Management
Utah State Office

Submitted by:
AECOM
Fort Collins, Colorado
February 2013

Utah Air Resource Management Strategy Modeling Project: Meteorological Model Performance Evaluation

AECOM
1601 Prospect Parkway
Fort Collins, Colorado 80525

970-493-8878 tel
970-493-0213 fax

February 1, 2013

Leonard Herr
BLM Utah State Office
440 West 200 South, Suite 500
Salt Lake City, UT 84101

Subject: Final Utah Air Resource Management Strategy Modeling Project: Meteorological Model Performance Evaluation

Dear Mr. Herr:

AECOM Technical Services, Inc. (AECOM) is pleased to submit the Final *Utah Air Resource Management Strategy Modeling Project: Meteorological Model Performance Evaluation* (referred to as the WRF MPE Report). The WRF MPE Report describes the meteorological model configuration, testing, and final results that were conducted for the Utah Bureau of Land Management's (BLM) Air Resource Management Strategy (ARMS) Modeling Project. The primary objective of the ARMS Modeling Project is to develop an air quality management tool that is appropriate to assess the potential air impacts from future activities occurring on BLM-administered land in the Uinta Basin. The reusable modeling framework developed as part of the ARMS Modeling Project also could be used to assess the potential cumulative impacts from future project-specific NEPA actions, which will facilitate consistency and efficiency with the planning activities in the area.

Enclosed with this letter is an electronic copy of the final WRF MPE Report, in Adobe format (PDF), for distribution to the Resource Technical Advisory Group (RTAG), as well as our response to all written comments received from the RTAG members.

If you have any questions relative to this protocol, or would like to discuss this study, please contact either Courtney Taylor or Zion Wang, or call (970) 493-8878.

Yours sincerely,

Courtney Taylor
Project Manager
Courtney.Taylor@aecom.com

Zion Wang
Senior Technical Advisor
Zion.Wang@aecom.com

cc: Stephen Reid
Kenneth Craig
Valerie Randall

Executive Summary

This report presents the Meteorological Model Performance Evaluation for the Air Resource Management Strategy (ARMS) Modeling Project. This study is being conducted by AECOM Technical Services, Inc., (AECOM) under the direction of the Bureau of Land Management (BLM) Utah State Office (Utah BLM).

The BLM is required to complete a National Environmental Policy Act (NEPA) analysis for each proposed project that would occur on BLM-administered federal land. The ARMS Modeling Project is designed to develop a reusable air management tool applicable to multiple projects for activities in the Uinta Basin. The ARMS modeling framework also could be used to assess the potential cumulative impacts associated with future project-specific NEPA actions, which will facilitate consistency and efficiency with the planning activities in the area. The Uinta Basin is an area in northeastern Utah that is projected to have extensive development of oil and gas reserves in the foreseeable future.

A gridded meteorological dataset is a necessary component of the modeling system proposed for the ARMS Modeling Project. The Weather Research and Forecasting (WRF) meteorological model was selected. The WRF configuration was tested extensively for the Uinta Basin Study Area to determine a preferred WRF configuration for the annual simulation. The result of these test led to two configurations: one for winter months and another for non-winter months. **Table ES-1** shows the WRF configuration. The primary differences between the two configurations are the planetary boundary layer scheme, the microphysics scheme, the short-wave radiation scheme, and the land surface model.

Both qualitative and quantitative (statistical) analyses were used to examine the performance of the final annual WRF simulation. Qualitative analyses of the meteorological model performance were conducted for four air quality episodes. The four selected episodes are: January 8 to 23, 2010; February 21 to March 8, 2010; August 19 to 29, 2010; and September 27 to October 5, 2010. The model results for these time periods were compared with:

- Observations of surface and upper-level pressure patterns;
- The spatial variability of observed precipitation, precipitation amounts, and snow cover; and
- The observed vertical profiles of wind speed, direction, temperature, and dew point.

In general, it was found that the WRF model was capable of reproducing the observed synoptic and precipitation patterns, including snow cover, during the events analyzed; however, the model tended to over-predict the extent of snow coverage during shoulder seasons. The model generally was able to simulate the vertical profiles of the atmosphere, including the vertical variability in wind direction and speed, as well as the height of the planetary boundary layer. However, the model had difficulty replicating sharp vertical changes in the dew point temperature. Altogether, the model's ability to reproduce important synoptic and vertical patterns provides confidence in the model's ability to reproduce important physical processes during periods with elevated concentrations of air pollutants.

The quantitative assessment of the 2010 annual simulation compared model results to observations using various statistical measures. The statistical results were evaluated over different temporal and spatial extents to assess the WRF model's performance for accuracy, consistency, and reasonableness with respect to available observations. Statistical summaries were generated for the 4-kilometer (km), 12-km, and 36-km model domains with a focus on the assessment of the 4-km results. In addition to domain-wide statistical summaries, the model performance was evaluated exclusively for the Uinta Basin Study Area to provide additional information about the area of interest for the ARMS study.

In general, the 2010 annual simulation performed well for all meteorological parameters evaluated. The model results were slightly better for the 4-km domain than the 12-km domain, likely as a result of both the finer resolution grid and the use of observation nudging in the 4-km domain. On an annual and seasonal basis, most meteorological parameters were within the traditional performance benchmarks. Moreover, when the results are evaluated relative to performance benchmarks for complex terrain, all results for the 4-km domain are within the accepted range.

Table ES-1 Key Parameters for the ARMS WRF 4-km Configuration

Parameter	Non-Winter Months[1]	Winter Months[2]
Horizontal grid	166 x 175 cells	166 x 175 cells
Vertical layers	36	36
Microphysics	Single moment (6-class)	Lin *et al.*., scheme
Cumulus parameterization	None[3]	None[3]
Planetary boundary layer (PBL)	Mellor-Yamada-Janjic (MYJ)	Asymmetric Convective Model Version 2 (ACM2)
Surface layer	Monin Obukov (Janic) scheme	Pleim-Xiu scheme
Land surface model (LSM)	Unified Noah Land Surface Model	Pleim-Xiu scheme
Long-wave radiation	Rapid Radiative Transfer Model (RRTM)	RRTM
Short-wave radiation	Goddard	Dudhia
Analysis nudging at the surface	n/a[4]	n/a[4]
Analysis nudging aloft	u/v/T/Q[4]	u/v/T/Q[4]
Observations nudging	u/v/T[4]	u/v/T[4]

[1] April, May, June, July, August, September, October, and November are defined as non-winter months for 2010.

[2] January, February, March, and December are defined as winter months for 2010.

[3] Typically, cloud convection is well resolved in the 4-km grid without the need for additional physics parameterizations.

[4] n/a = physics option is not applicable; u = velocity component in the east-west direction; and v = velocity component in the north-south direction, T = temperature, and Q = mixing ratio.

Based on the model performance evaluation, it is found that the model tends to under-predict wind speeds and temperature during winter, while over-predicting temperature in the fall and mixing ratio in the summer. In the Uinta Basin, the model wind speed tends to be biased slightly low independent of season with somewhat higher errors in the summer season. Model wind direction tends to be biased low during summer months and biased high in winter months. The model tends to over-estimate temperature in winter months and mixing ratio in summer months.

Based on these findings, the 2010 annual ARMS WRF modeling simulation demonstrated good performance and is considered suitable for use as input to an air quality model for the ARMS Modeling Project. Modeled air quality impacts will be evaluated with respect to the identified meteorological model limitation and biases.

List of Acronyms

ACM2	Asymmetric Convective Model Version 2
AECOM	AECOM Environment
ARMS	Air Resource Management Strategy
ARW	Advanced Research Weather Research and Forecasting
BLM	Bureau of Land Management
deg	degrees from true north
FDDA	four-dimensional data assimilation
g/kg	grams per kilogram
GMT	Greenwich Mean Time
IOA	Index of Agreement
K	degrees Kelvin
km	kilometer
LSM	land surface model
m/s	meters per second
MADIS	Meteorological Assimilation Data Ingest System
MDT	Mountain Daylight Time
MM5	Mesoscale Model Version 5
MPE	model performance evaluation
MST	Mountain Standard Time
MYJ	Mellor-Yamada-Janjic
NAM	North American Model
NCAR	National Center for Atmospheric Research
NIC IMS	National Ice Center Interactive Multisensor Snow and Ice Mapping System
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
OBSGRID	WRF preprocessing program
PBL	planetary boundary layer
$PM_{2.5}$	PM with an aerodynamic diameter less than or equal to 2.5 microns
RMSE	root mean square error
RRTM	Rapid Radiative Transfer Model
TDL	Techniques Development Laboratory
U.S.	United States
UCAR	University Corporation for Atmospheric Research
Utah BLM	Utah State Office
UV	ultraviolet

WRF	Weather Research and Forecasting

Contents

List of Appendices

List of Tables

List of Figures

1.0 Introduction

This report presents the results of the meteorological model performance evaluation developed for the Air Resource Management Strategy (ARMS) Modeling Project. This study is being conducted by AECOM Technical Services, Inc. (AECOM) under the direction of the Bureau of Land Management (BLM) Utah State Office (Utah BLM).

1.1 Air Quality Study Overview

The Utah BLM's goal is to develop a reusable modeling framework suitable for air quality management decisions affecting the Uinta Basin. For this modeling study, an air quality study area has been defined for the Uinta Basin. The Uinta Basin air quality study area is shown in **Figure 1-1**, and contains portions of Carbon, Duchesne, Daggett, and Uintah counties in Utah and Moffat and Rio Blanco counties in Colorado. The Uinta Basin study area encompasses most of the area administered by the BLM Vernal Field Office, as well as portions of the Price Field Office in Utah and White River Field Office and Little Snake Field Office in Colorado. In addition to BLM-administered land, the study area also includes state, private, and tribal lands and areas administered by other federal agencies. The Uinta Basin air quality study area was developed based on topographic features that influence air flow patterns, not political or geological boundaries. It is important to note that the Uinta Basin air quality study area does not completely contain the geological extents of the Uinta Basin's oil and gas reserves.[1] The air quality study area does completely contain the areas of the basin that have historically had elevated ozone concentrations (Energy Dynamics Lab 2011), which are of most concern for this air quality study.

1.2 Purpose of Meteorological Modeling and the Model Performance Evaluation

The purpose of the meteorological modeling is to develop a gridded meteorological dataset required for air quality modeling. The Weather Research and Forecasting (WRF) Model was selected for this study (AECOM 2012a). Since year 2010 was selected for the ARMS Modeling Project (AECOM 2012a), the meteorological model was configured to generate gridded meteorological data for year 2010. The WRF model configuration was tested extensively for use in this region. The results of these sensitivity tests are documented in **Appendix A** of this report. The final ARMS WRF simulation will be used in all air quality model simulations for the ARMS Project, including a 2010 base case simulation, a typical year simulation, and multiple future year simulations.

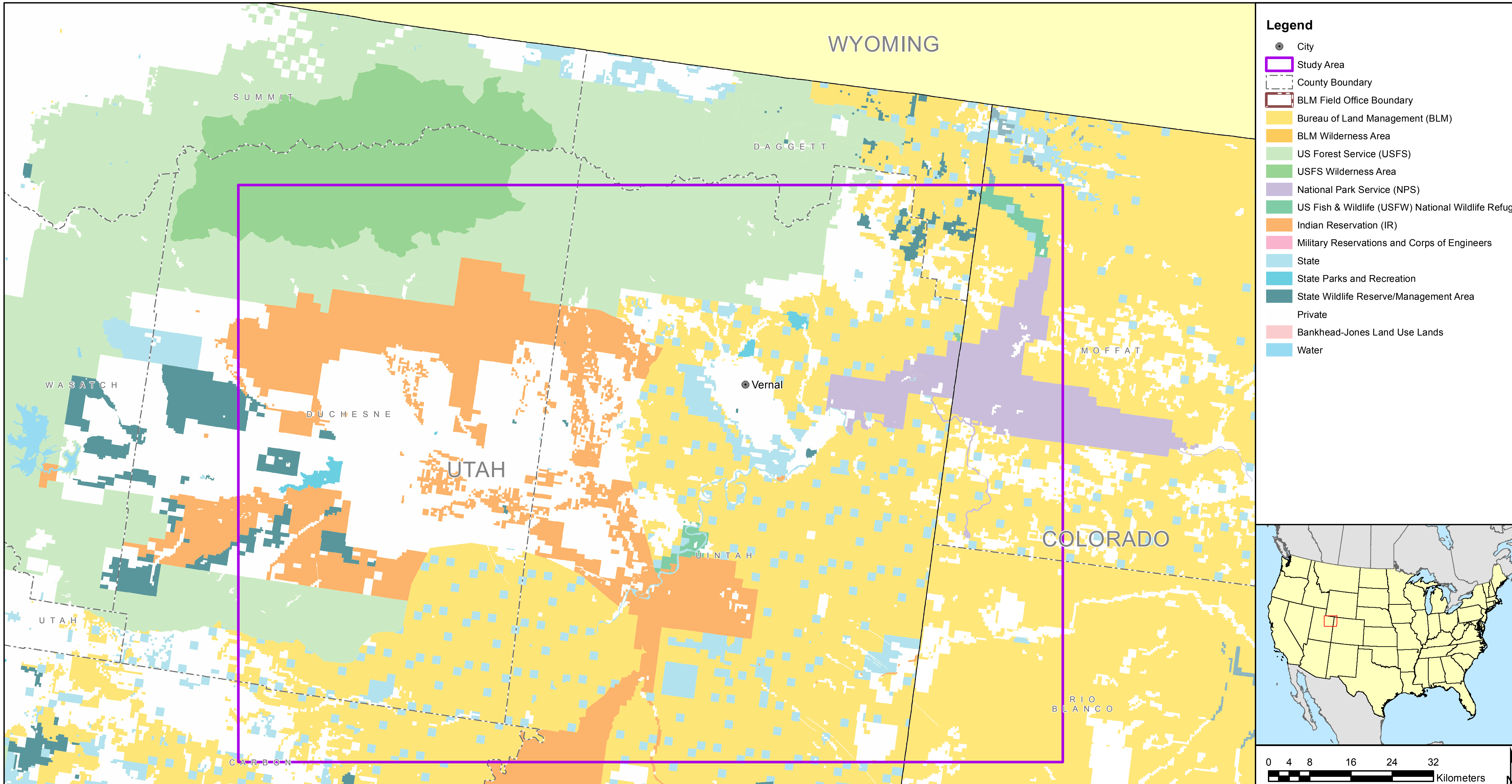
The purpose of a meteorological model performance evaluation (MPE) is to determine whether the WRF results are sufficiently accurate to properly characterize the transport, chemistry, and removal processes in the air quality model and to provide information about potential meteorological biases and errors that may affect model-predicted air quality concentrations. To this end, the ARMS WRF results were evaluated using a variety of quantitative (statistical) and qualitative analyses methods that are presented in the main body of this report.

[1] The exact extent of the Uinta Basin oil and gas reserves is unknown; however, extraction of Uinta Basin reserves is occurring in the northern portions of Grand and Emery counties, which are to the south and outside of the Uinta Basin air quality study area.

1.3 Organization of the WRF MPE

Following this introduction, Chapter 2.0 describes the WRF model, identifies the modeling domains, and provides an overview of the final WRF model configuration. Chapter 3.0 describes the MPE process, statistical metrics, and tools used to conduct the MPE. Chapter 4.0 presents the results of the ARMS WRF simulation and compares the results to observed measurements. Chapter 5.0 summarizes the ARMS WRF configuration performance. In addition to the information provided in the main body of this report, **Appendix A** presents the results of the WRF sensitivity tests that led to the selection of the final WRF configuration.

X:\0Projects\BLM_UTSO_ARMS_60225687\Figures\protocol\DEC_2012\Fig_1-1_ARMS_Study_Area.mxd



ARMS Modeling Project

Figure 1-1
Uinta Basin
Study Area

2.0 WRF Model Configuration

2.1 WRF Model Overview

In order to model air quality conditions for the ARMS Modeling Project, the selected air quality models require gridded meteorological data, which is typically generated by a meteorological model. A meteorological model produces gridded meteorological fields which may include wind speed and direction, atmospheric stability and vertical motion in the atmosphere, sunlight intensity, clouds and precipitation, and heat/moisture fluxes at the surface.

Traditionally, the Penn State University National Center for Atmospheric Research (NCAR) Mesoscale Model Version 5 (MM5) (Grell et al. 1994) has been used to generate gridded meteorological fields for air quality models. However, the MM5 model is being phased out and replaced by the WRF model (NCAR 2009; Skamarock et al. 2008). The scientific algorithms in WRF are developed and reviewed by NCAR, National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Prediction, and other United States (U.S.) government agencies and universities. For this project, the most current version of WRF (version 3.4) that was available at the start of the meteorological modeling effort was used.

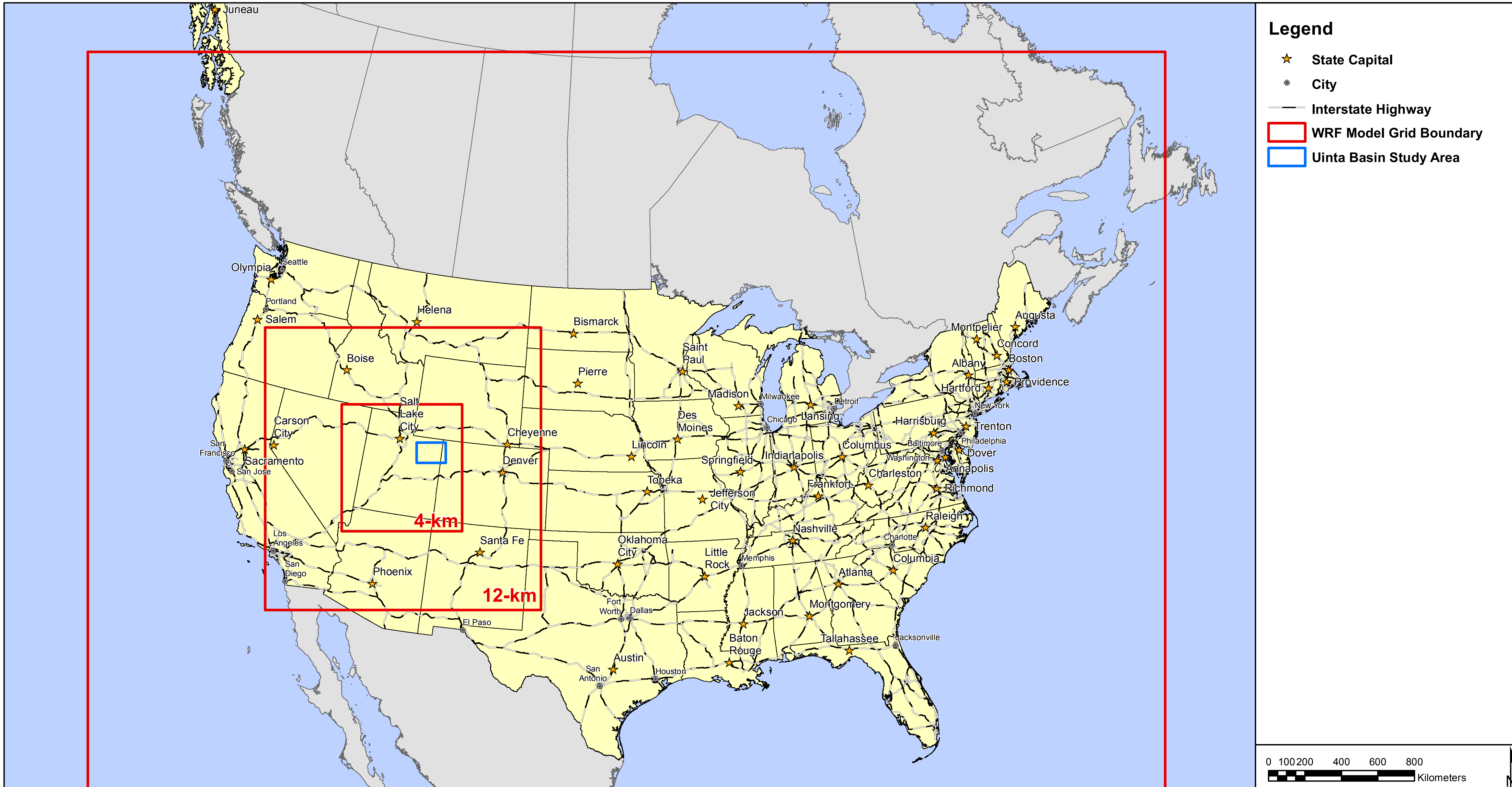
Currently, two dynamic cores (the part of the model code necessary to run the model adiabatically) are available in WRF: the Advanced Research WRF (ARW) (Skamarock et al. 2008) and the non-hydrostatic mesoscale model (Janjic and Gerrity 2001). The ARW core was selected for the ARMS Project because it is capable of performing four-dimensional data assimilation (FDDA). FDDA adjusts the gridded meteorological fields using additional sources of data, such as monitored meteorological data (e.g., observations) or output from other meteorological models.

As described in the Proposed WRF Modeling Approach memorandum (AECOM 2012b), several model simulations were conducted to test the performance of different meteorological model configurations. The configuration tests were conducted for the months of February and July 2010. February and July were selected for evaluation because the meteorological conditions during these months are markedly different and the model configurations may differ accordingly. Conditions in February may include cold-pool stagnation events, while summer month are generally characterized by diurnal patterns of mountain-valley flow and afternoon thunderstorms. The results of these tests are documented in **Appendix A** and led to the final ARMS WRF configuration described in this chapter.

2.2 ARMS WRF Modeling Domains

The ARMS WRF modeling domains include a coarse domain focused on the continental U.S., with a 36-kilometer (km) horizontal grid resolution and two more refined domains with 12-km and 4-km grid resolution focused on the ARMS Study Area. Importantly, as discussed in **Appendix A**, the 12- and 4-km domains were expanded relative to the original domains proposed in the Air Quality Modeling and Assessment Protocol (AECOM 2012a). The domains were expanded during sensitivity testing and these expanded domains were ultimately used for the complete annual ARMS WRF simulation. **Figure 2-1** shows the nested horizontal domains used for the ARMS WRF simulation. The WRF model projection specifications are identical to the Regional Planning Organization unified grid (AECOM 2012a).

The vertical grid is composed of 36 layers from the surface to the top of the model at 50 millibars (approximately 20 km above the surface) with thinner (more) layers near the surface to better resolve the planetary boundary layer (PBL) parameters important for air quality modeling. The vertical layers used for the ARMS modeling are defined in **Table 2-1**. The vertical layers shown in **Table 2-1** are used to model all months.



ARMS Modeling Project

Figure 2-1
WRF 4-km, 12-km, and 36-km Modeling Domains

Table 2-1 Vertical Layer Structure for WRF Annual Simulation

Model Layer	Sigma	Pressure (mb)	Height (meters)	Depth (meters)
36 – top	0.000	50	20,559	4,262
35	0.050	98	16,297	2,527
34	0.100	145	13,770	1,805
33	0.150	193	11,965	1,407
32	0.200	240	10,559	1,185
31	0.250	288	9,374	1,035
30	0.300	335	8,339	931
29	0.350	383	7,408	832
28	0.400	430	6,576	760
27	0.450	478	5,816	701
26	0.500	525	5,115	652
25	0.550	573	4,463	609
24	0.600	620	3,854	572
23	0.650	668	3,282	540
22	0.700	715	2,741	412
21	0.740	753	2,329	298
20	0.770	782	2,032	290
19	0.800	810	1,742	188
18	0.820	829	1,554	185
17	0.840	848	1,369	182
16	0.860	867	1,188	178
15	0.880	886	1,009	175
14	0.900	905	834	87
13	0.910	915	747	85
12	0.920	924	662	85
11	0.930	934	577	85
10	0.940	943	492	83
9	0.950	953	409	83
8	0.960	962	326	83
7	0.970	972	243	81
6	0.980	981	162	41
5	0.985	986	121	41
4	0.990	991	80	20
3	0.9929	993	60	20
2	0.995	995	40	20
1	0.9976	998	20	20
0 – ground	1.000	1,000	0	0

2.3 ARMS WRF Model Configuration

The WRF model was configured and tested for year 2010. For a complete annual air quality simulation for year 2010, the WRF model was initialized on December 29, 2009 12:00 Greenwich Mean Time (GMT) and ran through January 2, 2011 00:00 GMT. The WRF model was re-initialized at the beginning of each 5-day period to reduce error propagation throughout the simulation. Each 5-day period has an initial spin-up period of 12 hours that is excluded from the annual run. This process was used so that the air quality model could be started at 00:00 GMT without including the meteorological model re-initialization period.

As detailed in **Appendix A**, sensitivity tests were conducted for February and July 2010. Different configurations were found to perform better depending on the season evaluated. Based on this, two model configurations were developed: one for winter months and one for non-winter months. In order to determine which months should be modeled with the winter configuration and which months should be modeled with the non-winter configuration, an analysis of snow cover and solar radiation was conducted for the ARMS Study Area. Snow cover and solar radiation parameters were selected since longer nights and a high albedo from snowy surfaces lead to increasing occurrences of low PBL heights and stably stratified conditions in the ARMS Study Area. Based on these two parameters, the months of January, February, March, and December 2010 were modeled with the winter configuration and all other months were modeled with the non-winter configuration.

The final ARMS WRF configuration options common throughout the annual simulation are shown in **Table 2-2**. The final ARMS WRF configurations that vary between the winter and non-winter months are shown in **Table 2-3** and **Table 2-4**, respectively. The configuration differences between the two seasons are the microphysics scheme, the cumulus parameterization, the short-wave radiation scheme, the PBL scheme, the surface layer scheme, and the land surface model.

The WRF output files were created in the standard WRF netCDF output format.

Table 2-2 ARMS WRF Model Configuration Common to All Months[1]

Parameter	36-km Grid[2]	12-km Grid[2]	4-km Grid[2]
Horizontal grids	165 x 129 cells	127 x 130 cells	166 x 175 cells
Lower left X,Y (km)	-2952, -2304	-1980, -756	-1560, -324
Vertical layers	36	36	36
Sea Surface Temperature varies with time	Yes	Yes	Yes
Snow cover	Yes	Yes	Yes
Analysis nudging at the surface	n/a	n/a	n/a
Analysis nudging aloft	u/v/T/Q	u/v/T/Q	u/v/T/Q
Observations nudging	n/a	n/a	u/v/T
Vertical velocity damping	No	No	No
Diffusion option	Simple diffusion	Simple diffusion	Simple diffusion
K option	2d deformation	2d deformation	2d deformation
6th order horizontal diffusion	No	No	No
6th order numerical diffusion non-dimensional rate	0.12	0.12	0.12

Table 2-2 ARMS WRF Model Configuration Common to All Months[1]

Parameter	36-km Grid[2]	12-km Grid[2]	4-km Grid[2]
Base state sea level temperature degrees Kelvin (K)	290	290	290
Upper-level damping	No	No	No
Damping depth from model top (meters)	5,000	5,000	5,000
Damping coefficient	0.05	0.05	0.05
Non-hydrostatic mode	Yes	Yes	Yes
Moisture advection option	Monotonic	Monotonic	Monotonic
Scalar advection option	Monotonic	Monotonic	Monotonic
Total number of rows for specified boundary value nudging	5	5	5
Number of points in specified zone	1	1	1
Number of points in relaxation zone	4	4	4
Specified boundary conditions	True	False	False
Nested boundary conditions	False	True	True

[1] January, February, March, and December are defined as winter for the ARMS Study Area in 2010. All other months are defined as non-winter months.

[2] n/a = physics option will not be applicable to the 4-km, 12-km, or 36-km domain; u = velocity component in the east-west direction; v = velocity component in the north-south direction; T = temperature; and Q = water vapor mixing ratio.

Table 2-3 ARMS Winter WRF Model Configuration[1]

Parameter	36-km Grid	12-km Grid	4-km Grid
Microphysics	Lin et al., scheme	Lin et al., scheme	Lin et al., scheme
Cumulus parameterization	Kain-Fritsch scheme	Kain-Fritsch scheme	None[2]
PBL	Asymmetric Convective Model Version 2 (ACM2) scheme	ACM2 scheme	ACM2 scheme
Surface layer	Pleim-Xiu scheme	Pleim-Xiu scheme	Pleim-Xiu scheme
Land surface model (LSM)	Pleim-Xiu scheme	Pleim-Xiu scheme	Pleim-Xiu scheme
Long-wave radiation	Rapid Radiative Transfer Model (RRTM)	RRTM	RRTM
Short-wave radiation	Dudhia	Dudhia	Dudhia

[1] January, February, March, and December are defined as winter for the ARMS Study Area in 2010.

[2] Typically, cloud convection is well resolved in the 4-km grid without the need for additional cumulus parameterizations.

Table 2-4 ARMS Non-Winter WRF Model Configuration[1]

Parameter	36-km Grid	12-km Grid	4-km Grid
Microphysics	Single moment (6-class)	Single moment (6-class)	Single moment (6-class)
Cumulus parameterization	Grell-Devenyi Ensemble Scheme	Grell-Devenyi Ensemble Scheme	None[2]
PBL	Mellor-Yamada-Janic (MYJ) scheme	MYJ scheme	MYJ scheme
Surface layer	Monin-Obukov (Janic) scheme	Monin-Obukov (Janic) scheme	Monin-Obukov (Janic) scheme
LSM	Unified Noah Land Surface Model	Unified Noah Land Surface Model	Unified Noah Land Surface Model
Long-wave radiation	RRTM	RRTM	RRTM
Short-wave radiation	Goddard	Goddard	Goddard

[1] April, May, June, July, August, September, October and November are defined as "non-winter" conditions for the ARMS Study Area in 2010.

[2] Typically, cloud convection is well resolved in the 4-km grid without the need for additional cumulus parameterizations.

2.4 WRF Model Datasets

A variety of input data are needed to run the WRF model, including topographical information, vegetation cover and land use, initialization data, boundary conditions, water temperature for large bodies of water, and observation data used in data assimilation. High-resolution (e.g., 30-second to 5-minute) topographic, vegetation cover, and land use data were downloaded from University Corporation for Atmospheric Research's (UCAR) WRF User's website (UCAR 2012). In addition, Utah Division of Air Quality provided initial conditions and lateral boundary conditions from North American Model (NAM) data. NAM data is available for 12-km horizontal resolution. The NAM dataset also was used for water temperature inputs.

Observation datasets are required by WRF if FDDA options are invoked in the model configuration settings. For the ARMS WRF modeling, both analysis nudging and observation nudging were used throughout the 2010 annual simulation to nudge the model results. Analysis nudging was applied to wind, temperature, and water vapor mixing ratio on all three domains using the NAM data. The analysis nudging was performed at 6-hour intervals for the 3-dimensional fields above the PBL. Excluding the PBL from analysis nudging removes the undesirable possibility that well-resolved mesoscale forcings, which are important to PBL development, would be damped. If these processes are damped in the PBL, vertical fluxes of momentum, heat, and moisture between the free atmosphere and the surface would be affected.

Observation nudging was applied only to wind and temperature parameters within the 4-km domain. Sensitivity experiments (documented in **Appendix A**) indicated that observation nudging for relative humidity degraded model performance, and therefore observation nudging was not performed on water vapor mixing ratio.

Datasets used in observation nudging were derived from Meteorological Assimilation Data Ingest (MADIS) System coupled with two additional monitoring sites in the Uinta basin (data provided by Utah BLM). Locations of these additional sites are shown in **Table 2-5**. MADIS provides quality-controlled data from numerous synoptic and mesonet networks throughout North America. In addition, MADIS data has good observational coverage of Utah and the Uinta basin. **Figure 2-2** shows surface stations used for the observation nudging in the 4-km domain for temperature and winds. All available surface MADIS observations in the 4-km WRF domain were used for observational nudging and no sites were withheld for model performance evaluation. The primary reason for not withholding sites was to improve model performance in areas with a limited amount of surface observations.

The WRF preprocessing system was used to create input files for the WRF model. The NCAR MADIS2LITTLER utility was used to prepare the raw MADIS dataset for the WRF preprocessing program, OBSGRID. The processed MADIS dataset was merged with the additional Uinta basin data and processed through OBSGRID to create the FDDA input files. OBSGRID performs additional quality assurance tests on the observations, including a buddy check, which compares observations to other neighboring observations. In regions with relatively dense monitoring data and complex terrain, it was found that the buddy check was rejecting valid temperature data; therefore OBSGRID was run without buddy checks to ensure that valid observations were included in the observation nudging processes.

In addition to datasets required to run the WRF model, observation data also is used to evaluate the model performance. Several different sources of data were used for conducting the MPE, including the Techniques Development Laboratory (TDL) U.S. and Canada surface hourly observation data and the Utah BLM data collected at Ouray and Redwash (more information is provided in Chapters 3.0 and 4.0 and **Appendix A**).

Table 2-5 Location of Meteorological Monitoring Sites Provided by Utah BLM

Site Name	Latitude	Longitude
Ouray	40.088	-109.677
Redwash	40.197	-109.353

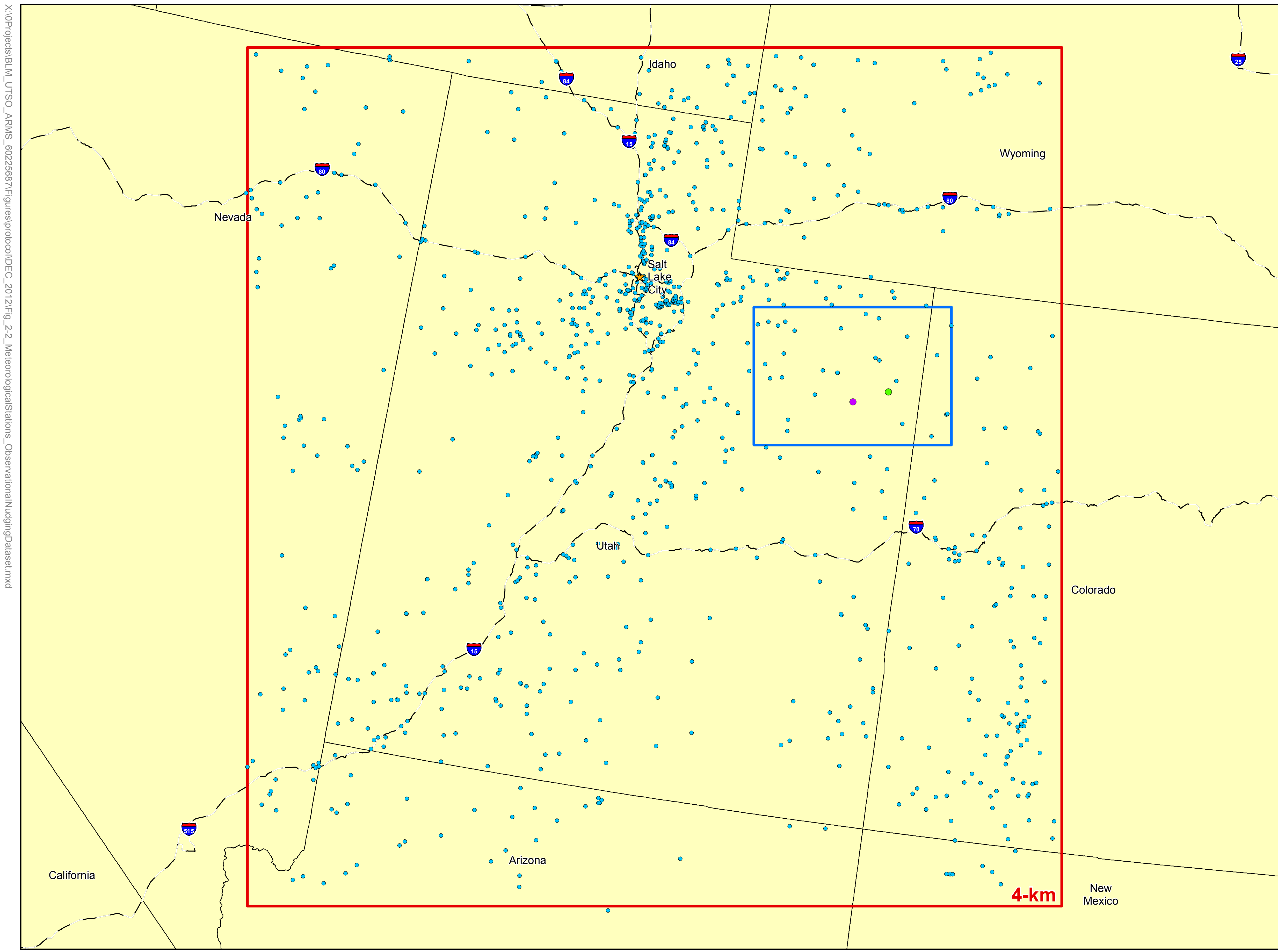


Figure 2-2
Meteorological Stations
Used in the
Observation Nudging Dataset

3.0 Meteorological Model Performance Evaluation Methodology

The generation of gridded, meteorological model data for the ARMS Study Area is the first step in developing a meteorology-emissions-air quality modeling system for use in assessing air quality conditions for the ARMS Modeling Project. The ability of the WRF model to reproduce the meteorological conditions in 2010 is a critical component of the overall modeling system's performance. The primary goal of the meteorological MPE is to determine whether the resulting meteorological fields are sufficiently accurate for the air quality model to properly characterize transport, chemical reactions, and removal processes. An additional benefit of the meteorological MPE is to provide context when evaluating the air quality modeling results in terms of known WRF modeling errors and biases.

The methods developed to evaluate the performance of meteorological models range from qualitative assessments of weather patterns to rigorous quantitative evaluations. To have a reasonable level of confidence in the pertinent meteorological fields, the WRF model should be able to qualitatively reproduce:

- Synoptic scale patterns of wind, temperature, and precipitation fields during frontal passages or blocking events;
- Diurnal variations in PBL height, temperature, and water vapor mixing ratio (mixing ratio);
- Mesoscale circulations such as sea breezes and mountain/drainage circulations; and
- The placement, intensity, and evolution of key weather phenomena.

In addition, quantitative evaluation of model performance often includes graphical and statistical evaluations for critical parameters such as wind speed, wind direction, temperature, and mixing ratio. In order to evaluate the ARMS WRF configuration, both qualitative and quantitative assessments were performed. The following sections provide a discussion of the evaluation methods, data used to conduct the evaluation, and the evaluation tools.

3.1 WRF Evaluation Methods and Datasets

3.1.1 Qualitative Comparisons

The purpose of the qualitative assessment is to establish a first-order acceptance/rejection of WRF's ability to replicate important synoptic and mesoscale patterns in the region of interest. Due to the large volume of data available from an annual WRF simulation, only certain days of model results were qualitatively evaluated to check for obvious model flaws and errors. The periods that are assessed were selected based on air quality events that occurred in the Uinta Basin and will be assessed during the air quality model evaluation. Events include both winter and summer conditions. Winter conditions include stagnant meteorological conditions in the ARMS Study Area. Stagnant conditions can be associated with elevated concentrations of air pollutants. Spatial and vertical plots of model results were compared with observations to assess the model's ability to reproduce important synoptic features; spatial variability of precipitation including precipitation amounts and snow cover; and vertical profiles of wind speed, direction, temperature, and water vapor mixing ratio.

In order to assess WRF's ability to reproduce important synoptic patterns, model results at the surface and 500 mb levels were compared with National Weather Service (NWS) daily weather charts. Also the modeled precipitation fields were compared with the NWS 24-hour cumulative precipitation area and amount charts, comparing predicted and observed occurrences of wide-scale precipitation patterns in the ARMS WRF 36-km and 4-km domains. The snow cover fields from the ARMS WRF 36-km and 4-km domains are compared to snow cover images from NOAA's National Ice Center Interactive Multisensor Snow and Ice Mapping System (NIC IMS) (NOAA 2013).

Radiosonde measurements are recorded twice per day at approximately 120 stations in the continental U.S. Vertical profiles of radiosonde measurements (commonly called skew-T diagrams) (NOAA 2011) were compared to vertical profiles from corresponding grid cells in the ARMS WRF 4-km model domain. Based on proximity to the ARMS Study Area, upper air measurements from SLC and GJT were used to evaluate the vertical performance of the ARMS WRF model. **Figure 3-1** shows the locations of these upper air stations relative to the ARMS WRF 12-km and 4-km model domains.

3.1.2 Quantitative Comparison

Quantitative analyses were conducted by comparing the WRF-predicted meteorological fields to available surface data collected, analyzed, and disseminated by the NWS. The surface station locations used in the MPE analysis are shown in **Figure 3-1**.

For the quantitative comparisons, the METSTAT analysis package was used to determine how well model surface output compared to surface station data for wind speed, wind direction, temperature, and mixing ratio. METSTAT produces various statistical metrics used to evaluate model performance on an hourly and daily basis, as described below.

Root Mean Square Error

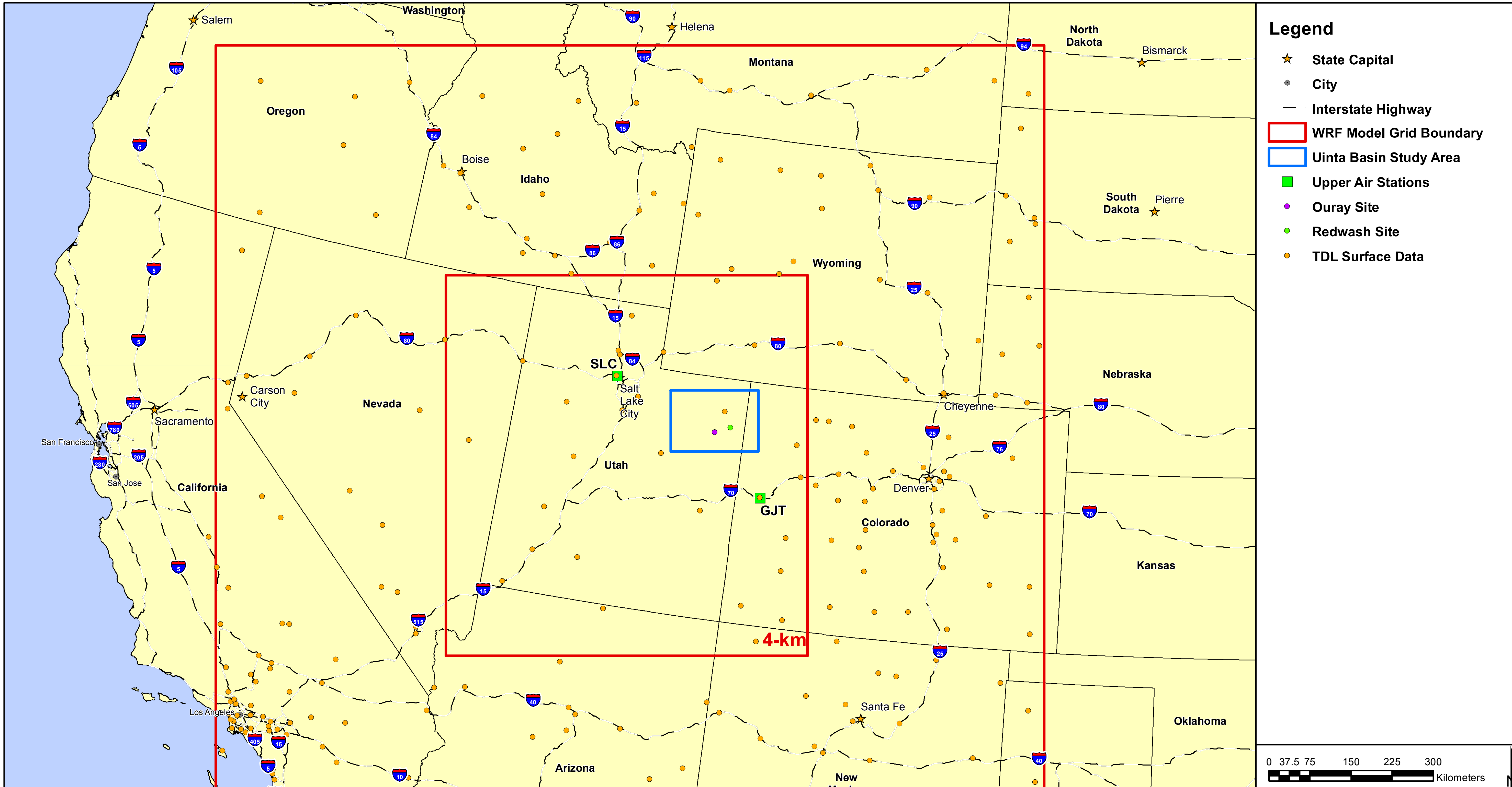
Root mean square error (RMSE) is an overall measure of model performance. The weighting of the difference between predicted and observed values by their squares tends to inflate RMSE, particularly when extreme values are present. RMSE values approaching zero indicate good model performance.

RMSE can be divided into a systematic and unsystematic component by least-squares regression. Differences described by systematic RMSE ($RMSE_S$) can be described by a linear function; therefore, they typically are relatively easy to dampen by a new parameterization of the model. Unsystematic RMSEs can be interpreted as a measure of potential accuracy or noise level (Emery et al. 2001). Good model performance occurs when the systematic difference approaches zero, while the unsystematic difference approaches RMSE. All three measures are computed in the METSTAT program; however, only the RMSE results are presented in the report.

The RMSE is calculated as the square root of the mean squared difference in predicted-observed pairings with valid data within a given analysis region and for a given time period (hourly or daily), as shown in the equation below:

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

where P_j^i is the individual predicted quantity at site *i* and time *j, 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*).



ARMS Modeling Project

Figure 3-1
Meteorological Stations
Used to Evaluate
Model Performance

X:\0Projects\BLM_UTSO_ARMS_60225687\Figures\protocol\DEC 2012\Fig_3-1_MeteorologicalStationsUsed_EvaluateModelPerformance.mxd

Bias

The bias error (bias) is the degree of correspondence between the mean predicted and the mean observed values, with lower absolute values indicative better model performance. Values less than zero indicate model under-prediction. The bias is calculated as the mean difference in predicted-observed pairings with valid data within a given analysis region and for a given time period (hourly or daily), as shown in the equation below:

$$Bias = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\left(P_j^i - O_j^i\right)$$

Gross Error

The gross error, or mean absolute error, is the mean of the absolute value of the residuals from a fitted statistical model. Lower numbers indicate better model performance. The gross error is calculated as the mean absolute difference in predicted-observed pairings, with valid data within a given analysis region and a given time period (hourly or daily), as shown in the equation below:

$$GrossError = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\left|P_j^i - O_j^i\right|$$

Index of Agreement

Index of Agreement (IOA) is a relative measure of the degree to which predictions are error-free. The denominator accounts for the model's deviation from the mean of the observations, as well as to the observations deviation from their mean. It does not provide information regarding systematic and unsystematic errors. The IOA approaches one when model performance is best. The IOA is calculated by condensing all of the differences between model estimates and observations within a given analysis region and for a given time period (hourly and daily) into one statistical quantity, as shown in the equation below:

$$IOA = 1 - \left[\frac{IJ \cdot RMSE^2}{\sum_{j=1}^{J}\sum_{i=1}^{I}\left|P_j^i - M_o\right| + \left|O_j^i - M_o\right|}\right]$$

where mean observation (M_o) was calculated from all sites with valid data within a given analysis region and for a given time period (hourly or daily):

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

It should be noted that not all statistics are appropriate metrics for evaluating every meteorological parameter. For example, RMSE is an appropriate statistical metric for wind speed, while gross error is more appropriate for reporting the error associated with wind direction, temperature, and mixing ratio.

Domain-wide statistics for each of the 36-km, 12-km, and 4-km domains were generated and summarized in this report. In addition to domain-wide statistics, statistical evaluations were performed for Uinta Basin Study Area using the 4-km and 12-km grid resolution results.

3.2 Statistical Benchmarks Used to Evaluate Performance

Statistics were computed for the annual simulation and compared against a set of statistical benchmarks derived by Tesche et al. (2002) for establishing acceptable model performance (**Table 3-1**). These benchmarks were developed based on evaluation of approximately 30 episodic meteorological simulations used for air quality modeling over a 5- to 10-year period. Since the WRF model was not operational at the time these benchmarks were established, the benchmarks may not be appropriate for the WRF model. In addition to comparing WRF results to the benchmarks developed by Tesche et al. (2002), additional benchmarks have been developed specifically for areas with complex terrain in recognition of the difficulties associated with modeling these types of regions (Kemball-Cook, *et al.*, 2005).

Table 3-1 Statistical Benchmarks for Evaluating Meteorological Model Performance

Statistic		Wind Speed (meters per second [m/s])	Wind Direction (degrees from true north [deg])	Temperature (K)	Water Vapor Mixing Ratio (grams per kilogram [g/kg])
RMSE	Tesche et al. (2002)	≤ 2	--[1]	--[1]	--[1]
	Complex Terrain	≤ 2.5	--[1]	--[1]	--[1]
Bias	Tesche et al. (2002)	±0.5	±10	±0.5	±1
	Complex Terrain	--[1]	--[1]	± 2	± 1
Gross error	Tesche et al. (2002)	--[1]	≤ 30	≤ 2	≤ 2
	Complex Terrain	--[1]	≤ 55	≤ 3.5	≤ 2
IOA	Tesche et al. (2002)	≥ 0.6	--[1]	≥ 0.8	≥ 0.6
	Complex Terrain	--[1]	--[1]	--[1]	--[1]

[1] Benchmark not appropriate for this meteorological parameter and statistical metric.

The benchmarks shown in **Table 3-1** will be used in this report for informational purposes. Statistical benchmarks are not used as an acceptance and/or rejection criteria of the ARMS WRF simulation. Rather, they put the ARMS WRF performance into perspective relative to previous simulations performed in the U.S. In addition, by comparing model results to these benchmarks, it allows for identification of potential problems in the WRF results. Results exceeding the benchmarks are reviewed with more scrutiny.

3.3 Evaluation Tools

The METSTAT program spatially and temporally pairs the WRF model predictions with observed data for selected periods and locations. Domain-wide analyses were performed for the three ARMS WRF domains; however, analysis of the 36-km domain was restricted to the area contained within the 12-km domain. The analysis of the 36-km domain is restricted since model performance outside of the 12-km domain is not of interest for the study, provided that the model performs adequately in the 4-km domain.

The METSTAT program calculates statistics for the parameters identified below.

- Wind Speed, Temperature, and Water Vapor Mixing Ratio:
  - Mean observed
  - Mean predicted
  - RMSE
  - Bias
  - Gross error
  - IOA
- Wind Direction:
  - Mean observed
  - Mean predicted
  - Bias
  - Gross error

The RMSE and IOA typically are not used to quantify error for wind direction and, thus, are not calculated by the program.

Daily statistics were plotted as bar charts to show daily performance over an episode, and daily results from multiple WRF runs can be plotted together to provide an inter-comparison of performance. Hourly statistics were plotted as time series summaries, to show the diurnal variation of model performance.

In addition to tabular summaries and time series plots, the performance statistics are presented using Bakergrams to assess the seasonal variability in performance. Bakergrams provide a graphical representation of statistical metrics over time. The daily averages of a particular statistical metric (e.g., bias, gross error, RMSE, or IOA) are shown for an entire year in a Bakergram. In the annual Bakergram, each column represents 1 month, with the month of January at the far left and December at the far right. Each row represents a daily average with day 1 at the top and day 31 at the bottom. For the months with less than 31 days, the daily average value is treated as zero. The domain-wide daily average bias, gross error, RMSE, and IOA values are shown by different colors. Bias data shown in grey represents data that are within the benchmarks developed by Tesche et al. (2002). Bakergrams visually consolidate information in a way that makes it relatively easy to assess performance changes over an annual simulation.

4.0 Evaluation of WRF Model Performance

The accuracy of the WRF simulation will directly impact the air quality model's ability to represent transport (via advection and diffusion), removal (via deposition), and chemical reactions that are a function of meteorological conditions such as humidity and temperature. In order to better understand the WRF modeling, both qualitative and quantitative evaluations were conducted depending on the parameter being analyzed and the type of data available. The evaluations included:

- Qualitative assessment of select periods including the surface patterns, upper-air patterns, precipitation, and snow cover; and
- Quantitative assessment of the surface layer, including:
  - Assessment of the model performance seasonally and annually; and
  - Assessment of the model performance domain-wide and for the Uinta Basin Study Area.

The qualitative analyses provide a detailed assessment of the surface and upper-air model performance including pressure patterns, atmospheric profiles, and precipitation. These analyses help to frame the quantitative, statistical analyses by reviewing the model's ability to reproduce synoptic and mesoscale patterns. Proper model placement and timing of systems provides confidence in the WRF simulation and supports conclusions that the model is reproducing observed events for the correct reasons. The nature of qualitative analyses requires intensive review, which limits the amount of analyses that may be conducted. As a result, only select periods are qualitatively reviewed and the results are assumed to be representative of other time periods.

The quantitative analysis provides an objective measure of model performance as compared to observations. These analyses provide detailed information about model performance at a specific location, geographical region, and/or modeling domain and can be provided for any averaging period such as hourly, daily, monthly, seasonally or annually. The analysis presented in this chapter focuses on daily, seasonal, and annual model performance for specific model domains and geographic regions.

4.1 Qualitative Model Performance Evaluation

One of the goals of conducting the WRF model performance evaluation is to examine whether the meteorological fields from the ARMS WRF simulation properly characterized the large-scale weather patterns in the western U.S. An evaluation of the large-scale (or synoptic) weather patterns provides insight into model performance regarding the placement, timing and progression of systems. Proper model characterization of synoptic patterns provides confidence that the model is working properly and that the model-generated meteorological fields represent observed physical phenomena.

To qualitatively assess how well the ARMS WRF configuration performed synoptically, model outputs were compared to NWS surface weather maps, 500-mb height contour charts, 24-hour cumulative precipitation amounts, and the snow cover. The comparison to the NWS surface maps provides insight into the model's placement of systems. The comparison to the 500-mb heights provides insight into the placement and timing of systems in the upper-levels of the model. The comparison of the precipitation plots provides insight into the overall physical and dynamical performance of the model since precipitation occurs for many different and complex reasons. In addition, the presence of snow is important to assess as the high albedo of snow surfaces impacts the amount of ultraviolet (UV) radiation available for photochemistry, To qualitatively assess how well the ARMS WRF snow cover reproduces actual conditions in 2010, model outputs were compared with Northern Hemisphere snow coverage maps produced by NIC IMS. Following these surface analyses, upper-air data at select locations in the 4-km domain were compared with the model's vertical profiles of key parameters.

These qualitative comparisons were conducted for four different episodes:

- January 8-January 23, 2010 (elevated ozone and PM with an aerodynamic diameter less than or equal to 2.5 microns [$PM_{2.5}$]);
- February 21-March 8, 2010 (elevated ozone);
- August 19-August 29, 2010 (elevated ozone and $PM_{2.5}$); and
- September 27-October 5, 2010 (elevated $PM_{2.5}$).

These episodes were selected based on air quality conditions in the Uinta Basin and will be assessed in detail during the evaluation of the air quality model performance. To support and inform the review of the air quality model performance, an assessment of the meteorological conditions during these events is provided in the following sections.

4.1.1 January 8 to 23, 2010 Air Quality Event

4.1.1.1 Synoptic Conditions

Figures 4-1, **4-2**, and **4-3** compare WRF model results to observations for surface weather, 500-mb height contours, and 24-hour precipitation amounts, respectively. The surface charts and 500-mb height contours are shown for 1200 GMT, which corresponds with 5:00 a.m. Mountain Standard Time (MST). The precipitation plots show cumulative amounts over a 24-hour period. For **Figures 4-1**, **4-2**, and **4-3** the charts on the left side are from the NWS; those on the right are from the ARMS WRF simulation. Each figure shows three days in the middle of the January 8-23 event: January 12, 16, and 20, 2010.

During January 8-18, a high pressure system dominated much of the western and central U.S. A series of weak to moderate frontal systems migrated from Gulf of Alaska across southern British Columbia, Midwest U.S., and into eastern U.S. Throughout this period, the study area was constantly under the influence of a high pressure system with a center pressure of 1032 mb. This resulted in dry conditions with light winds and no clouds. Both the surface weather chart and WRF predictions from January 12 (**Figure 4-1a**) shows two high pressure systems over much of U.S. divided by a trough. The strengths and locations of the high pressure systems and the trough are similar between observations and the WRF model.

From January 14-18, the high pressure system over much of the southwestern U.S. slowly weakened but the high pressure center continued to remain over the state of Utah. Dry conditions with no rain continue to persist in the study area. The January 16 (**Figure 4-1b**) surface weather charts and WRF predictions show good agreement for the location of the high pressure center over the study area, as well as the frontal systems over the Northwest Pacific and Canada.

From January 20-23, low pressure systems passed through Arizona, New Mexico, and southern parts of Utah leading to increased precipitation over the southwestern U.S. The January 20 surface plots (**Figure 4-1c**) show good agreement with the location of the frontal systems over the Pacific and southeastern U.S.

On January 12, 2010, the mid-level (500-mb) pattern diagnosed by NWS from observations shows a ridge of high pressure that covered the western U.S. and a trough of low pressure that covered the eastern U.S. (**Figure 4-2a**). The location and strength (5,760 meters) of the western ridge, eastern trough and other upper level features (e.g., the low over Mexico and parts of Texas) from the ARMS WRF simulation compares favorably with the NWS observations. As the ridge slowly moved eastward, a low develops over Texas and Oklahoma on January 16 (**Figure 4-2b**). The ARMS WRF simulation was able to reproduce the NWS observations. Additionally, the ARMS WRF wind speeds and directions also agree well with the NWS at the 500-mb level. On January 20, 2010, the NWS 500-mb pattern shows a weak trough that covered the western U.S. and a low pressure that covered the northeastern U.S. (**Figure 4-2c**). Again, the ARMS WRF simulation was able to reproduce the NWS observations.

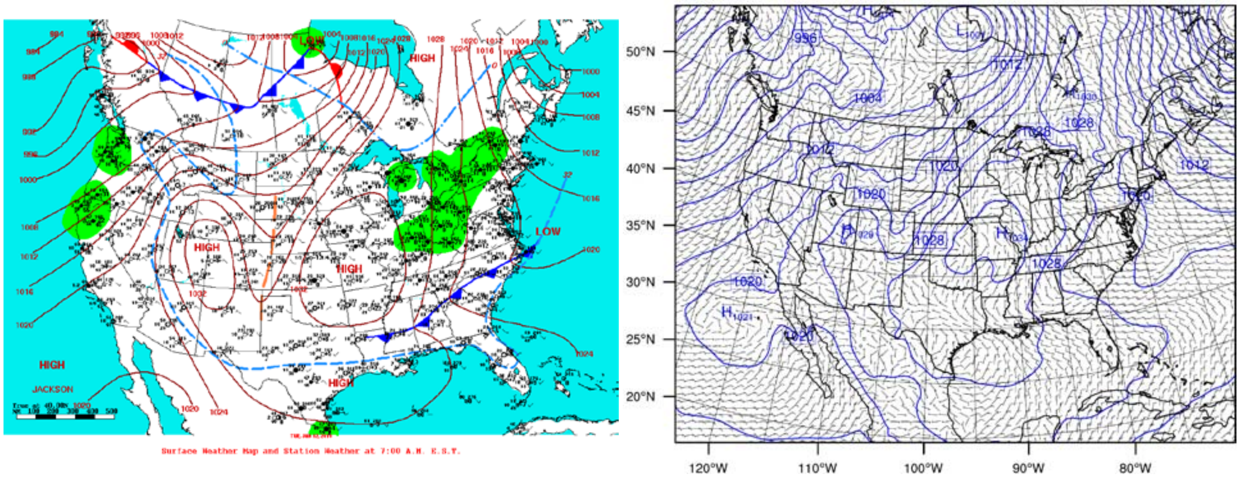
In terms of precipitation in the 36-km domain, on January 12, 2010 (**Figure 4-3a**), the NWS reported precipitation over the Pacific Northwest and the Great Lakes. The ARMS WRF simulation predicted slightly more precipitation in Canada and Atlantic Ocean. Otherwise, the overall ARMS WRF precipitation area matched with NWS observations. On January 16 (**Figure 4-3b**), the ARMS WRF precipitation area also matched up well with NWS observations over the entire U.S. On January 20 (**Figure 4-3c**), the NWS reported precipitation over much of the western U.S., parts of central U.S., and northeastern U.S. The ARMS WRF precipitation areas are consistent with the NWS observations.

In the 4-km domain, the NWS reported very little precipitation throughout the state of Utah for January 12 (**Figure 4-4a**). The ARMS WRF simulation predicted a small amount of precipitation northwest of the Great Salt Lake and in the Uinta Basin that did not actually occur. The one-day accumulated precipitation is between 0.01 to 0.1 inches in the ARMS WRF simulation. On January 16, 2010, the precipitation pattern reported by NWS and ARMS WRF simulation (**Figure 4-4b**) is similar to January 12. On January 30, 2010, NWS reported precipitation over much of the state of Utah (**Figure 4-4c**) which is well matched by the WRF simulation. In the Uinta Basin, the NWS reported one-day precipitation amounts of less than 0.01 inches while the ARMS WRF simulation reported precipitation amounts of 0.01-0.1 inches. During this period, the ARMS WRF 4-km domain precipitation area is in agreement with NWS observations; however, the WRF simulation predicts slightly more total precipitation over Utah than the NWS.

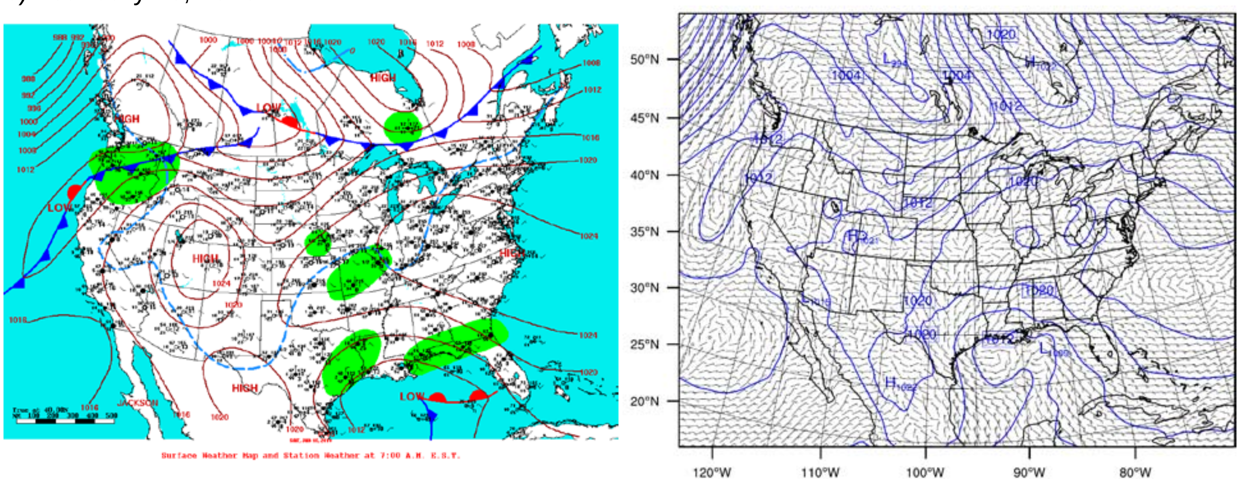
In **Figure 4-5**, the WRF 36-km and 4-km domain snow cover fields are compared with NIC IMS observation data on January 16, 2010. Snow cover data were reviewed for other days during the period January 8 through 23, but the results were similar to January 16 and not reproduced here. In **Figure 4-5**, the chart on the left side is from the NIC IMS (NOAA 2013). The chart in the middle is from the ARMS WRF 36-km simulation and one on the right is from the 4-km simulation. In the NIC IMS plots, areas with snow cover are shown in white, ice is shown in yellow, and uncovered land surfaces is shown in green. In the WRF plots, snow and ice are shown in white and uncovered surfaces in green. The WRF snow cover field reproduces the observed snow cover well in both the 36-km domain and the 4-km domain during this period.

NWS Observations ARMS WRF 36-km

a) January 12, 2010



b) January 16, 2010



c) January 30, 2010

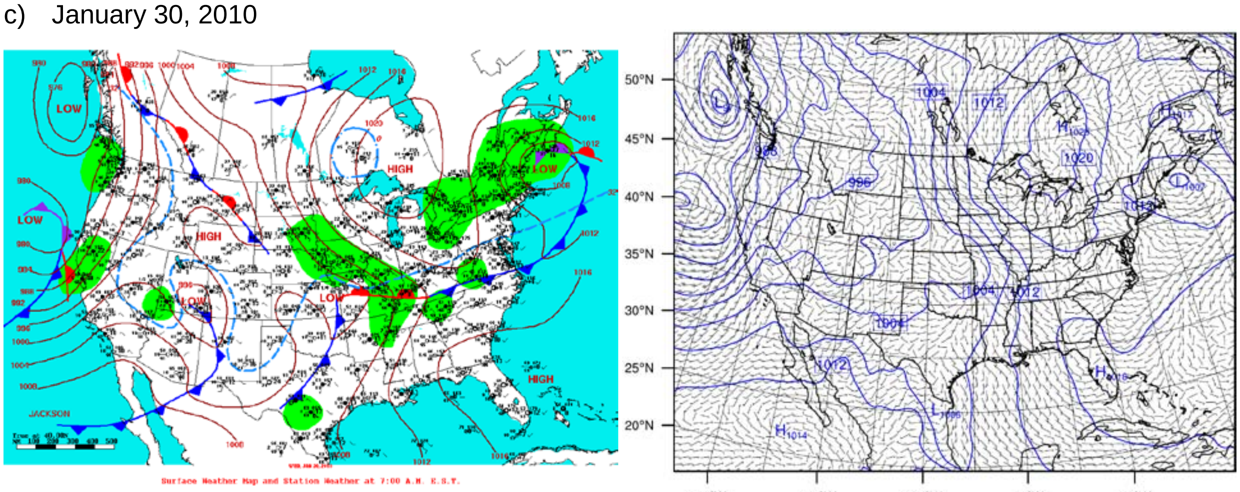


Figure 4-1 Comparison of Surface Weather Charts for January 12, 16, and 20, 2010

NWS Observations ARMS WRF 36-km

a) January 12, 2010

b) January 16, 2010

c) January 30, 2010

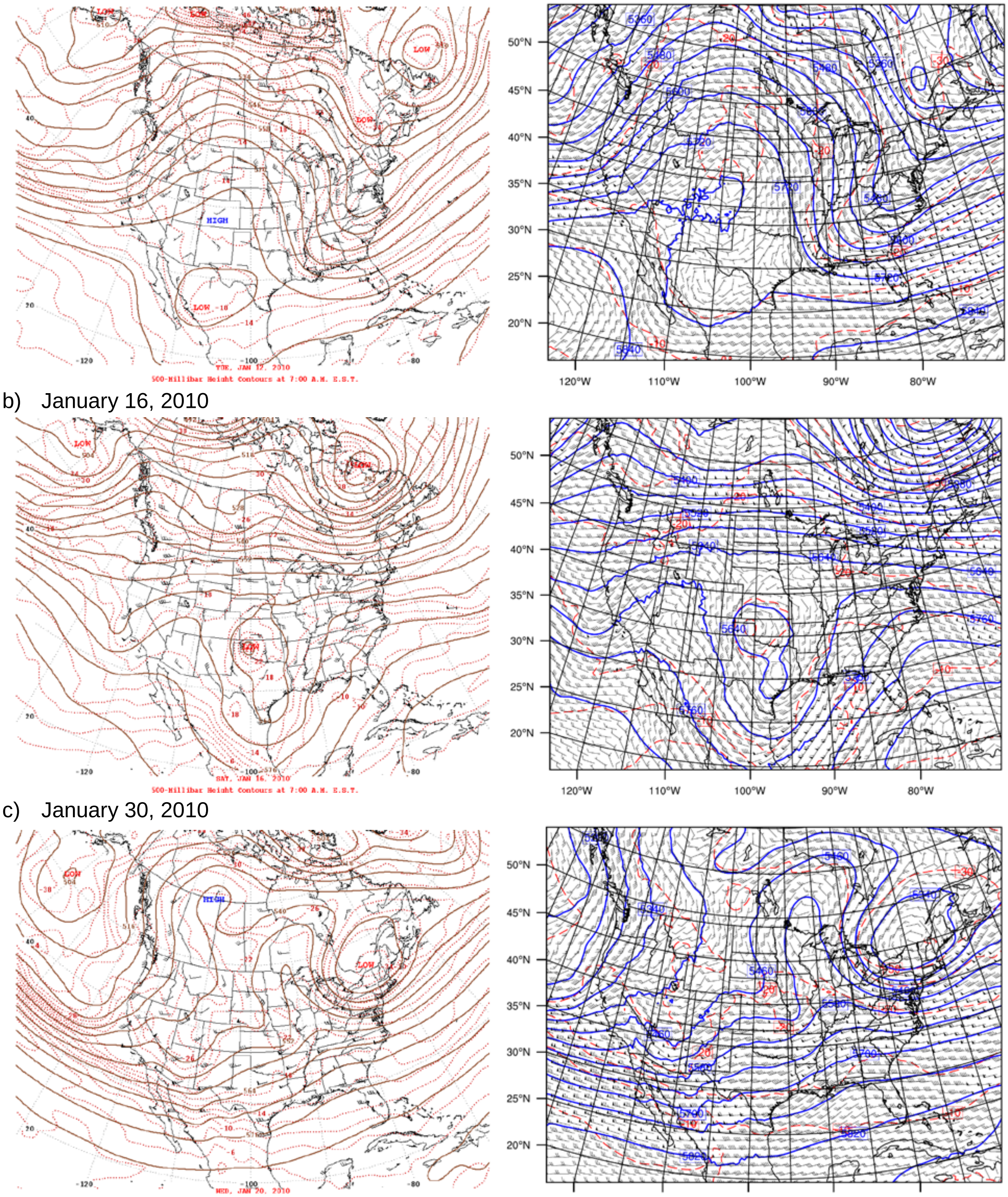
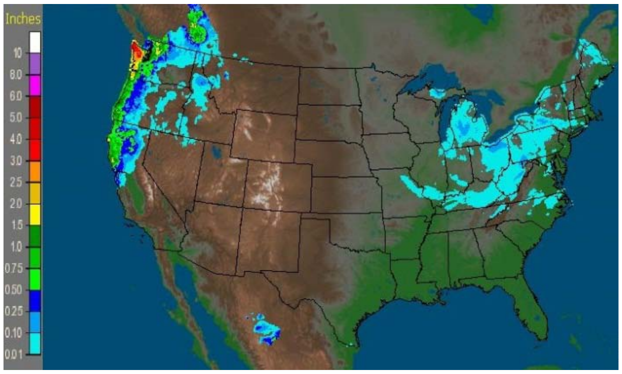


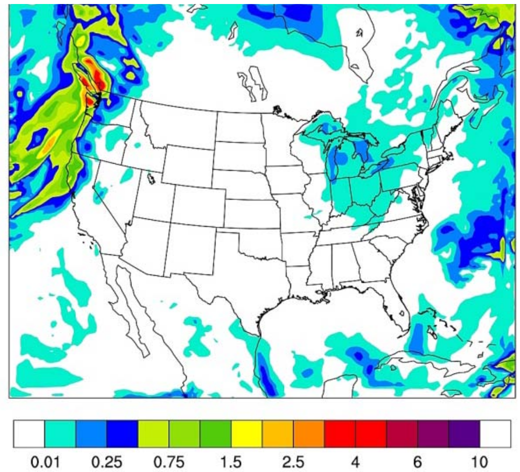
Figure 4-2 Comparison of 500-mb Height Contour Charts for January 12, 16, and 20, 2010

NWS Observations

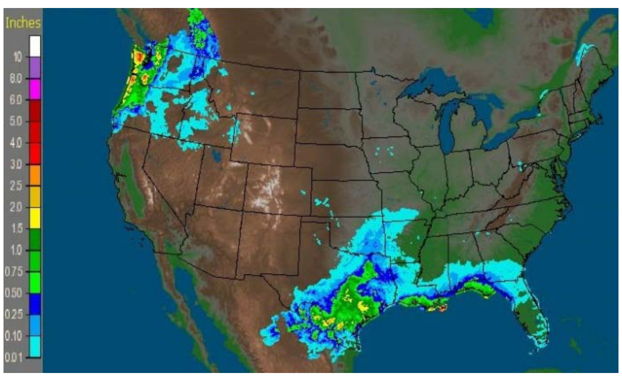
ARMS WRF 36-km

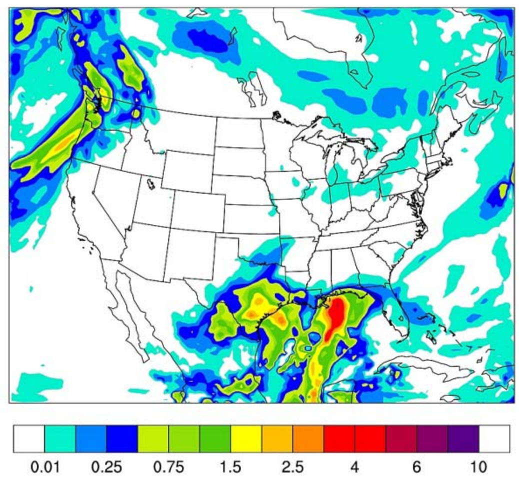
a) January 12, 2010



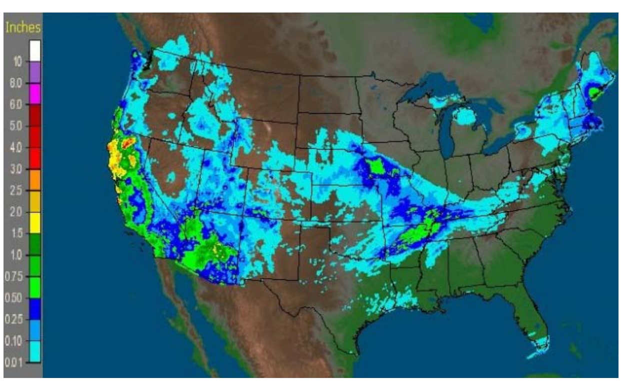


b) January 16, 2010





c) January 30, 2010



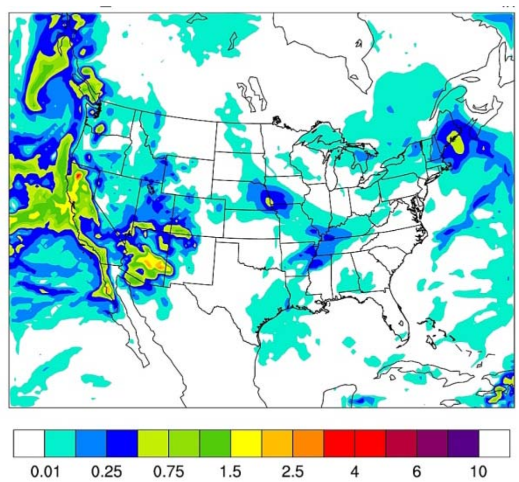
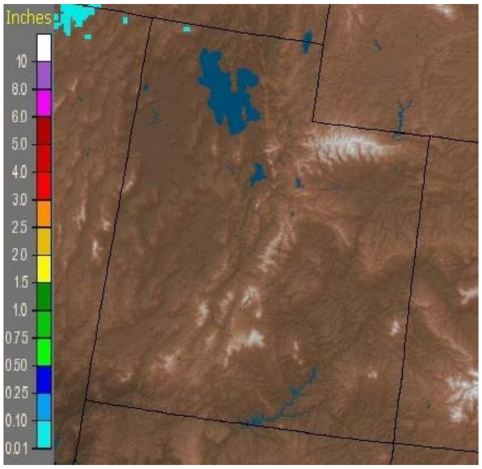


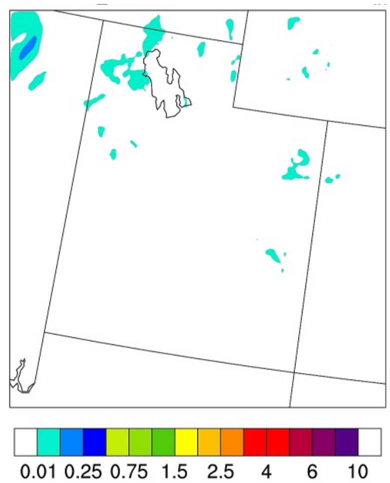
Figure 4-3 Comparison of 36-km Domain Precipitation Amounts for January 12, 16, and 20, 2010

NWS Observations

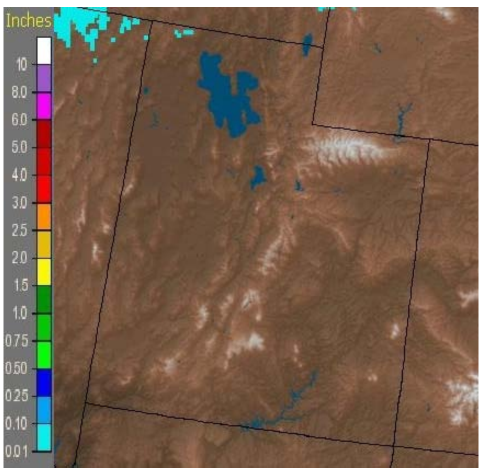
ARMS WRF 4-km

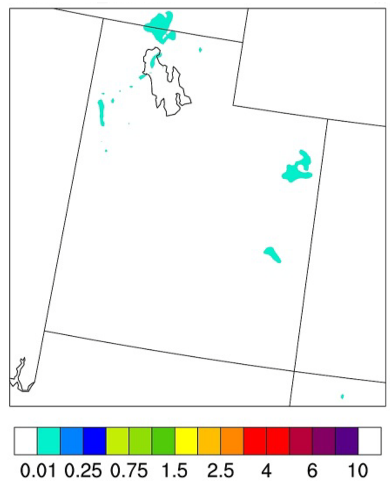
a) January 12, 2010



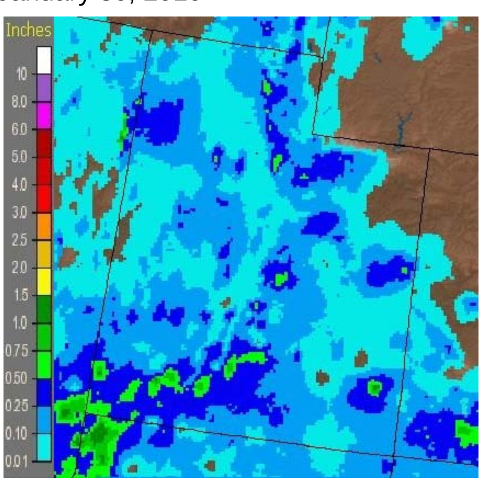


b) January 16, 2010





c) January 30, 2010



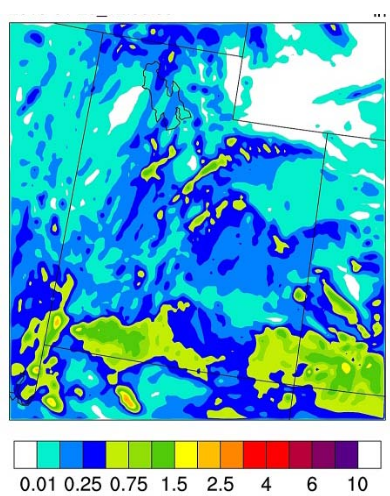


Figure 4-4 Comparison of 4-km Domain Precipitation Amounts for January 12, 16, and 20, 2010

4.1.1.2 Conditions Aloft

The NWS collects upper-air data twice per day at 0000 and 1200 Greenwich Mean Time (GMT) (or 5:00 p.m. and 5:00 a.m. local standard time, respectively) from locations across the U.S. The data is typically referred to as upper-air soundings, and the information collected at each location includes vertical profiles of temperature, moisture, and wind speed and direction. From this data additional information can be obtained such as atmospheric stability and mixing heights. For analyzing WRF performance, the morning soundings from the Salt Lake City, Utah, (SLC) and Grand Junction, Colorado (GJT) stations were compared with the WRF-simulated vertical structure of the atmosphere. The locations of these stations are shown relative to the ARMS WRF model domains in **Figure 3-1**.

Morning soundings at SLC and GJT were compared with the WRF model simulated vertical structure of the atmosphere at the grid cell that contains the SLC station and the GJT station in **Figures 4-6a** and **4-6b**, respectively. The vertical model performance is analyzed for the same three days selected for the surface layers (January 12, 16, and 20, 2010).

In general, both the temperature (solid lines) and dew point temperature (dashed lines) profiles from the ARMS WRF simulations followed the profiles from the observed soundings. The temperature profile performed better than the dew point temperature profile. The model had difficulty replicating the sharp changes in the dew point temperature that is related to the water vapor mixing ratio. The model generally was able to simulate the vertical variability in wind direction and speed. Importantly, the model was capable of reproducing the height of the inversions as well as the strengths.

4.1.2 February 21 to March 8 Air Quality Event

4.1.2.1 Synoptic Conditions

Figures 4-7, **4-8**, and **4-9** compare WRF model results to observations for surface weather, 500-mb height contours, and 24-hour precipitation amounts, respectively. The surface charts and 500-mb height contours are shown for 1200 GMT, which corresponds with 5:00 a.m. MST. The precipitation plots show cumulative amounts over a 24-hour period. The charts on the left side are from the NWS; those on the right are from the ARMS WRF simulation. Each figure shows three days in the middle of the February 21 – March 8 event: February 28, March 2, and March 5, 2010.

During February 21 – March 8, 2010, a series of low pressure system tracked through western U.S. In between the low pressure systems, weak high pressure systems covered much of the western U.S. In the study area, 24-hr accumulated precipitation was observed for many days during this episode.

On February 28 (**Figure 4-7a**), a low pressure system with associated precipitation moved over parts of southwest U.S. By March 2 (**Figure 4-7b**), a large slow moving high pressure system formed over central U.S. extending into parts of southwest U.S. By March 5 (**Figure 4-7c**), another low pressure from the Pacific Ocean tracked through the southwest U.S. into the central U.S.

The February 28 NWS 500-mb height chart (**Figure 4-8a**) shows a low pressure system (5,820 meters) persistent over the southwest U.S. A low pressure system with a 500-mb height of 5,580 meters was located over northeast U.S. The ARMS WRF simulation was able to reproduce the NWS observations. In addition, the ARMS WRF simulation was able to produce wind speeds and directions comparable with the NWS at the 500-mb level. On March 2 (**Figure 4-8b**), plots show evidence of a strong high pressure ridge slowly developing, which moved into the central plain states. The location and strength of the high pressure system, the jet stream, and low pressure system were reproduced successfully by the ARMS WRF simulation. On March 5, the ARMS WRF simulation was able to reproduce the location and strengths of low pressure over the study area and the Northwest Pacific. For all three days analyzed, the ARMS WRF simulation was able to reproduce the 500-mb flow patterns and strengths of the NWS observations.

NIC IMS Data ARMS WRF 36-km ARMS WRF 4-km

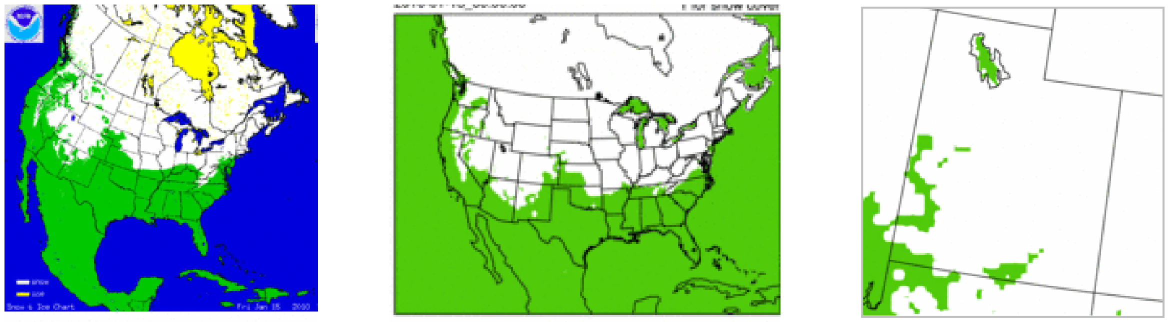
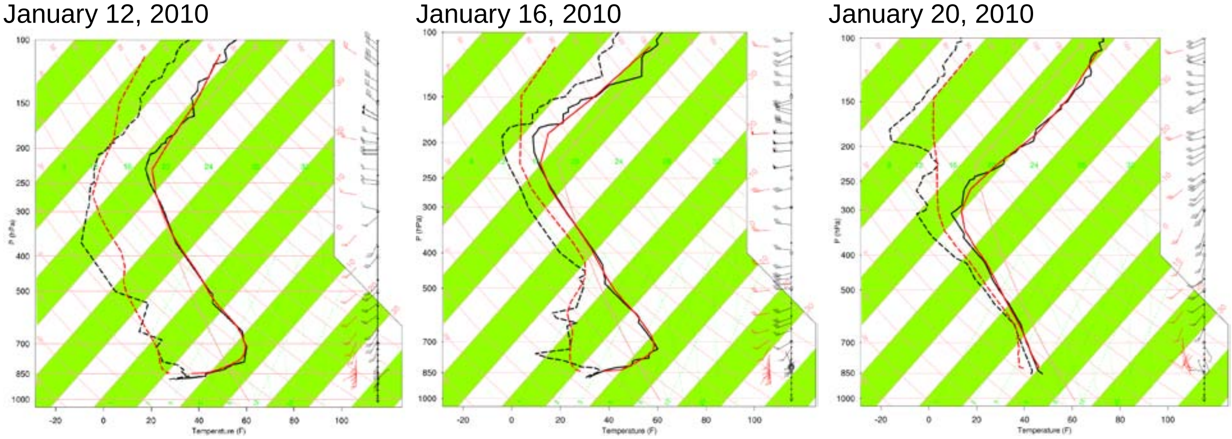


Figure 4-5 Comparison of Snow Cover on January 16, 2010

a) Salt Lake City (SLC)

January 12, 2010 January 16, 2010 January 20, 2010



b) Grand Junction (GJT)

January 12, 2010 January 16, 2010 January 20, 2010

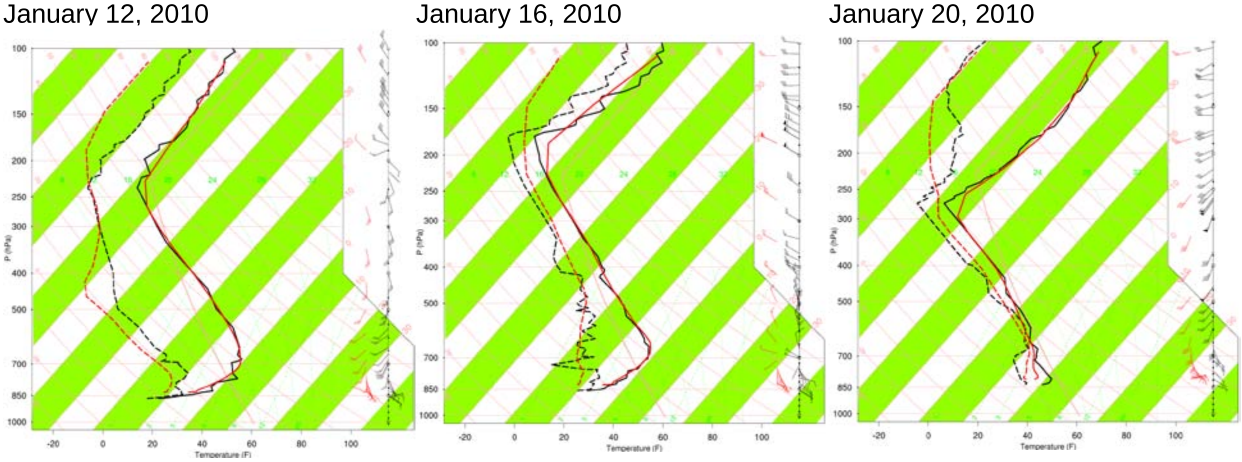
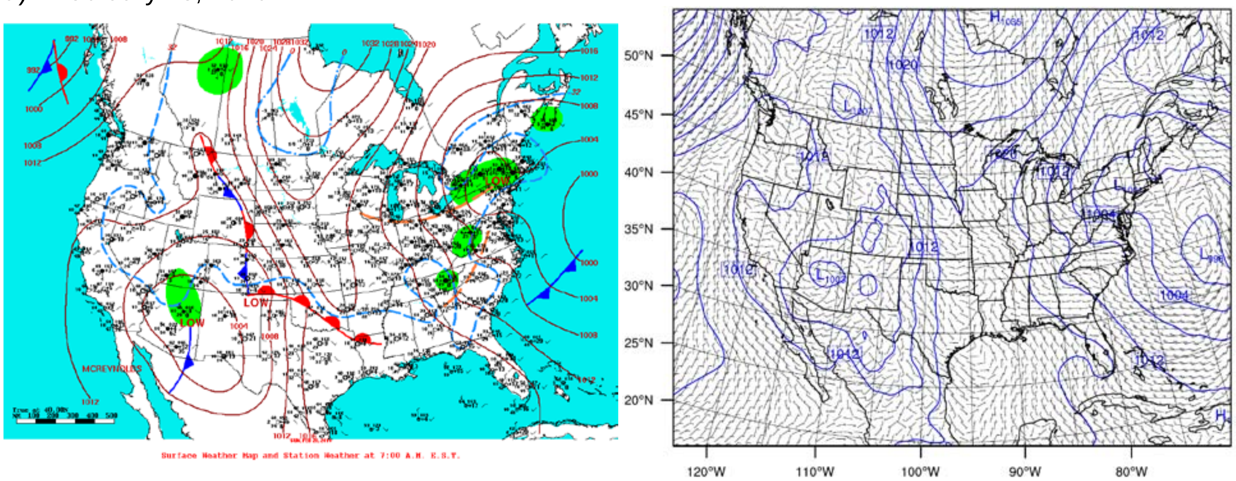


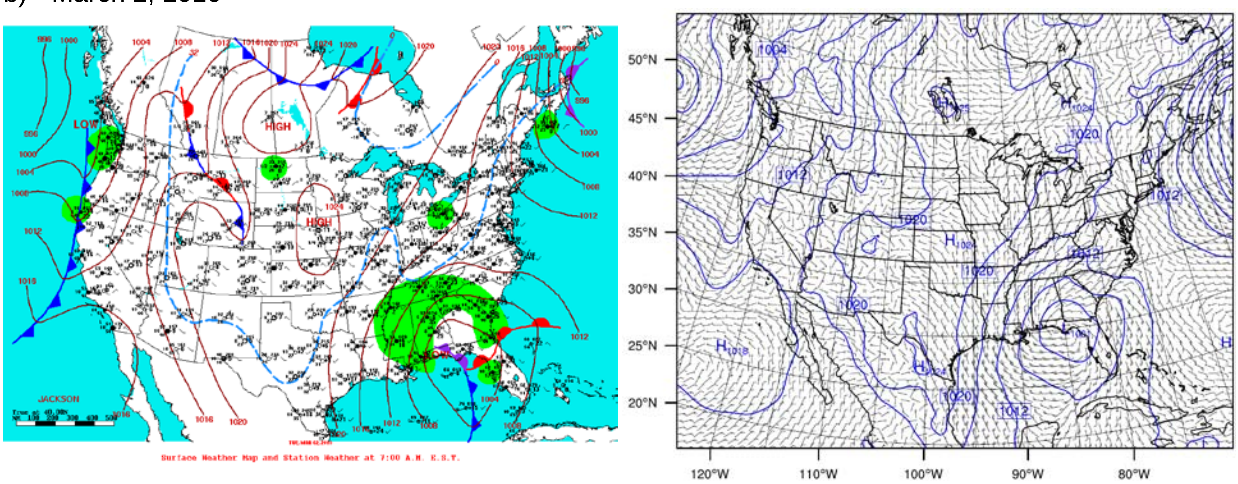
Figure 4-6 Comparison of Vertical Profiles at SLC and GJT for January 12, 16, and 20, 2010

NWS Observations ARMS WRF 36-km

a) February 28, 2010



b) March 2, 2010



c) March 5, 2010

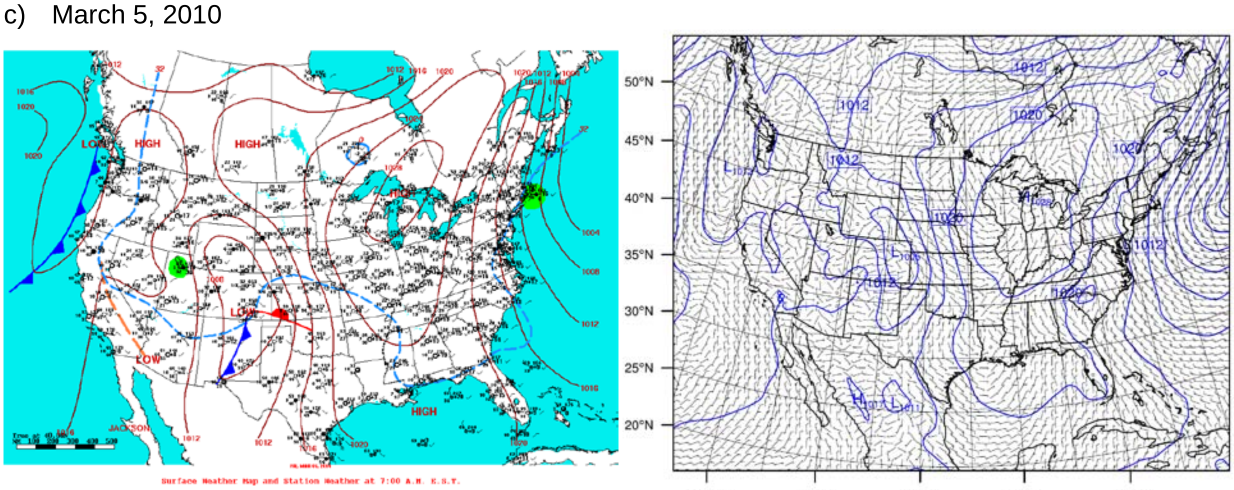
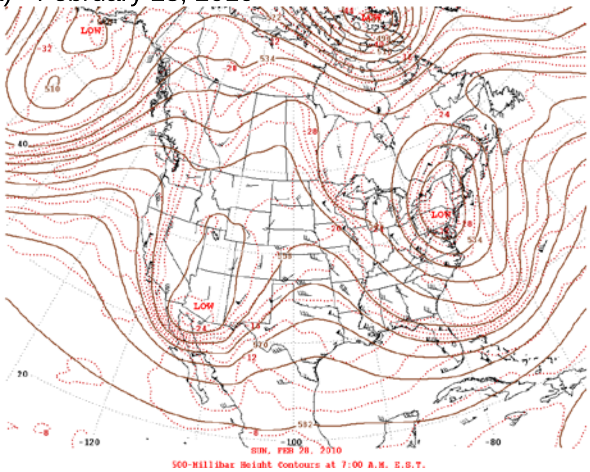


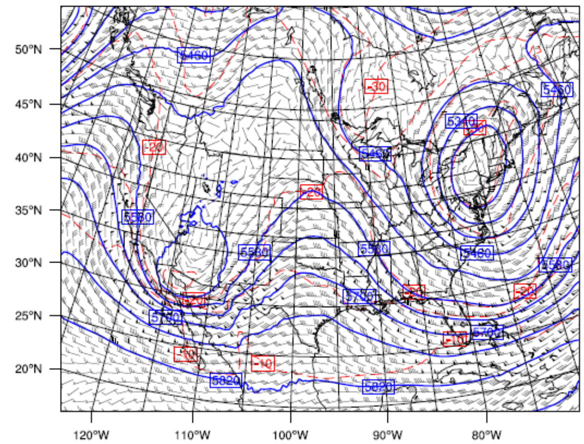
Figure 4-7 Comparison of Surface Weather Charts for February 28, March 2, and March 5, 2010

NWS Observations

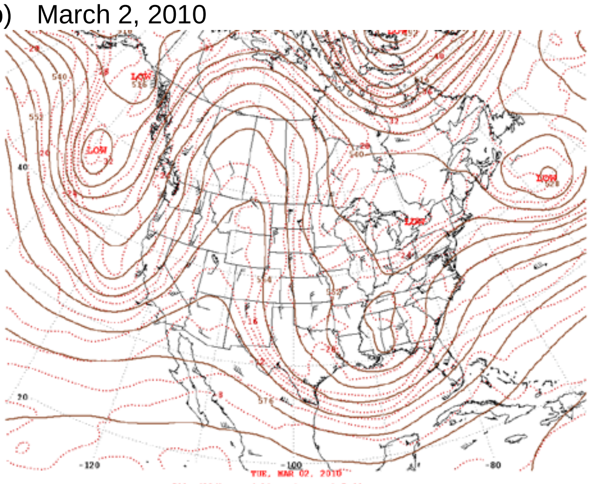
ARMS WRF 36-km

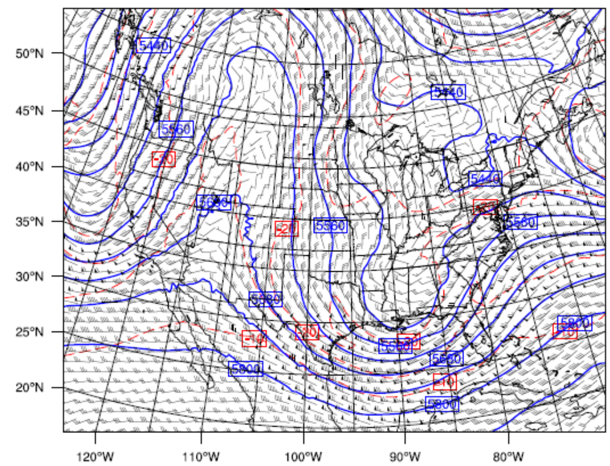
a) February 28, 2010



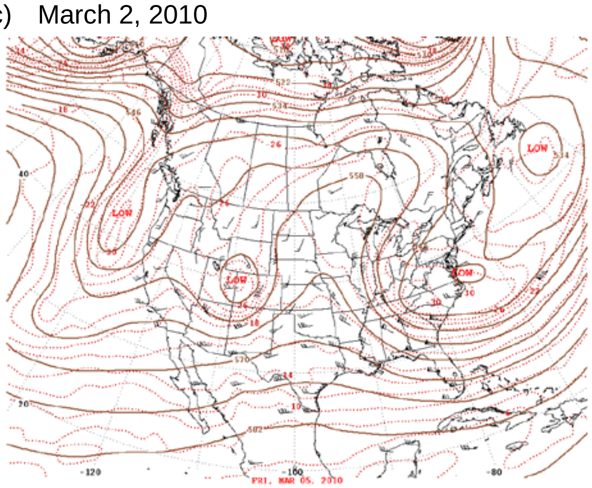


b) March 2, 2010





c) March 2, 2010



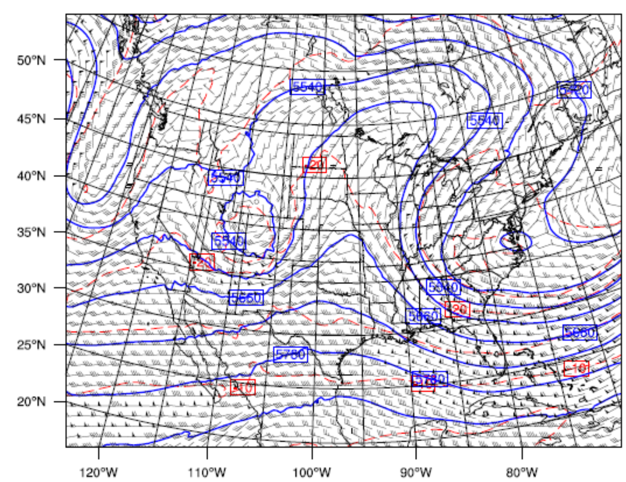


Figure 4-8 Comparison of 500-mb Height Contour Charts for February 28, March 2, and March 5, 2010

On February 28, 2010 (**Figure 4-9a**), the NWS reported precipitation over much of the southwest U.S. and the Pacific Northwest. Additional precipitation was reported over the northeast and Florida. The ARMS WRF 36-km simulation predicted more precipitation in the study area than actually occurred. Other than this area, the ARMS WRF precipitation patterns generally matched NWS observations. No precipitation was observed in the study area from March 2-4. **Figure 4-9b** shows that on March 2 ARMS WRF precipitation areas matched up well with NWS observations over the western U.S. Although, compared with the NWS observations, the ARMS WRF simulation produced slightly more precipitation in Colorado and less precipitation in New Mexico. The passage of the low pressure system on March 5-8 brought precipitation to the study area. On March 5 (**Figure 4-9c**), the NWS reported precipitation over Utah, Nevada, Idaho, and parts of Wyoming and Colorado. The ARMS WRF precipitation areas are fairly consistent with the NWS observations.

In the 4-km domain, on February 28, 2010, the NWS reported precipitation throughout parts of the state of Utah (**Figure 4-10a**). The ARMS WRF simulation predicted slightly more precipitation both in area and intensity. On March 2, 2010, NWS reported one-day precipitation amounts of less than 0.01 inches, while the ARMS WRF simulation (**Figure 4-10b**) predicted precipitation in the mountains east of Salt Lake City. On March 5, 2010, NWS reported precipitation over much of the state of Utah (**Figure 4-10c**) which matched well with the WRF predicted precipitation amount and extent. Overall, both ARMS WRF precipitation area and intensity are in agreement with NWS observations during this period.

In **Figure 4-11**, the WRF 36-km and 4-km domain snow cover fields are compared with NIC IMS observation data on March 6, 2010. Snow cover data were reviewed for other days during the period February 21 through March 8 and the results were similar to March 6, so they are not reproduced here. In **Figure 4-11**, the chart on the left side is from the NIC IMS (NOAA 2013). The chart in the middle is from the ARMS WRF 36-km simulation and one on the right is from the 4-km simulation. In the NIC IMS plots, areas with snow cover are shown in white, ice is shown in yellow, and uncovered land surfaces is shown in green. In the WRF plots, snow and ice are shown in white and uncovered surfaces in green. The WRF snow cover field reproduces the observed snow cover well in both the 36-km domain and the 4-km domain during this period.

4.1.2.2 Conditions Aloft

Morning soundings at SLC and GJT were compared with the WRF model simulated vertical structure of the atmosphere in **Figures 4-12a** and **4-12b**, respectively. The vertical model performance is analyzed for the same 3 days selected for the surface layers (February 28, March 2, and March 5).

In general, the temperature profiles (shown with solid lines) from the ARMS WRF simulation agrees with the observed vertical profiles. However, on February 28 and March 5, the ARMS WRF simulation predicted much lower surface temperature than the NWS observation at both sites. The low temperature bias appears to be limited to the surface predictions and has better agreement aloft. The temperature profile performed better than the dew point temperature profile (shown with dashed lines). The model had difficulty replicating the sharp changes in the dew point temperature that is related to the water vapor mixing ratio. This is especially apparent on March 2 where the ARMS WRF simulations failed to follow the dew point temperature profile between approximately 650 to 400 mb at both sites. Otherwise, the model was able to follow the dew point temperature profile up to 300 mb. The WRF model generally was able to simulate the vertical variability in wind direction and speed. In general, the model was capable of reproducing the height of the inversions as well as the strengths.

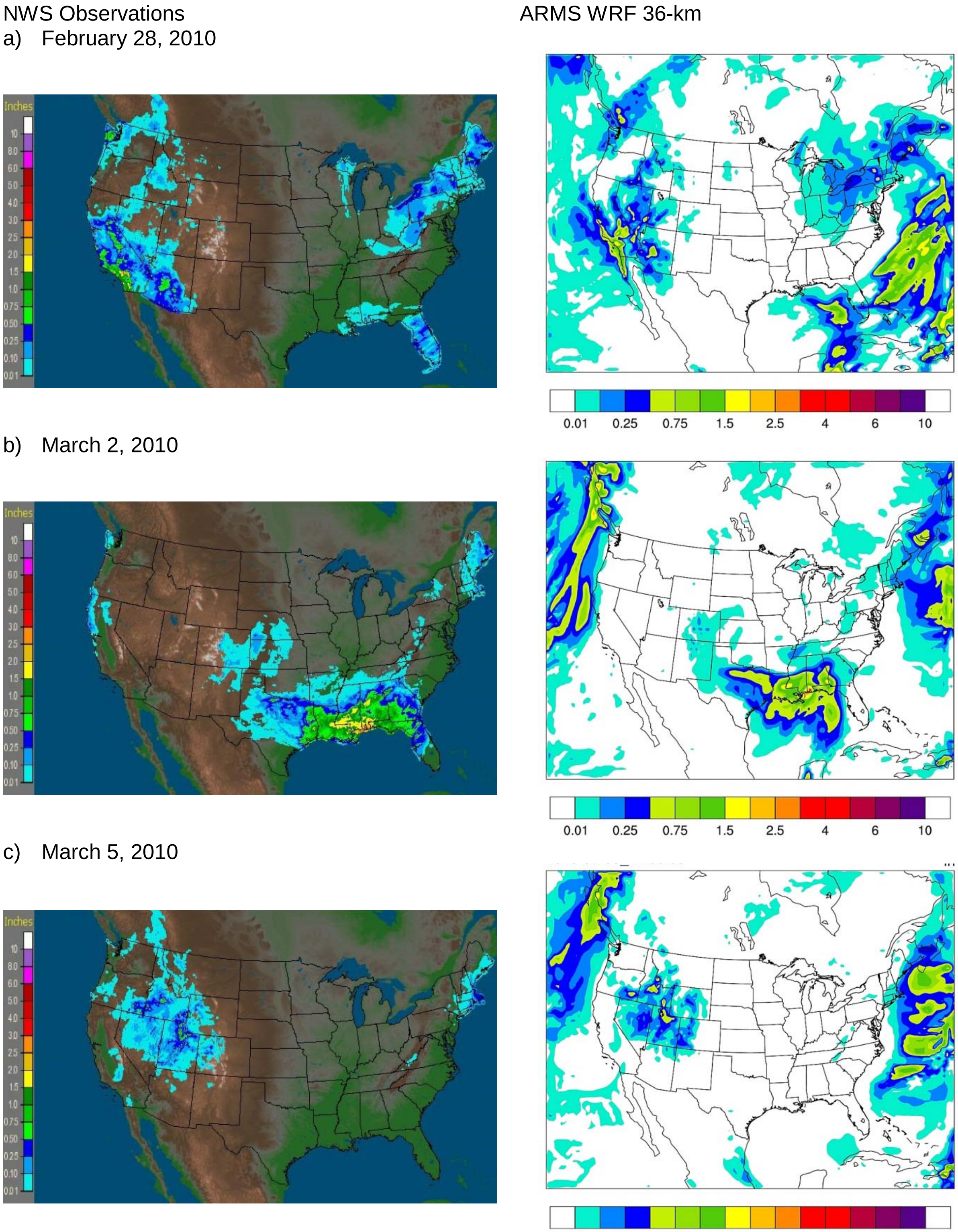
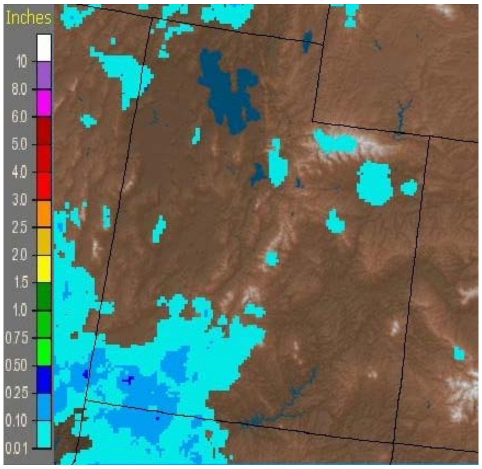


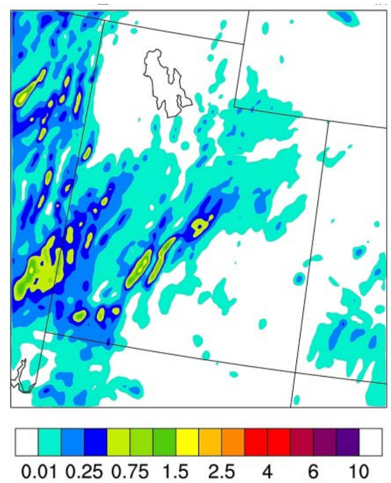
Figure 4-9 Comparison of 36-km Domain Precipitation Amounts for February 28, March 2, and March 5, 2010

NWS Observations

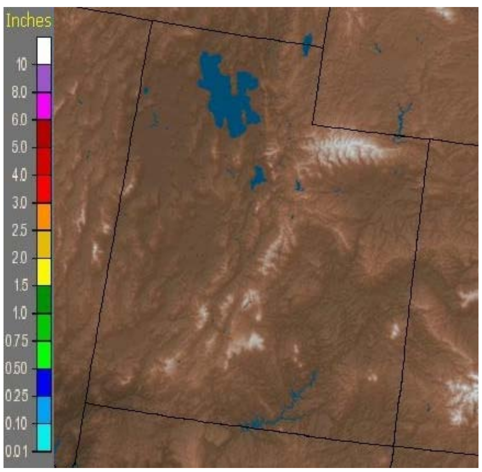
ARMS WRF 4-km

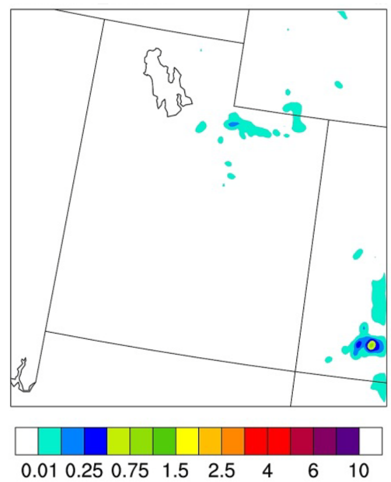
a) February 28, 2010



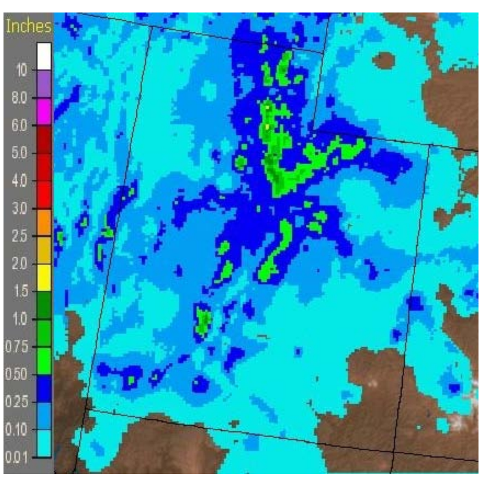


b) March 2, 2010





c) March 5, 2010



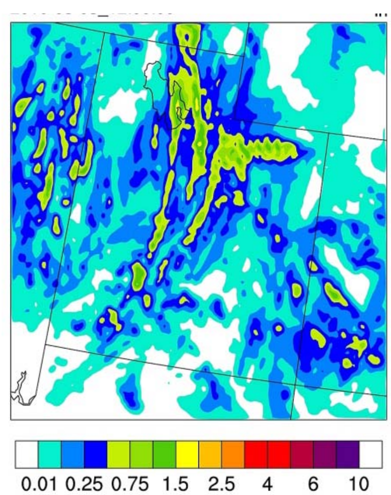


Figure 4-10 Comparison of 4-km Domain Precipitation Amounts for February 28, March 2, and March 5, 2010

NIC IMS Data ARMS WRF 36-km ARMS WRF 4-km

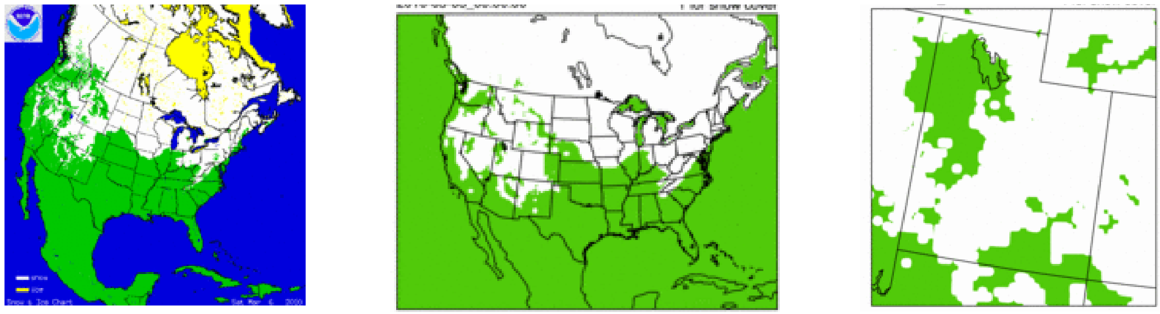
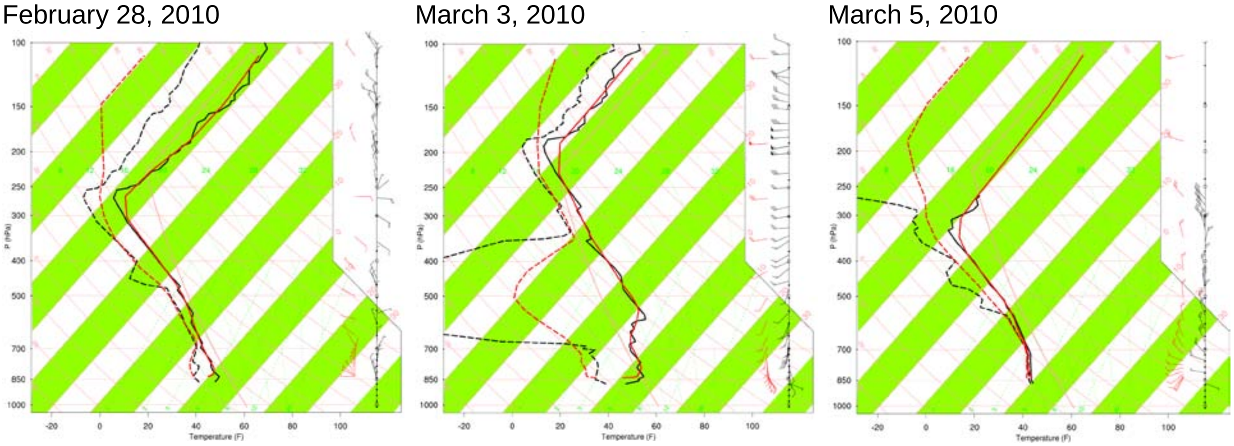


Figure 4-11 Comparison of Snow Cover on March 6, 2010

a) Salt Lake City (SLC)

February 28, 2010 March 3, 2010 March 5, 2010



b) Grand Junction (GJT)

February 28, 2010 March 3, 2010 March 5, 2010

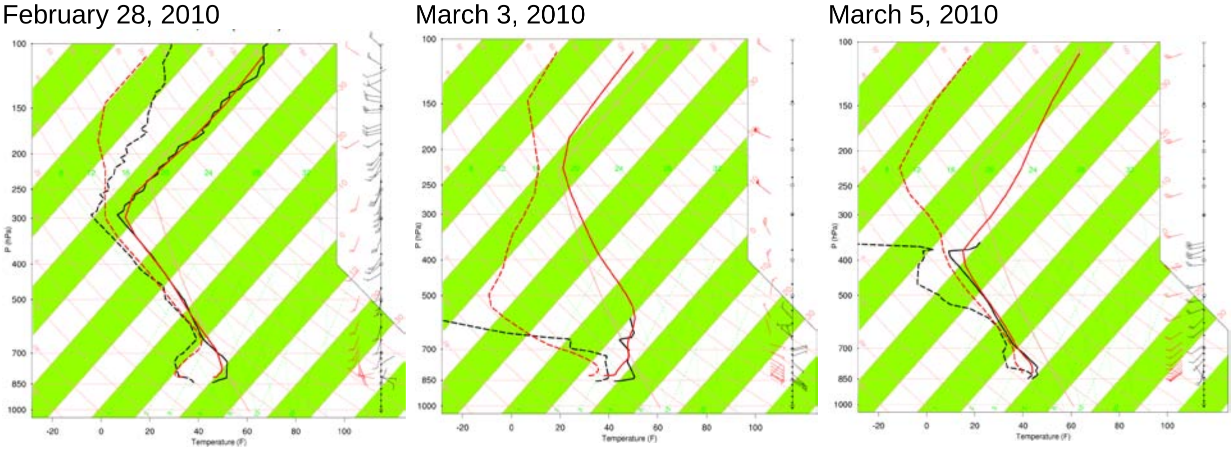


Figure 4-12 Comparison of Vertical Profiles at SLC and GJT for February 28, March 2, and March 5, 2010

4.1.3 August 19 to 29, 2010 Air Quality Event

4.1.3.1 Synoptic Conditions

Figures 4-13, **4-14**, and **4-15** compare WRF model results to observations for surface weather, 500-mb height contours, and 24-hour precipitation amounts, respectively. The surface charts and 500-mb height contours are shown for 1200 GMT, which corresponds with 6:00 a.m. Mountain Daylight Time (MDT). The precipitation plots show cumulative amounts over a 24-hour period. The charts on the left side are from the NWS; those on the right are from the ARMS WRF simulation. Each figure shows three days in the middle of the August 19 - 29 event: August 20, 23, and 26, 2010.

Throughout a majority of the period August 19- 29, some portion of the Western U.S. was impacted by a surface high pressure system and a ridge aloft. Limited precipitation fell across the Western U.S. during this time period. A few frontal systems moved through the Northwestern U.S. to central Canada and brought light to moderate precipitation to the Rocky Mountains region.

On August 20 a weak 1,012-mb high pressure system over Four Corners region slowly moved east to the Central U.S. region. Precipitation developed across the Central Rockies associated with a weak shortwave trough and westerly flow aloft. The ARMS WRF simulation was able to reproduce the NWS observations (shown in **Figure 4-13a**). By August 23, a surface low pressure system over Northern Central Plains was starting to deepen and progress northeast. The trailing cold front passed through western U.S. bringing heavier precipitation to the north with lighter precipitation in the south. The location of the cold front and the high pressure system pushing in from the Pacific Northwest were successfully captured by the ARMS WRF simulation (**Figure 4-13b**). A surface high pressure system developed behind the low dominates the area for next several days. An upper level ridge builds from the south over the Western U.S. With the high pressure system over much of the U.S. on August 26, little precipitation and no clouds develop can be observed with the exception of the far Northwest and east. A low pressure moves into the Northwest from the Gulf of Alaska. The ARMS WRF simulation once again was able to reproduce the NWS observations (**Figure 4-13c**). Over the next few days, the low pressure moves east across the U.S. Canada border finally pushing the high pressure system out of the plains and breaking down the ridge aloft.

The mid-level (500-mb) pattern diagnosed by NWS from observations on August 20 shows a weak trough over parts of Colorado and a weak ridge that covered the eastern U.S. (**Figure 4-14a**). The location and strength of the western trough and eastern ridge from the ARMS WRF simulation compares favorably with the NWS observations. By August 23, a trough develops over Canada and parts of western U.S. while a high pressure ridge formed over the Great Lakes (**Figure 4-14b**). Additionally, the ARMS WRF simulation produced wind speeds and directions also agree well with the NWS at the 500-mb level. On August 26, 2010, the NWS 500-mb pattern shows a high pressure system that covered much the western U.S. (**Figure 4-14c**). The ARMS WRF simulation was able to reproduce the NWS observations.

On August 20, 2010 (**Figure 4-15a**), the NWS reported precipitation over much of the study area. Additional precipitation was reported over parts of the central states, the Great Lakes, and the northeast. The ARMS WRF 36-km simulation was able to reproduce the observed precipitation patterns. On August 23, precipitation was observed in the study area and along the east coast. **Figure 4-15b** show the ARMS WRF precipitation areas matched up well with NWS observations over the western U.S. Although, compared with the NWS observations, the ARMS WRF simulation produced less precipitation over the study area. On August 26 (**Figure 4-15c**), the NWS reported little precipitation over most of the western and central U.S. The ARMS WRF precipitation areas match up well the NWS observations over the study area. However, the ARMS WRF simulation predicted slightly more precipitation over southern California, Nevada, and parts of Arizona.

In the 4-km domain, on August 20, 2010, the NWS reported precipitation over much of the state of Utah (**Figure 4-16a**). The ARMS WRF simulation predicted slightly less precipitation than NWS over portions of the state, but higher precipitation than the NWS in the Uinta Basin area. On August 23, 2010, NWS reported precipitation over the northeastern portion of Utah and the four corners area (**Figure 4-16b**). The ARMS WRF simulation reported slightly less precipitation coverage than NWS. Both NWS and ARMS WRF reported precipitation over the Uinta Basin. On August 26, 2010, both NWS and ARMS WRF reported very little precipitation over much of the state of Utah (**Figure 4-16c**).Overall, there was very good agreement between the NWS and the WRF model precipitation during this period.

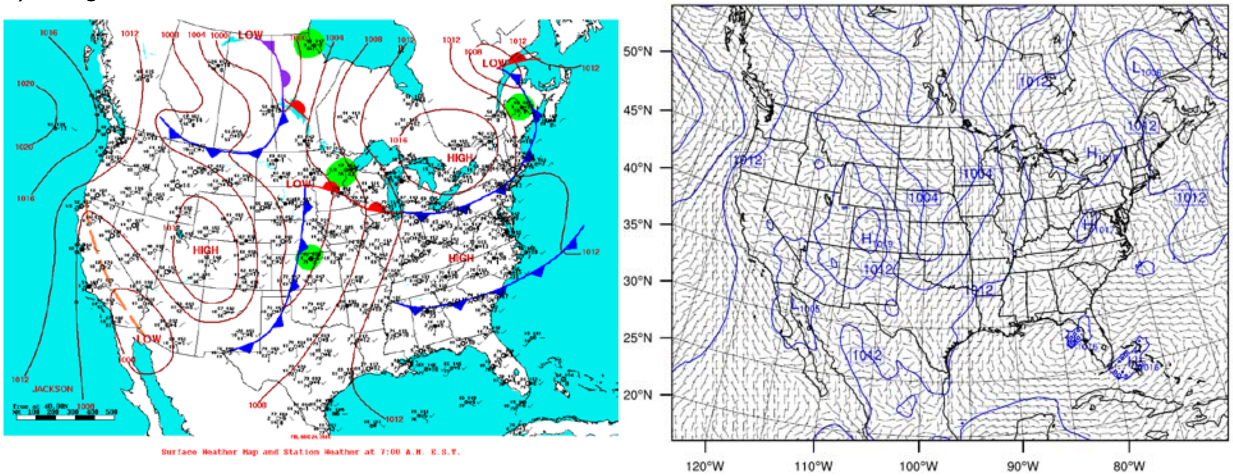
4.1.3.2 Conditions Aloft

Morning soundings at SLC and GJT were compared with the WRF model simulated vertical structure of the atmosphere in **Figures 4-17a** and **4-17b**, respectively. The vertical model performance is analyzed for the same 3 days selected for the surface layers (August 20, 23, and 26).

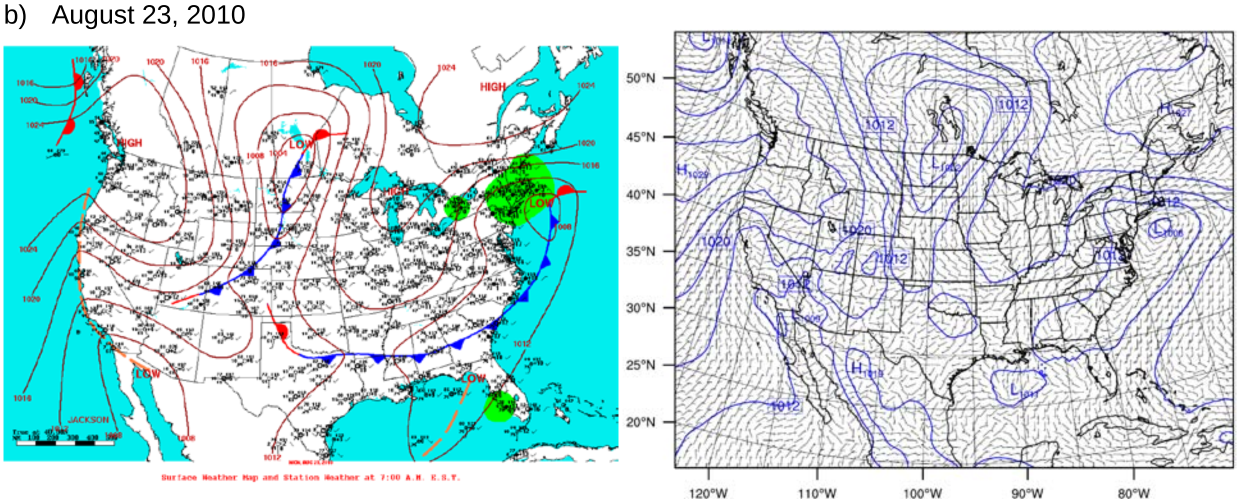
The temperature (solid lines) profiles from the ARMS WRF simulations followed the profiles from the observed soundings. At the SLC upper air sounding site, the model under-predicted temperature at the surface, but was able to capture the early morning temperature inversion. The temperature profile at the GJT site showed that the model under-predicted temperature at the lowest levels. Otherwise, the model's temperature profile followed observation reasonably well. Overall, the temperature profile performed better than the dew point temperature profile (dashed lines). The model had difficulty replicating the sharp changes in the dew point temperature that is related to the water vapor mixing ratio. The model generally was able to simulate the vertical variability in wind direction and speed.

NWS Observations ARMS WRF 36-km

a) August 20, 2010



b) August 23, 2010



c) August 26, 2010

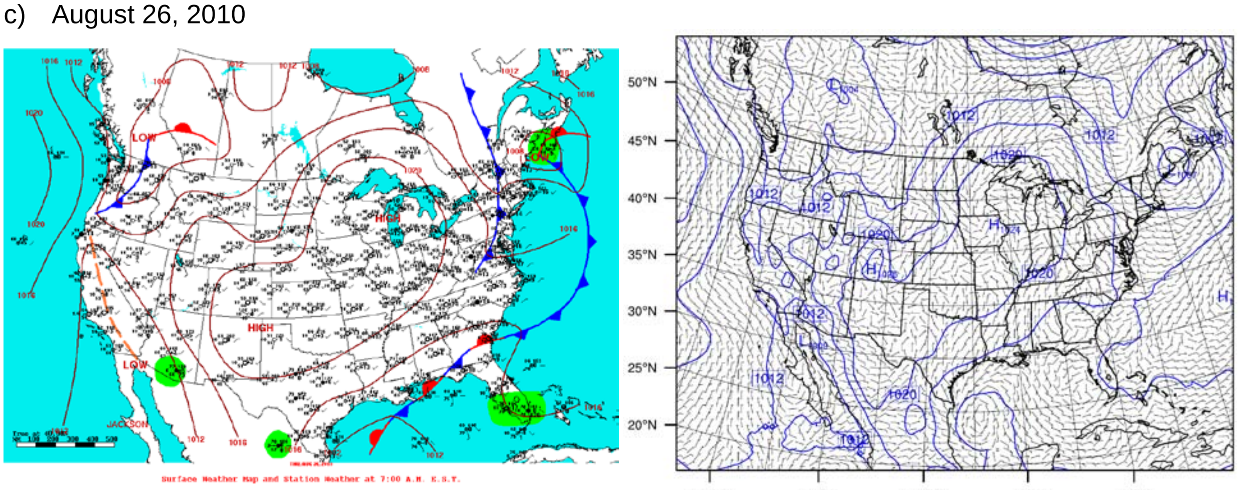
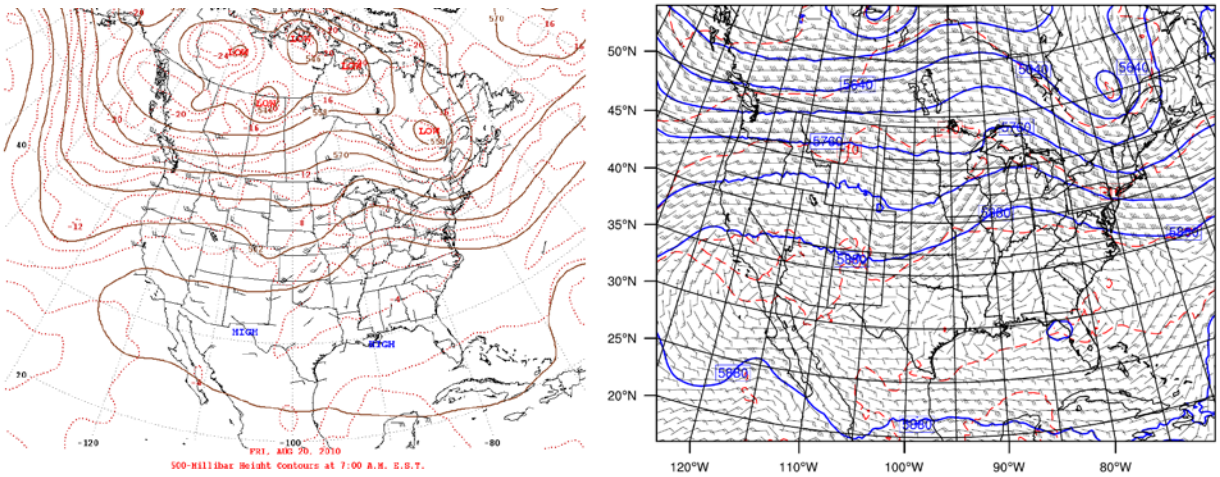


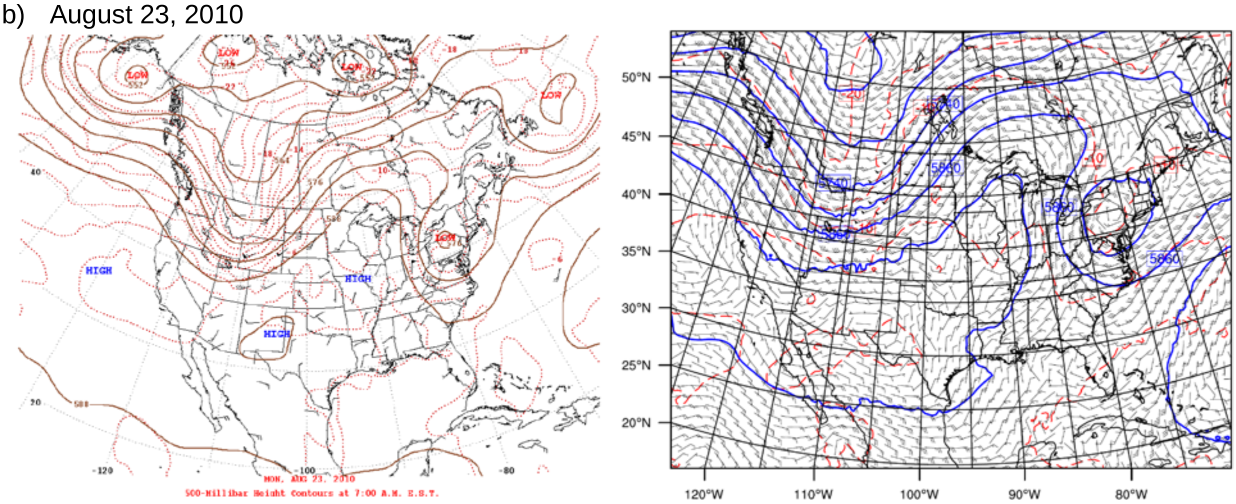
Figure 4-13 Comparison of Surface Weather Charts for August 20, 23, and 26, 2010

NWS Observations ARMS WRF 36-km

a) August 20, 2010



b) August 23, 2010



c) August 26, 2010

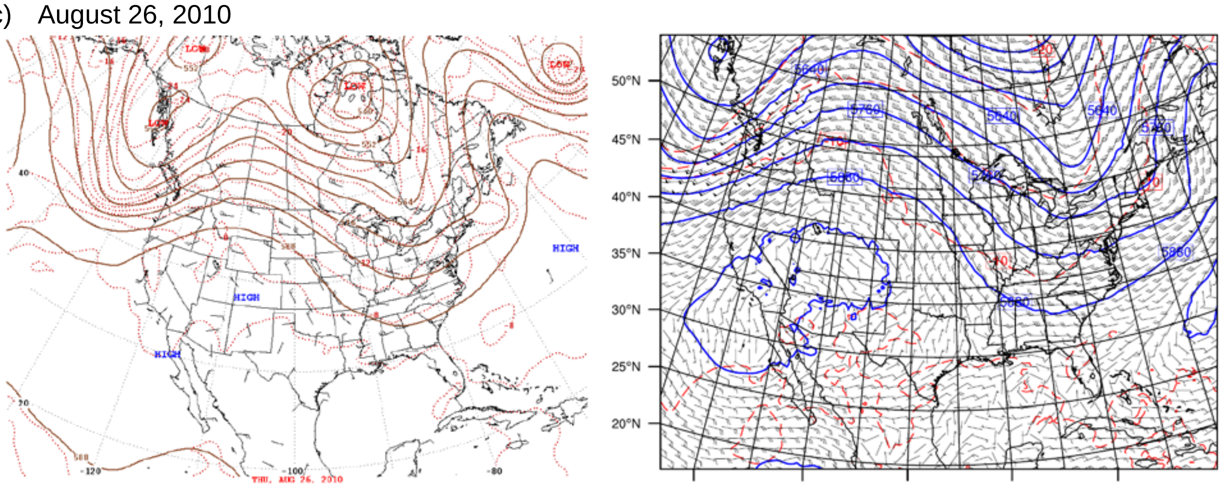


Figure 4-14 Comparison of 500-mb Height Contour Charts for August 20, 23, and 26, 2010

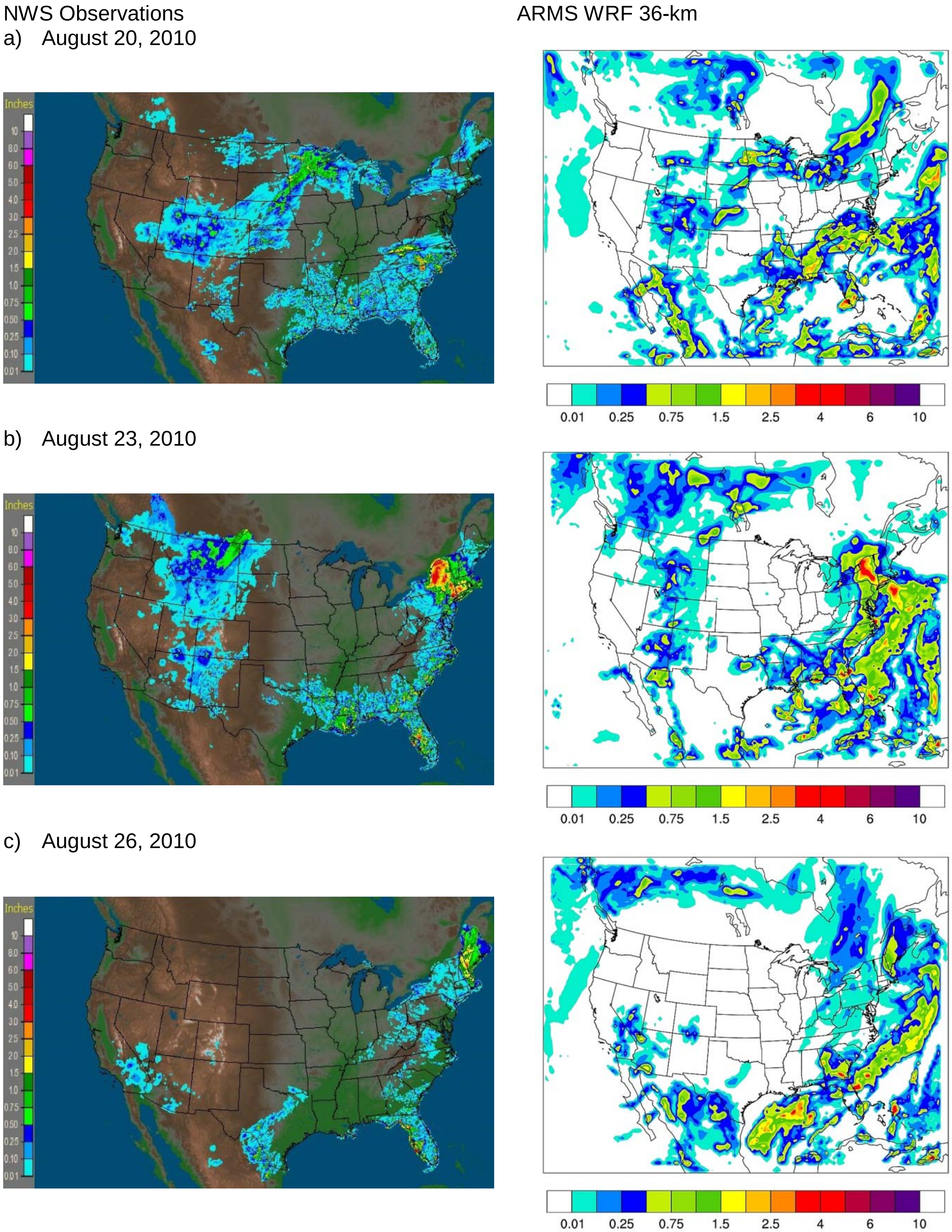
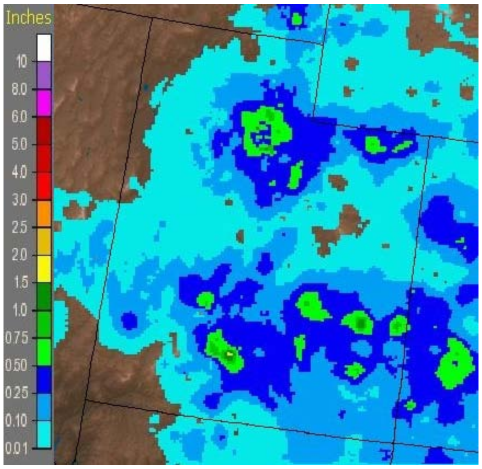


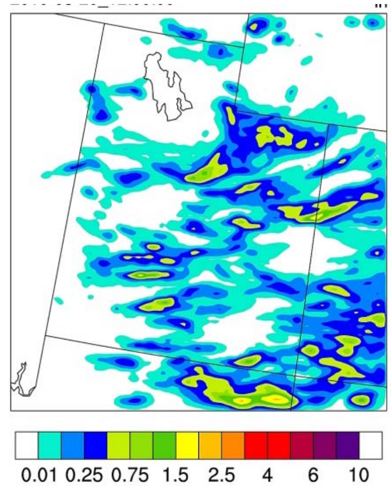
Figure 4-15 Comparison of 36-km Domain Precipitation Amounts for August 20, 23, and 26, 2010

NWS Observations

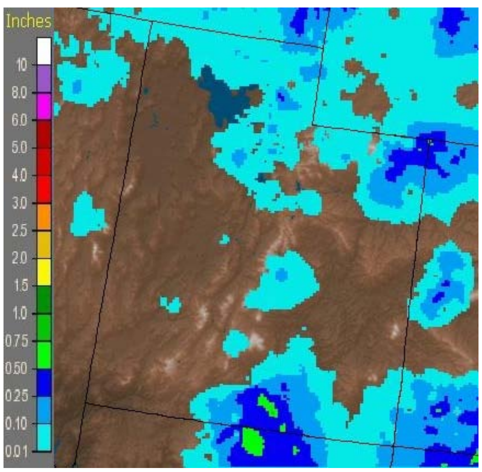
ARMS WRF 4-km

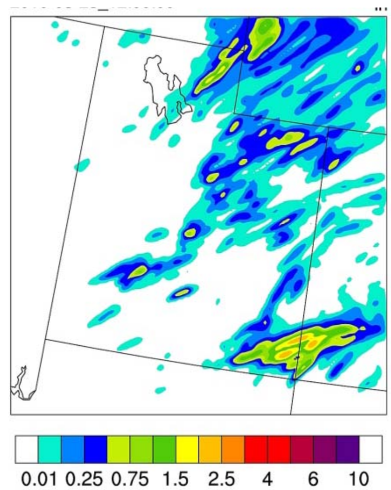
a) August 20, 2010



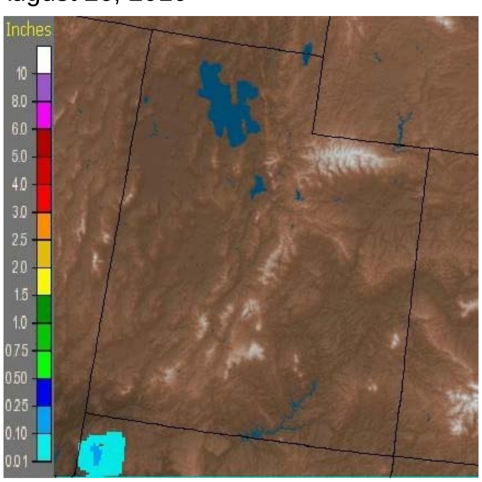


b) August 23, 2010





c) August 26, 2010



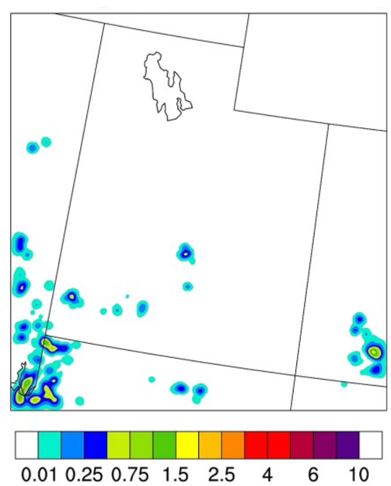


Figure 4-16 Comparison of 4-km Domain Precipitation Amounts for August 20, 23, and 26, 2010

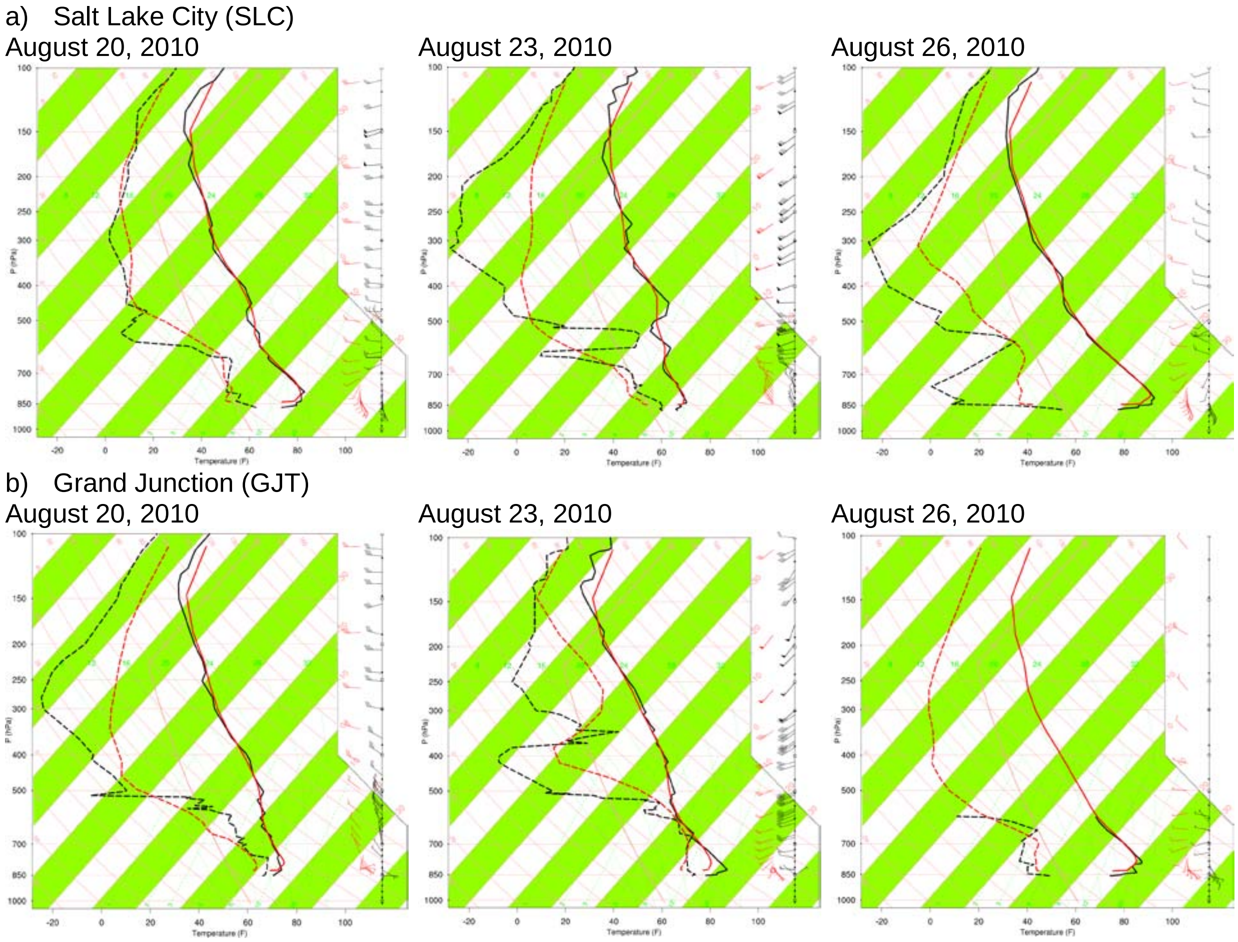


Figure 4-17 Comparison of Vertical Profiles at SLC and GJT for August 20, 23, and 26, 2010

4.1.4 September 27 to October 7, 2010 Air Quality Event

4.1.4.1 Synoptic Conditions

Figures 4-18, **4-19**, and **4-20** compare WRF model results to observations for surface weather, 500-mb height contours, and 24-hour precipitation amounts, respectively. The surface charts and 500-mb height contours are shown for 1200 GMT, which corresponds with 6:00 a.m. MDT. The precipitation plots show cumulative amounts over a 24-hour period. The charts on the left side are from the NWS; those on the right are from the ARMS WRF simulation. Each figure shows 3 days in the middle of the September 27 – October 8 event: September 28, October 1, and 4, 2010.

During the September 27 – October 7, 2010, an upper level ridge over the southwest slowly moved northeast and strengthened over the central plains. The eastern U.S. is largely influenced by an upper level trough and associated surface frontal system. The frontal system brought moderate to heavy precipitation to the east and northeast throughout the period. In the regions impacted by surface high pressure system and upper level ridge, little to no precipitation developed.

On September 28, 2010 (**Figure 4-18a** and **4-19a**) the upper level ridge over the southwest continued to slowly build northeast while the upper level trough over the east dug south. This pattern remained stagnant over the next several days. By October 1, 2010 (**Figure 4-18b** and **4-19b**) an upper level cutoff low began to develop in the Gulf of Alaska. Slowly the upper level low deepened resulting in a longwave trough along the west coast on October 4, 2010 (**Figure 4-18c** and **4-19c**). The ARMS WRF simulation was able to reproduce the NWS observations.

Precipitation (**Figure 4-20**) had been minimal across the western U.S. until weak surface low pressure system developed along the west coast ahead of the upper level trough. Towards the end of the high $PM_{2.5}$ event, a surface low pressure system has moved over the southwest and northeast with a high pressure system in the central plains. Light to moderate precipitation develop across the west with little to no precipitation in the Midwest by October 6, 2010. In general, the ARMS WRF 36-km domain model tends to under-predict the amount of precipitation, particularly in the west during the beginning of this event.

In the 4-km domain, for September 28, 2010 and October 1, 2010, both the NWS and ARMS WRF simulation reported no precipitation throughout the state of Utah (**Figure 4-21a** and **Figure 4-21b**). On October 4, 2010, NWS reported precipitation over much of the southern parts of Utah and in areas northeast of the Great Salt Lake (**Figure 4-21c**). On October 4, the ARMS WRF simulation reported slightly less precipitation over these areas, but intense pockets of precipitation in the four corners area. During this event, the NWS and the ARMS WRF precipitation have good agreement over the 4-km domain and within the Uinta Basin.

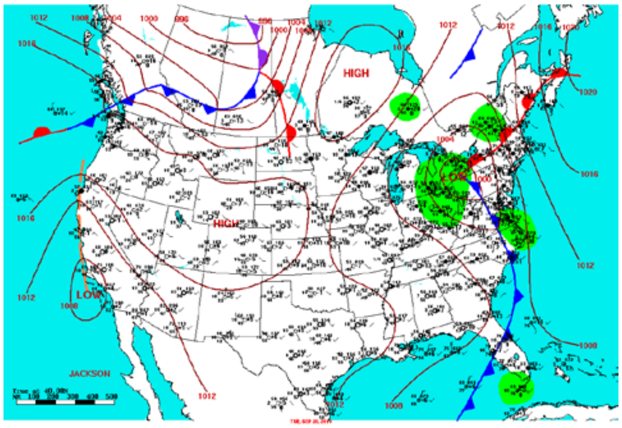
4.1.4.2 Conditions Aloft

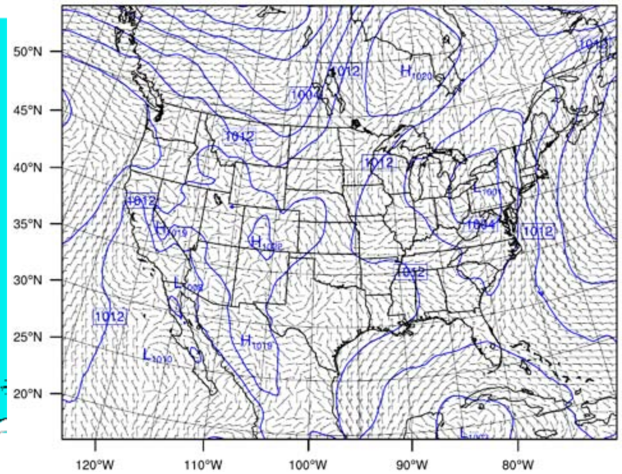
Morning soundings at SLC and GJT were compared with the WRF model simulated vertical structure of the atmosphere in **Figures 4-22a** and **4-22b**, respectively. The vertical model performance is analyzed for the same three days selected for the surface layers (September 28, October 1 and 4).

Similar to the vertical sounding profiles from the August 19-29 episode, the temperature (solid lines) profiles from the ARMS WRF simulations followed the profiles from the observed soundings. The temperature inversion was better captured at the SLC upper air sounding site than the GJT site. Both sites showed that the model under-predicted temperature at the lowest levels. Otherwise, the model's temperature profile followed observation reasonably well. Overall, the temperature profile performed better than the dew point temperature profile (dashed lines). Once again, the model had difficulty replicating the sharp changes in the dew point temperature that is related to the water vapor mixing ratio. The model generally was able to simulate the vertical variability in wind direction and speed.

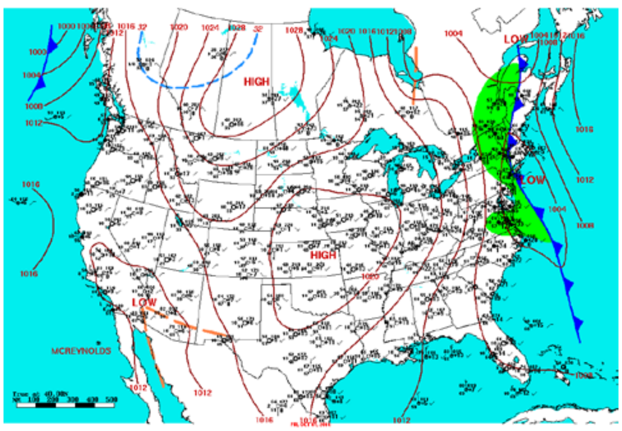
NWS Observations ARMS WRF 36-km

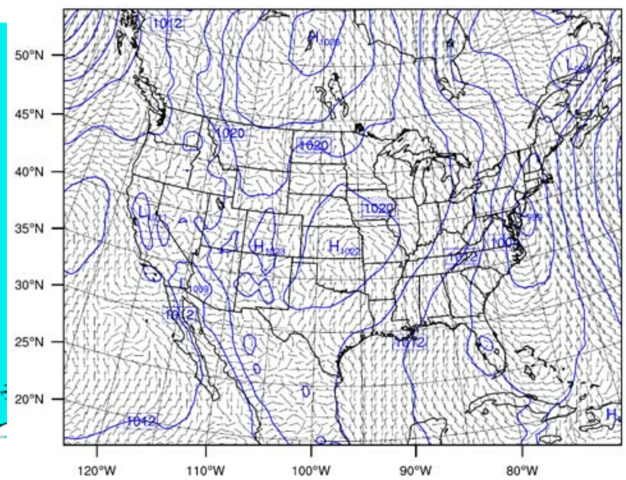
a) September 28, 2010



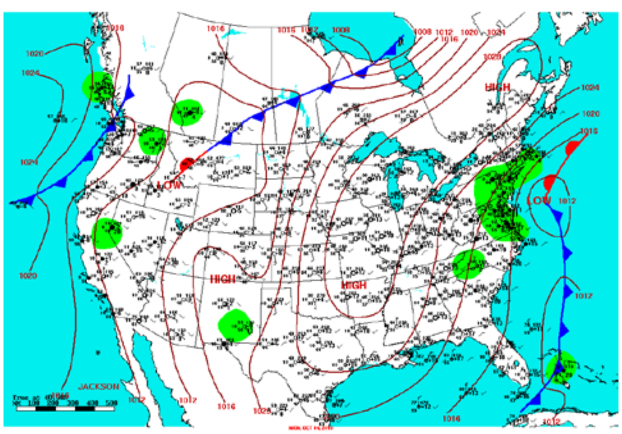


b) October 1, 2010





c) October 4, 2010



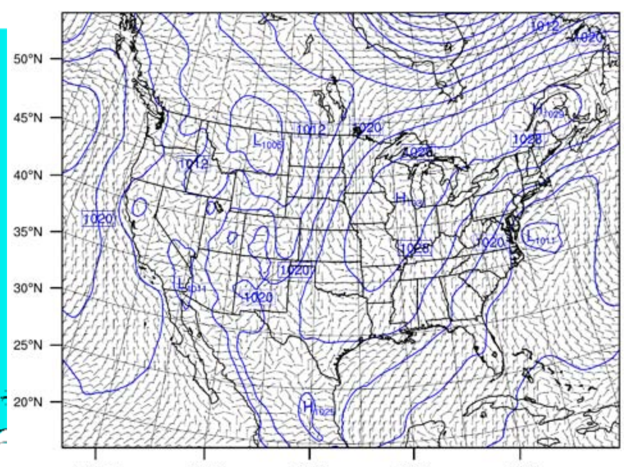
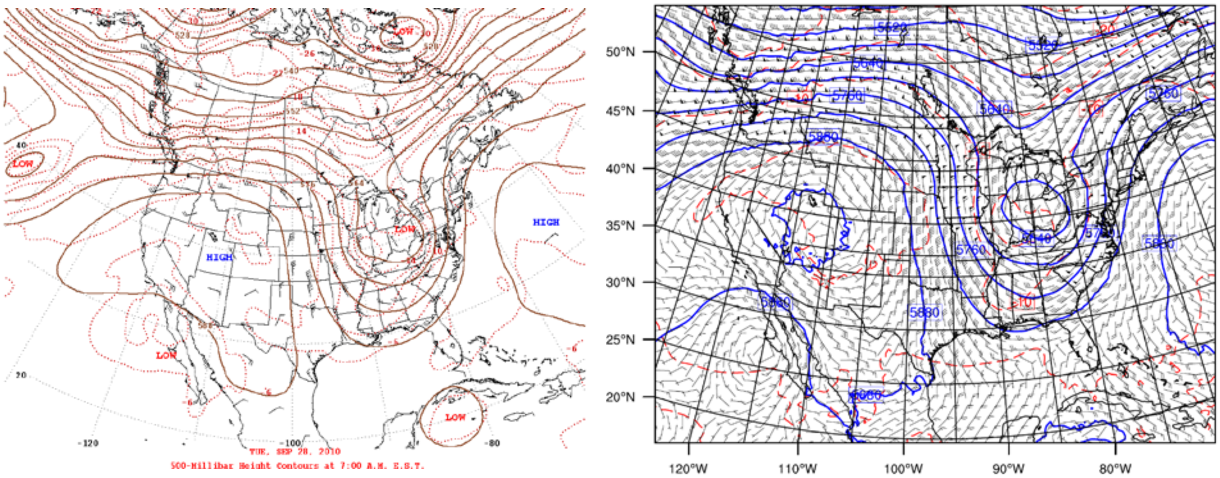


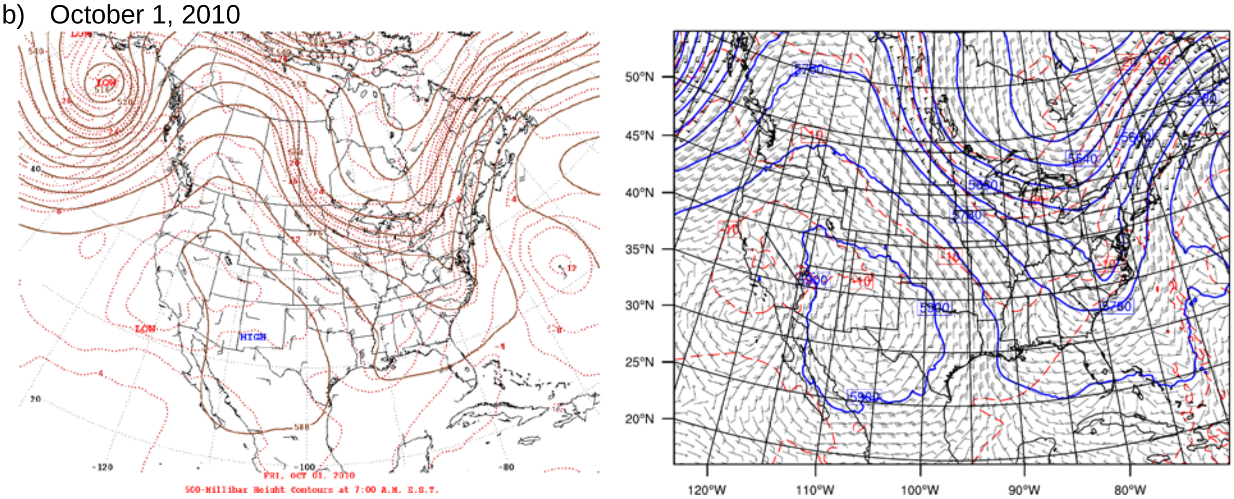
Figure 4-18 Comparison of Surface Weather Charts for September 28, October 1, and 4, 2010

NWS Observations ARMS WRF 36-km

a) September 28, 2010



b) October 1, 2010



c) October 4, 2010

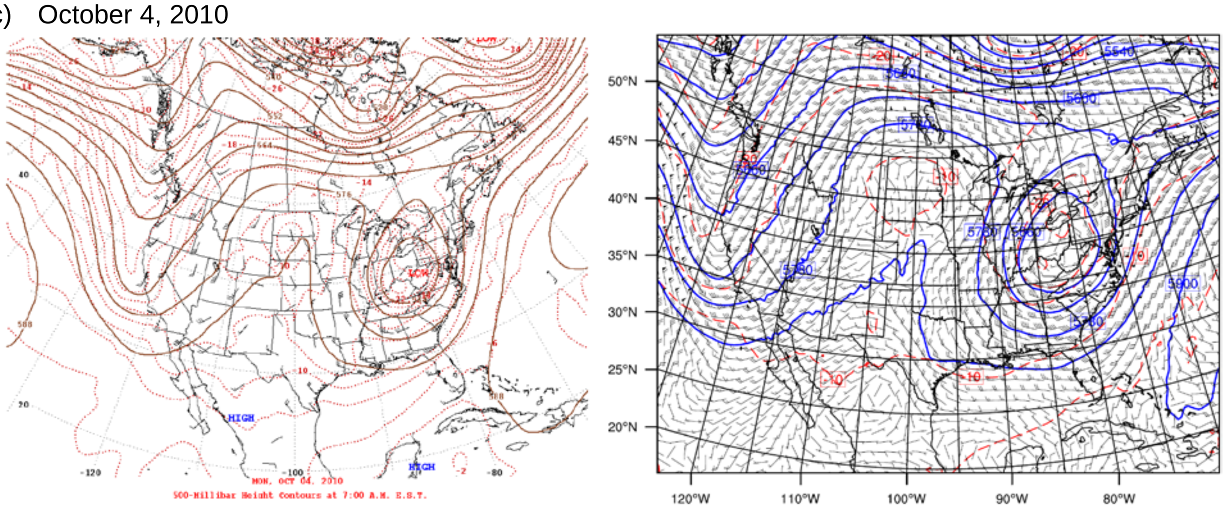
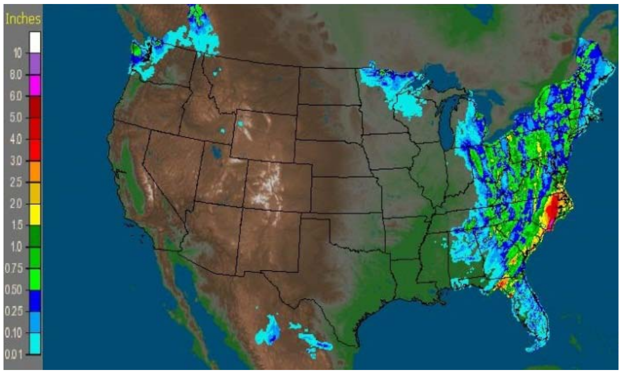


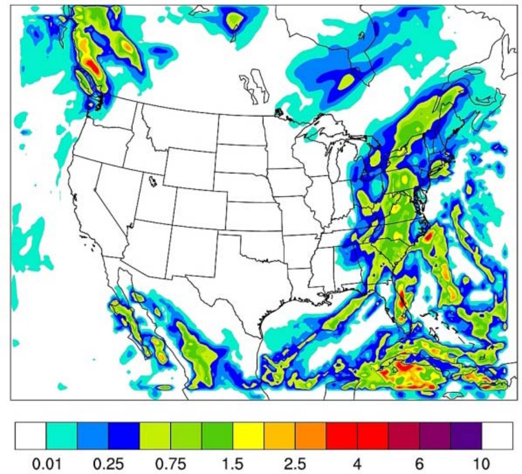
Figure 4-19 Comparison of 500-mb Height Contour Charts for September 28, October 1, and 4, 2010

NWS Observations

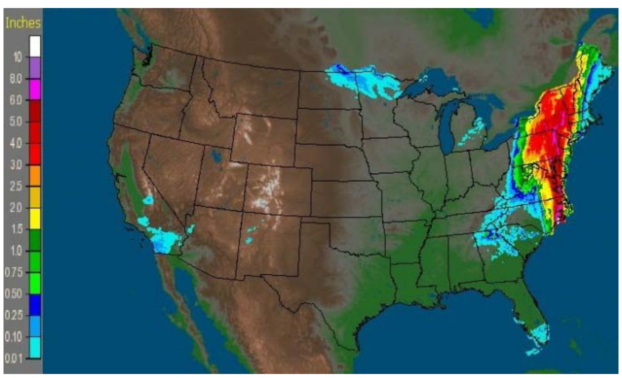
ARMS WRF 36-km

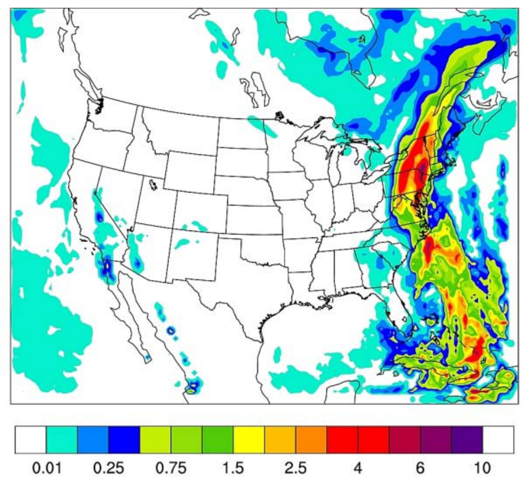
a) September 28, 2010



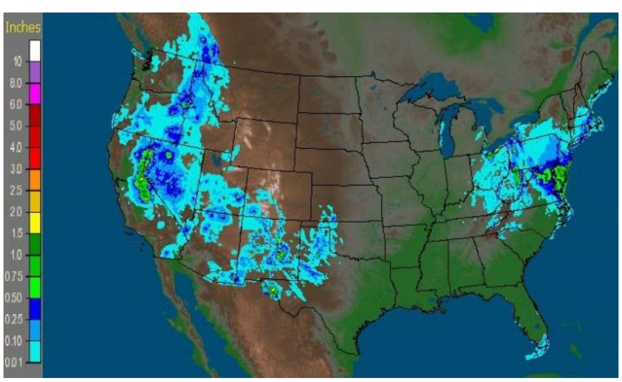


b) October 1, 2010





c) October 4, 2010



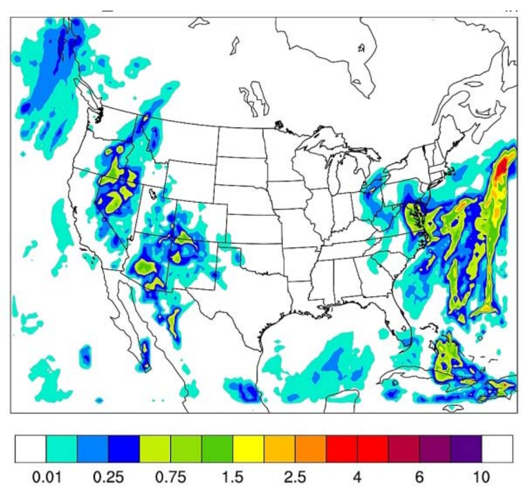


Figure 4-20 Comparison of 36-km Domain Precipitation Amounts for September 28, October 1, and 4, 2010

NWS Observations ARMS WRF 4-km

a) September 28, 2010

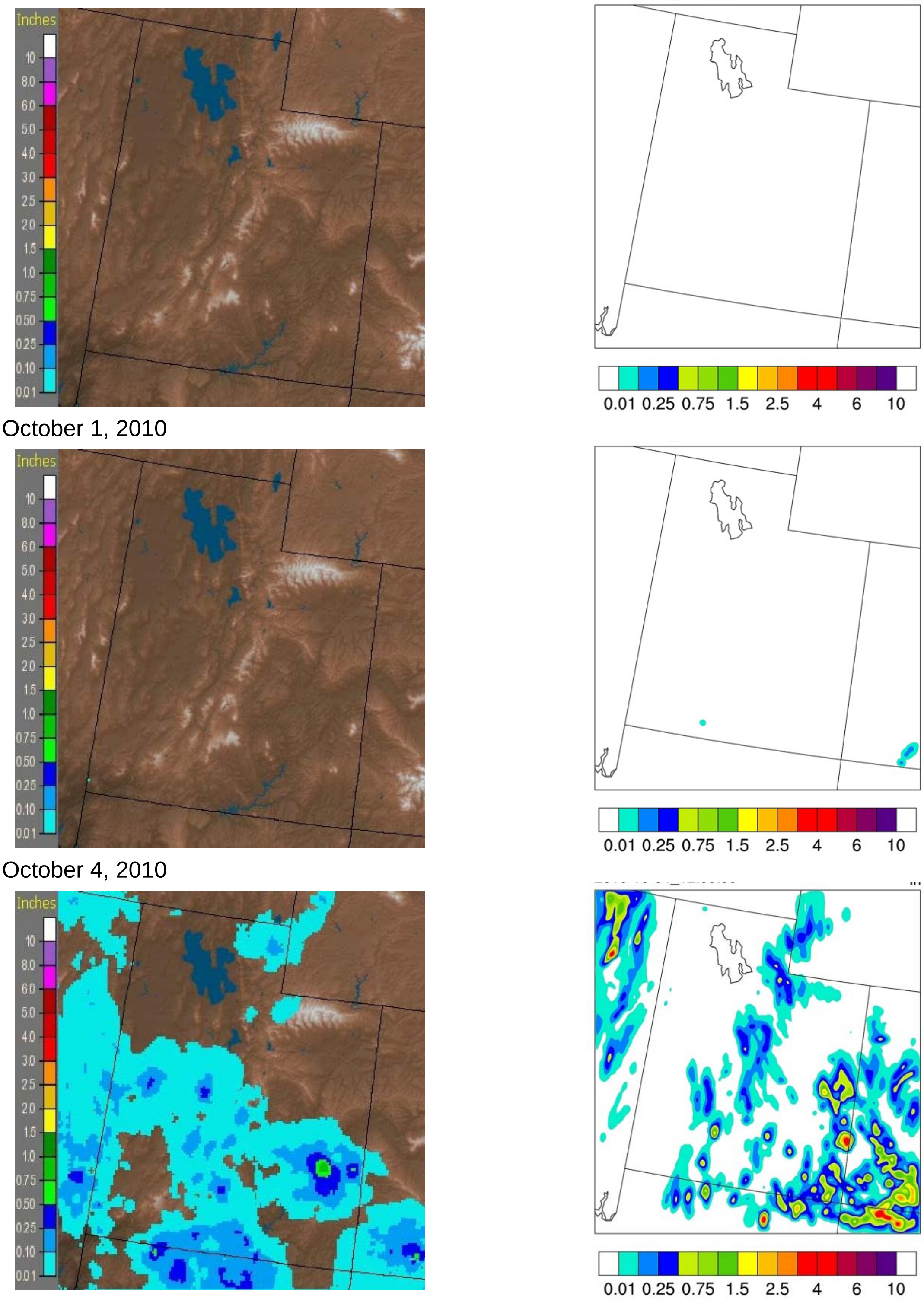


Figure 4-21 Comparison of 4-km Domain Precipitation Amounts for September 28, October 1, and 4, 2010

4.1.5 Evaluation of Snow Cover Extent

Given the influence snow cover has on surface albedo and reflected UV radiation, it is important to assess the extent and timing of snow cover throughout the annual simulation and compare the result to actual conditions in 2010. Given the relatively slow rate of change of snow cover relative to other meteorological parameters, the first day of each month was evaluated and results are considered to be representative of seasonal model performance for this parameter. The modeled snow cover for the first day of each month in 2010 was compared to snow cover from NIC IMS in **Figures 4-23**, **4-24**, and **4-25**. In these three figures, the charts on the left side are from the NIC IMS (NOAA 2013). The charts in the middle row are from the ARMS WRF 36-km simulation and the row on the right is from the 4-km simulations. In the NIC IMS plots areas with snow cover are shown in white, ice is shown in yellow, and uncovered land surfaces is shown in green. In the WRF plots, snow and ice are shown in white and uncovered surfaces in green.

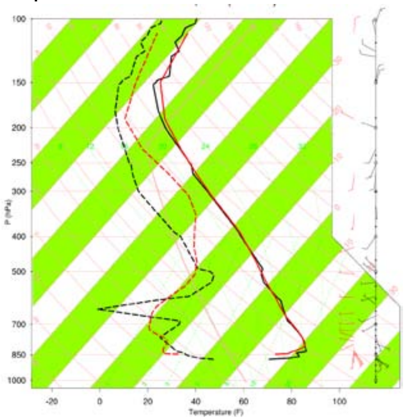
During January, February, March and April, 2010 (**Figure 4-23**), WRF 36-km and 4-km snow cover patterns agree well with NIC IMS snow cover. The WRF 4-km snow cover pattern is consistent with NIC IMS snow cover over most of Utah. In general, the WRF model maintained the widespread snow coverage throughout the Uinta Basin during winter and adequately captured the timing of the spring snowmelt. Similar to NIC IMS snow coverage, neither the 36-km nor the 4-km WRF results have snow or ice over the Great Salt Lake. On February 1, 2010 (**Figure 4-23b**), WRF 36-km reported slightly more snow cover along the west coast and less snow cover over parts of the central and southeastern U.S. than compared to NIC IMS..

On May 1, 2010 (**Figure 4-24a**), the WRF 36-km simulation over-predicted snow coverage over the entire western United States and the WRF 4-km simulation over-predicted snow over much of Utah, although WRF accurately predicts no snow in the Uinta Basin by this time of year. By June 1, 2010 the WRF snow cover is almost completely gone throughout the 36-km and 4-km domains, which is consistent with the NIC IMS snow cover, and remains that way until November, 2010 (**Figure 4-25c**). On November 1, 2010, WRF 36-km and 4-km simulated slightly more snow cover than NIC IMS reports. By December 1, 2010 (Figure 4-23d), the WRF 36-km and 4-km simulated snow coverage are generally consistent with NIC IMS snow cover.

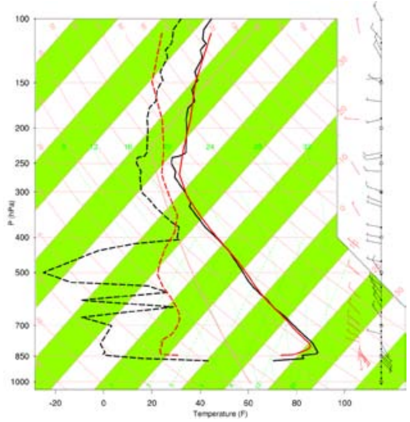
In general, the WRF model reproduced the extent of observed seasonal snow cover throughout the 36-km and 4-km domains, but had a tendency to over-estimate the extent of the snow cover during the shoulder seasons of late spring and fall.

a) Salt Lake City (SLC)

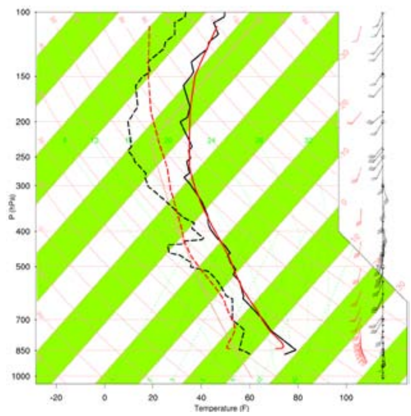
September 28, 2010



October 1, 2010

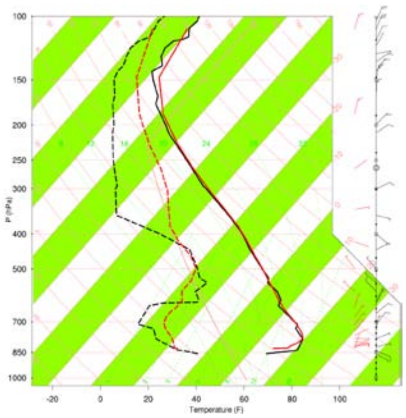


October 4, 2010

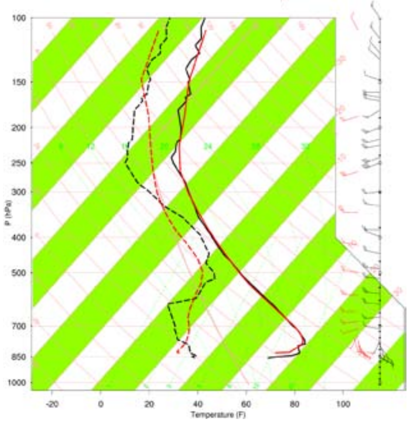


b) Grand Junction (GJT)

September 28, 2010



October 1, 2010



October 4, 2010

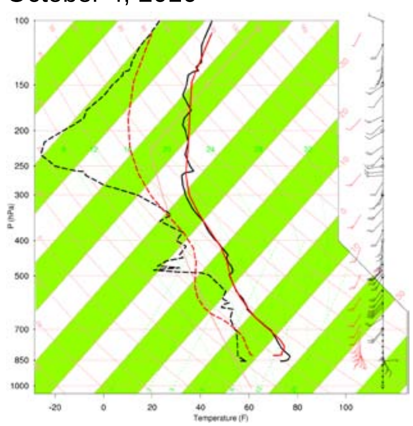


Figure 4-22 Comparison of Vertical Profiles at SLC and GJT for September 28, October 1, and 4, 2010

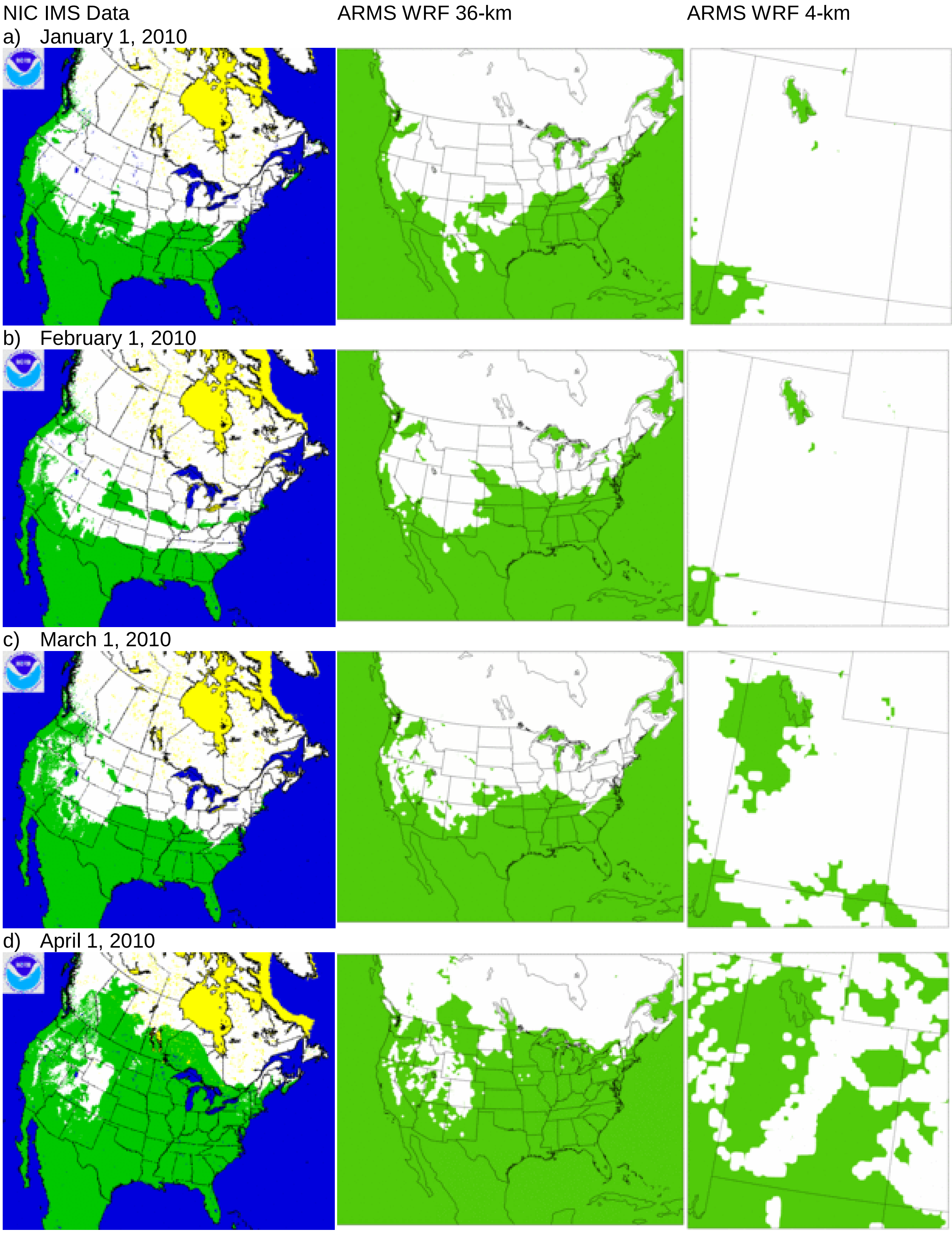


Figure 4-23 Snow Cover on January 1, February 1, March 1, and April 1, 2010

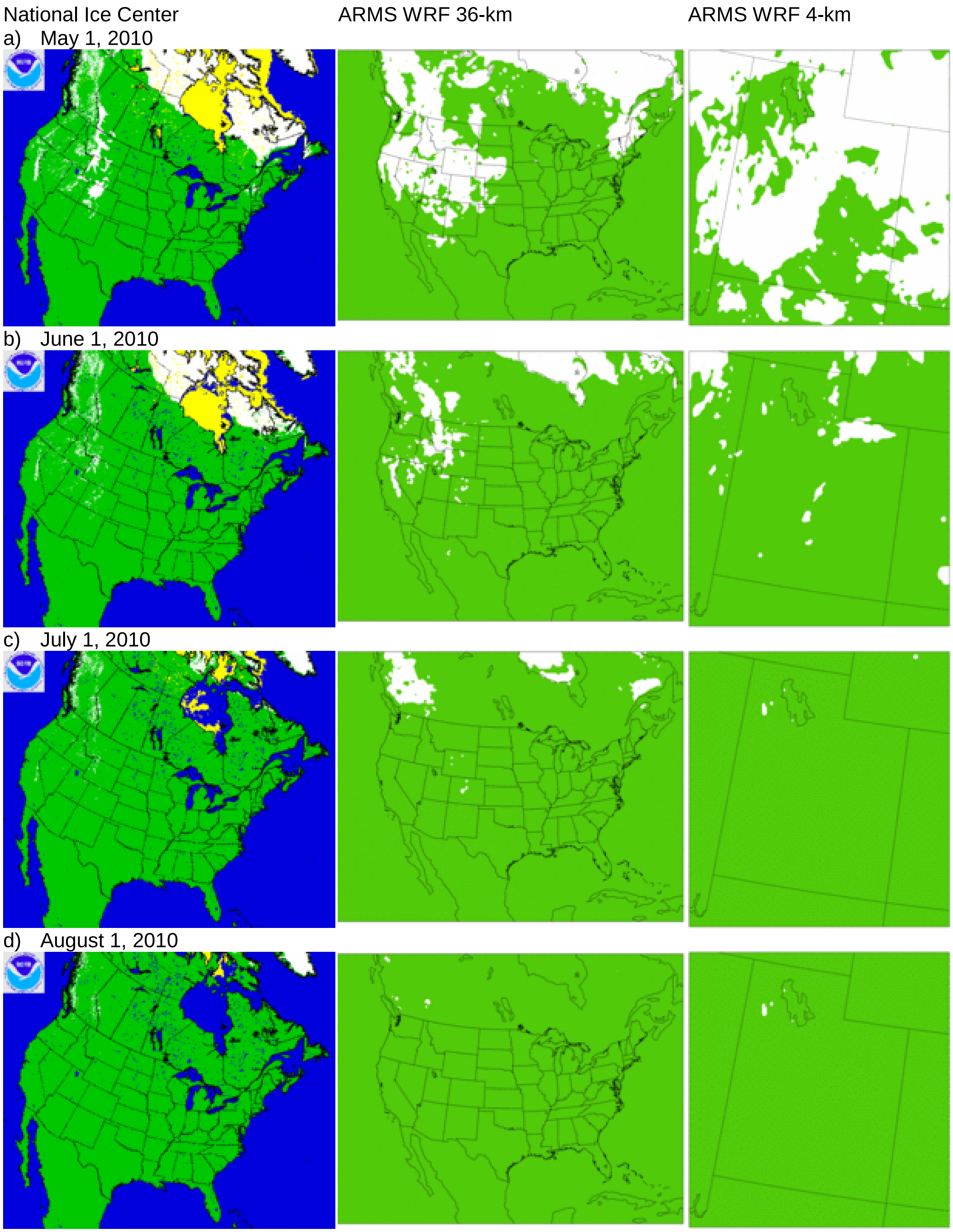


Figure 4-24 Snow Cover for Mary 1, June 1, July 1, and August 1, 2010

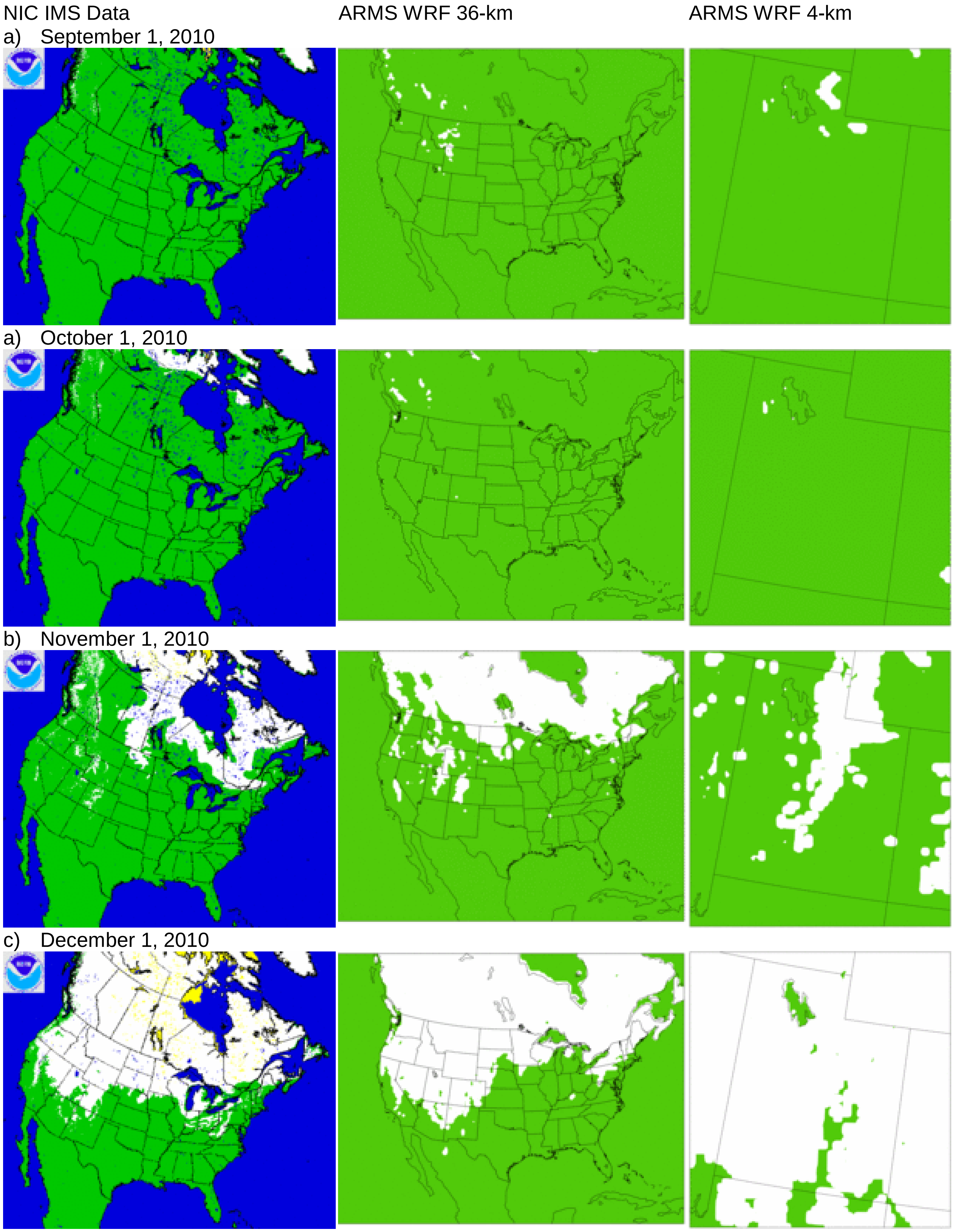


Figure 4-25 Snow Cover for September 1, October 1, November 1, and December 1, 2010

4.2 Annual and Seasonal Model Performance Evaluation

4.2.1 Domain-wide Model Performance

Domain-wide statistics were generated to provide insight regarding the ARMS WRF annual model performance, as well as information regarding how the performance varies for different seasons and domains. The statistical analysis performed for the 36-km domain only assesses for the spatial area within the 12-km domain since model performance outside of the 12-km domain is not of interest for the study. In addition to domain-wide statistical summaries, statistical evaluations also were conducted for the Uinta Basin Study Area.

The seasons have a non-standard definition for this study due to the separation of model configurations in winter and non-winter months, as explained in Chapter 2.0. For this study, the months of January through March were grouped; April through June; July through September, and October through December. **Tables 4-1**, **4-2**, **4-3**, and **4-4** show the key statistical results for the ARMS WRF 4-km domain, the 12-km grid resolution within the 4-km domain, the 12-km domain, and the 36-km domain grid resolution within the 12-km domain, respectively. The results are presented in this way to isolate whether the results differ due to different model resolutions or the geographic extent of the domains. For example, differences between the data in **Tables 4-1** and **4-2** are entirely due to different model grid resolutions, likewise differences between the data in **Tables 4-3** and **4-4** also are entirely due to different model grid resolutions. Differences between the data in **Tables 4-2** and **4-3** are due to model performance differences for different geographic areas. Statistical performance benchmarks also are provided in each table to provide context for the model performance and allow for an initial identification of potential problems with the ARMS WRF configuration.

As shown in **Table 4-1**, the ARMS WRF 4-km simulation was within recommended performance benchmarks for all meteorological parameters, with the exception of the average wind speed bias during January-March. Often the performance was within the more restrictive benchmarks established by Tesche et al. (2002). In general, the simulation under-predicts wind speed, as shown by a negative wind speed bias, for the full annual simulation as well for each season. The ARMS WRF 4-km domain had the most difficulty reproducing wind direction, which is likely due to complex terrain throughout much of the domain. While the average wind direction bias is relatively low, the gross errors consistently exceed the benchmarks recommended by Tesche et al. (2002), but are within the recommended benchmarks for complex terrain. The 4-km domain performs very well for temperature with very low biases, reasonably small gross errors, and high IOAs. There is a tendency to under-predict temperature during January-March with slight over-predictions in other months. The WRF mixing ratio performed reasonably well and was always within the benchmarks established by Tesche et al. (2002).

A comparison between ARMS WRF 4-km grid resolution (**Table 4-1**) and 12-km grid resolution (**Table 4-2**) shows that the 12-km grid resolution performance is similar to the 4-km grid resolution (with a few exceptions), which is anticipated due to the nested grid configuration. The performance between the two grid resolutions is similar with the exception of wind speed bias and wind direction bias. The wind speed bias was generally better (i.e. lower) for the 4-km grid resolution than the 12-km grid resolution. The wind direction bias was generally better (i.e. lower) for the 12-km grid resolution than the 4-km grid resolution; however, the wind direction errors are slightly smaller for the 4-km domain.

Table 4-1 WRF 4-km Domain Model Performance

Parameter	Statistics	Statistical Benchmark		Average Values				
		(Tesche et al. 2002)	Complex Terrain	Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Wind speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.93***	***1.71***	*2.11*	*2.10*	***1.80***
	Bias	≤ ±0.5		**-0.39**	-0.80	**-0.19**	**-0.11**	**-0.46**
	IOA	≥ 0.6		**0.73**	**0.69**	**0.79**	**0.72**	**0.72**
Wind direction (deg)	Bias	≤ ±10		**1.52**	**1.64**	**1.25**	**0.98**	**2.21**
	Gross Error	≤ 30	≤ 55	*38.70*	*38.04*	*37.76*	*40.86*	*38.15*
Temperature (K)	Bias	≤ ±0.5	≤ ±2	***0.09***	***-0.23***	***0.01***	***0.21***	***0.38***
	Gross Error	≤ 2	≤ 3.5	***1.69***	***1.96***	***1.53***	***1.76***	***1.51***
	IOA	≥ 0.8		**0.96**	**0.93**	**0.97**	**0.96**	**0.95**
Mixing ratio (g/kg)	Bias	≤ ±1	≤ ±1	***0.03***	***-0.08***	***0.36***	***-0.07***	***-0.11***
	Gross Error	≤ 2	≤ 2	***0.81***	***0.55***	***0.85***	***1.22***	***0.61***
	IOA	≥ 0.6		**0.76**	**0.75**	**0.73**	**0.77**	**0.77**

Bold indicates passing benchmark for Tesche et al. 2002. Italics indicates passing benchmark for complex terrain.

Table 4-2 WRF 12-km Resolution Model Performance Within 4-km Domain

Parameter	Statistics	Statistical Benchmark		Average Values				
		(Tesche et al. 2002)	Complex Terrain	Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Wind speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.92***	***1.82***	*2.05*	***1.99***	***1.83***
	Bias	≤ ±0.5		-0.55	-1.02	**-0.30**	**-0.34**	-0.53
	IOA	≥ 0.6		**0.72**	**0.67**	**0.79**	**0.73**	**0.72**
Wind direction (deg)	Bias	≤ ±10		**0.47**	**0.12**	**0.68**	**-0.57**	**1.66**
	Gross Error	≤ 30	≤ 55	*41.09*	*40.90*	*38.91*	*43.05*	*41.51*
Temperature (K)	Bias	≤ ±0.5	≤ ±2	***-0.16***	***-0.35***	***-0.38***	***-0.08***	***0.15***
	Gross Error	≤ 2	≤ 3.5	***1.78***	*2.06*	***1.65***	***1.84***	***1.59***
	IOA	≥ 0.8		**0.95**	**0.93**	**0.97**	**0.96**	**0.94**
Mixing ratio (g/kg)	Bias	≤ ±1	≤ ±1	***-0.03***	***-0.08***	***0.27***	***-0.15***	***-0.14***
	Gross Error	≤ 2	≤ 2	***0.78***	***0.54***	***0.80***	***1.19***	***0.60***
	IOA	≥ 0.6		**0.76**	**0.76**	**0.75**	**0.77**	**0.77**

Bold indicates passing benchmark for Tesche et al. 2002. Italics indicates passing benchmark for complex terrain.

Table 4-3 WRF 12-km Domain Model Performance

Parameter	Statistics	Statistical Benchmark (Tesche et al. 2002)	Statistical Benchmark Complex Terrain	Average Values Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Wind speed (m/s)	RMSE	≤ 2	≤ 2.5	*2.18*	*2.03*	*2.37*	*2.09*	*2.23*
	Bias	≤ ±0.5		**-0.01**	-0.51	**0.17**	**0.20**	**0.08**
	IOA	≥ 0.6		**0.74**	**0.73**	**0.77**	**0.73**	**0.74**
Wind direction (deg)	Bias	≤ ±10		**3.37**	**2.40**	**4.04**	**3.50**	**3.52**
	Gross Error	≤ 30	≤ 55	*45.2*	*45.82*	*42.37*	*47.27*	*45.35*
Temperature (K)	Bias	≤ ±0.5	≤ ±2	***-0.03***	***-0.32***	***-0.17***	***0.09***	***0.30***
	Gross Error	≤ 2	≤ 3.5	*2.16*	*2.22*	*2.00*	*2.25*	*2.16*
	IOA	≥ 0.8		**0.96**	**0.97**	**0.97**	**0.96**	**0.96**
Mixing ratio (g/kg)	Bias	≤ ±1	≤ ±1	***0.01***	***-0.10***	***0.13***	***-0.01***	***0.04***
	Gross Error	≤ 2	≤ 2	***0.97***	***0.67***	***1.01***	***1.45***	***0.74***
	IOA	≥ 0.6		**0.86**	**0.87**	**0.86**	**0.86**	**0.86**

Bold indicates passing benchmark for Tesche et al. 2002. Italics indicates passing benchmark for complex terrain.

Table 4-4 WRF 36-km Resolution Model Performance Within the WRF 12-km Domain

Parameter	Statistics	Statistical Benchmark (Tesche et al. 2002)	Statistical Benchmark Complex Terrain	Average Values Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Wind speed (m/s)	RMSE	≤ 2	≤ 2.5	*2.18*	*2.11*	*2.35*	*2.05*	*2.22*
	Bias	≤ ±0.5		**-0.22**	-0.69	**-0.02**	**-0.10**	**-0.07**
	IOA	≥ 0.6		**0.73**	**0.70**	**0.77**	**0.72**	**0.72**
Wind direction (deg)	Bias	≤ ±10		**5.92**	**5.27**	**5.75**	**5.97**	**6.71**
	Gross Error	≤ 30	≤ 55	*47.82*	*49.13*	*44.23*	*49.63*	*48.28*
Temperature (K)	Bias	≤ ±0.5	≤ ±2	***-0.39***	*-0.58*	*-0.65*	***-0.29***	***-0.03***
	Gross Error	≤ 2	≤ 3.5	*2.33*	*2.36*	*2.22*	*2.42*	*2.30*
	IOA	≥ 0.8		**0.96**	**0.96**	**0.96**	**0.95**	**0.96**
Mixing ratio (g/kg)	Bias	≤ ±1	≤ ±1	***0.03***	***-0.12***	***0.14***	***0.02***	***0.06***
	Gross Error	≤ 2	≤ 2	***0.94***	***0.65***	***0.97***	***1.41***	***0.72***
	IOA	≥ 0.6		**0.87**	**0.88**	**0.88**	**0.86**	**0.88**

Bold indicates passing benchmark for Tesche et al. 2002. Italics indicates passing benchmark for complex terrain.

As shown in **Table 4-3**, the ARMS WRF 12-km simulation was within recommended performance benchmarks for all meteorological parameters, with the exception of the average wind speed bias during January-March. Often the performance was within the more restrictive benchmarks established by Tesche et al. (2002), except for the wind speed RMSE, wind direction gross error, and temperature gross error. All three of these statistical parameters are within the recommended performance benchmarks for complex terrain for all seasons. In general, the simulation does not have systematic biases (i.e., a parameter biased high during one season is biased low during a different season), with the exception of wind direction, which the model tends to over-predict by a few degrees. Similar to the 4-km domain, the 12-km domain simulation had the most difficulty reproducing wind direction, which is likely due to complex terrain throughout much of the domain. The wind direction gross errors are approximately 8-9 degrees higher in the 12-domain than the 4-km domain (comparing **Tables 4-1** and **4-3**). While some of this difference is due to grid resolution (as discussed when comparing **Tables 4-1** and **4-2**), most of is due to geographic differences between the 4-km domain and the 12-km domain. The 12-km domain performs reasonably well for temperature with very low biases, average gross errors of approximately 2 degrees K, and high IOAs. The WRF mixing ratio performed reasonably well and was always within the benchmarks established by Tesche et al. (2002).

As expected, the 36-km grid resolution performance (shown in **Table 4-4**) is fairly similar to the 12-km grid resolution (shown in **Table 4-3**), but with slightly higher errors and larger biases.

As described in Chapter 3, Bakergram plots are included to supplement the tabular statistical summaries. **Figures 4-26** and **4-27** show Bakergram plots of the daily average statistics for the 4-km and 12-km domains, respectively. While **Table 4-1** summarizes seasonal and annual statistical averages, **Figure 4-26** shows the daily average statistics and visually represents seasonal patterns. All Bakergram plots present information in a similar order. From left to right, the top row shows bias for wind speed, wind direction, temperature, and mixing ratio. The second row shows wind speed RMSE, wind direction gross error, temperature gross error, and mixing ratio gross error. The third row shows the IOA for wind speed, temperature, and mixing ratio.

As shown in **Figure 4-26**, the wind speed bias Bakergram shows that the ARMS WRF 4-km performance is within the performance benchmark for most days of the year (indicated with the gray color) and tends to under-predict wind speed during the winter and spring (indicated with shades of blue). Similar to the wind speed bias, the wind speed RMSE Bakergram shows that the ARMS WRF 4-km domain performed well for at least half of the year, with a tendency to have larger errors in summer months. For most of the year, the WRF wind speed IOA performance is better than the benchmark. During most of the year the wind direction bias is within model performance benchmarks. However, the wind direction gross error often exceeded the Tesche et al. (2002) benchmarks, but is generally within the benchmarks for complex terrain. The wind direction gross error is highest in July and August. Temperature biases suggest that the ARMS WRF performed well in summer, with a small cold bias in the spring and warm bias in fall. The temperature IOA indicates that WRF simulation performed well throughout the full annual simulation. Mixing ratio bias and error suggests that the ARMS WRF simulation performance is better in winter months than in summer months. For most of the year, the mixing ratio IOA performance is within the statistical benchmarks.

As shown in **Figure 4-27**, the wind speed bias Bakergram shows that the ARMS WRF 12-km performance is within the performance benchmark for most days of the year (indicated with the gray color) and tends to under-predict wind speed during the winter and spring (indicated with shades of blue) with a small over-prediction during the shoulder seasons (indicated with shades of yellow). Similar to the results shown in **Table 4-3**, the wind speed RMSE Bakergram shows that the ARMS WRF 12-km domain exceeds the benchmarks established by Tesche et al. (2002), but is often within the complex terrain benchmarks. There is a tendency to have larger wind speed errors in summer months. The WRF wind speed IOA performance is better than the benchmark throughout the full annual simulation. During most of the year the wind direction bias is within model performance benchmarks. However, the wind direction gross error often exceeded the Tesche et al. (2002) benchmarks, but is generally within the benchmarks for complex terrain. The wind direction gross error is highest in July and August, but also has high errors during days in winter. Temperature biases suggest that the ARMS WRF performed well overall, with a small cold bias in the spring and warm bias in fall. The temperature IOA indicates that WRF simulation performed well throughout the full annual simulation. Mixing ratio bias and error suggests that the ARMS WRF simulation performance is better in winter months than in summer months. The mixing ratio IOA performance is within the statistical benchmarks throughout the full annual simulation.

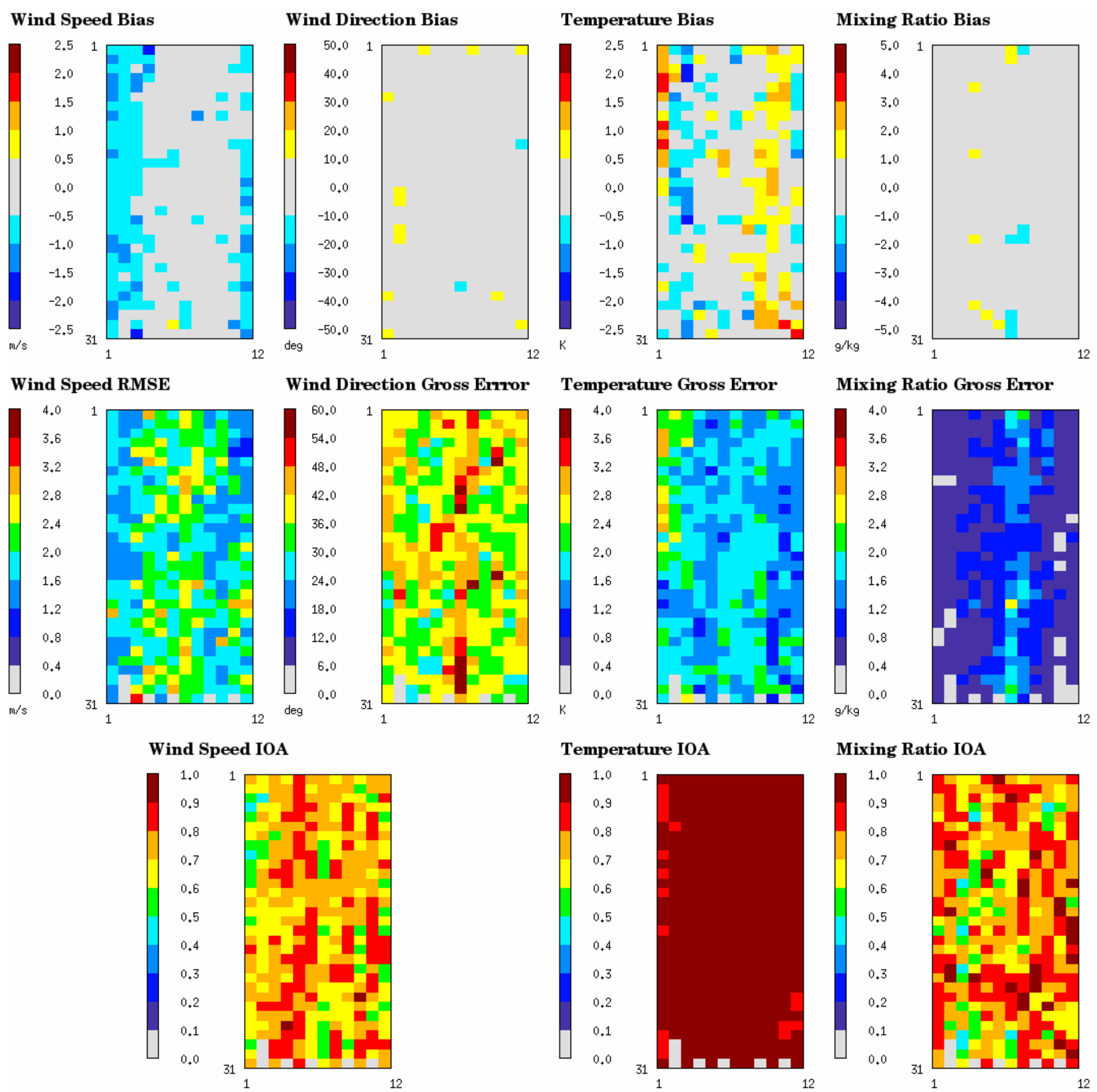


Figure 4-26 WRF 4-km Domain Bakergrams

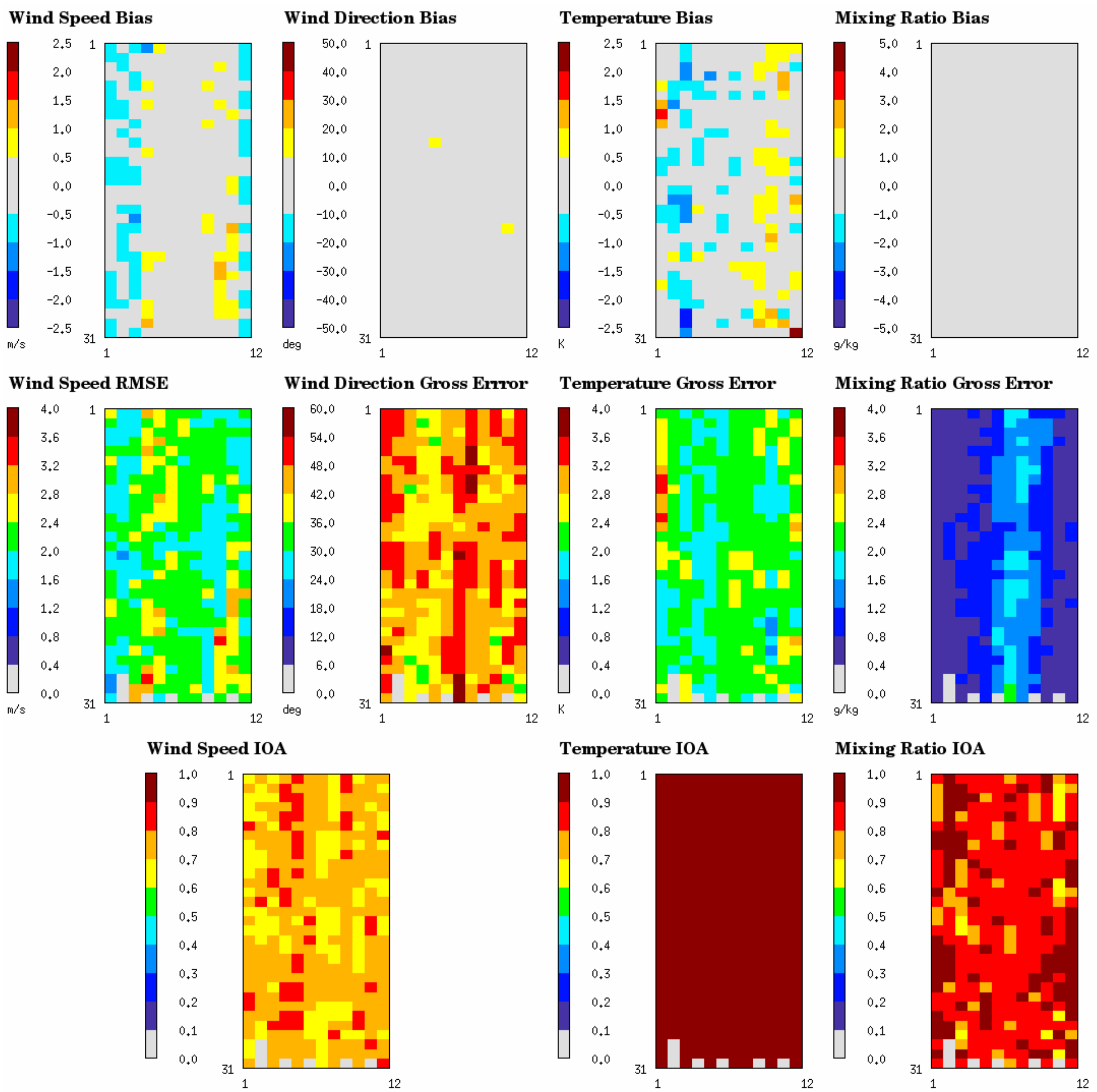


Figure 4-27 WRF 12-km Domain Bakergrams

In general, model bias and error were slightly better in the 4-km domain than for the 12-km domain; however, the IOA for all parameters was better in the 12-km domain than the 4-km domain.

4.2.2 Uinta Basin Model Performance

The ARMS WRF model performance was evaluated in more detail by comparing WRF 4-km grid resolution results to observations located in the Uinta Basin. By evaluating the model performance within the Uinta Basin Study Area, the results can be evaluated in the context of the ARMS study objectives and additional information will be available to assess the performance of the air quality model. Similar to the analysis of domain-wide performance, the 4-km annual average and seasonal average statistics for the Uinta Basin Study Area are tabulated in **Table 4-5**. In addition to tabular summaries, **Figure 4-28** presents Bakergram plots presenting daily average model performance statistics for the Uinta Basin Study Area.

Table 4-5 WRF 4-km Resolution Model Performance Within the Uinta Basin Study Area

Parameter	Statistics	Statistical Benchmark		Average Values				
		(Tesche et al. 2002)	Complex Terrain	Annual	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Wind speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.52***	***1.14***	***1.94***	***1.74***	***1.26***
	Bias	≤ ±0.5		**-0.42**	**-0.49**	-0.54	**-0.41**	**-0.25**
	IOA	≥ 0.6		**0.68**	0.56	**0.75**	**0.74**	**0.65**
Wind direction (deg)	Bias	≤ ±10		**1.26**	**3.24**	**-3.18**	**0.15**	**4.85**
	Gross Error	≤ 30	≤ 55	*48.56*	*53.93*	*46.87*	*43.56*	*49.89*
Temperature (K)	Bias	≤ ±0.5	≤ ±2	***0.47***	*1.12*	***-0.01***	***0.37***	***0.38***
	Gross Error	≤ 2	≤ 3.5	***1.44***	*2.24*	***1.17***	***1.18***	***1.18***
	IOA	≥ 0.8		**0.91**	**0.82**	**0.97**	**0.98**	**0.87**
Mixing ratio (g/kg)	Bias	≤ ±1	≤ ±1	***0.15***	***0.30***	***0.76***	***-0.19***	***-0.28***
	Gross Error	≤ 2	≤ 2	***0.77***	***0.56***	***0.96***	***1.02***	***0.56***
	IOA	≥ 0.6		0.54	0.55	0.47	0.55	0.57

Bold indicates passing benchmark for Tesche et al. 2002. Italics indicates passing benchmark for complex terrain.

As shown in **Table 4-5**, in the Uinta Basin the ARMS WRF 4-km simulation was within recommended performance benchmarks for all meteorological parameters, with the exception of the wind speed bias during April-June, the wind speed IOA during January-March and the mixing ratio IOA throughout the full year. Often the performance was within the more restrictive benchmarks established by Tesche et al. (2002). In general, the simulation under-predicts wind speed, as shown by a negative wind speed bias, for the full annual simulation as well for each season. The model simulation had difficulty reproducing wind direction, which is likely due to complex terrain throughout much of the domain. While the average wind direction bias is relatively low, the gross errors consistently exceed the benchmarks recommended by Tesche et al. (2002), but are within the recommended benchmarks for complex terrain. The model performs reasonably well for temperature except during January- March when the model tends to over-predict temperature and has larger gross errors. The model simulation in the Uinta Basin Study Area had the most difficulty reproducing mixing ratio.

As shown in **Figure 4-28**, the wind speed bias Bakergram shows that the ARMS WRF 4-km performance in the Uinta Basin Study Area is within the performance benchmark for most days of the year (indicated with the gray color) and tends to under-predict wind speed (indicated with shades of blue) independent of season. The wind speed RMSE Bakergram shows that the ARMS WRF 4-km performance in the Uinta Basin was best in the winter with larger errors in summer months. For most of the year, the WRF wind speed IOA performance is better than the benchmark though the performance is generally better in the summer. During most of the year the wind direction bias is within model performance benchmarks. However, the wind direction gross error often exceeded the Tesche et al. (2002) benchmarks and even exceeds the benchmarks for complex terrain, particularly during winter months. Temperature biases suggest that the ARMS WRF performed well in summer, but with notable

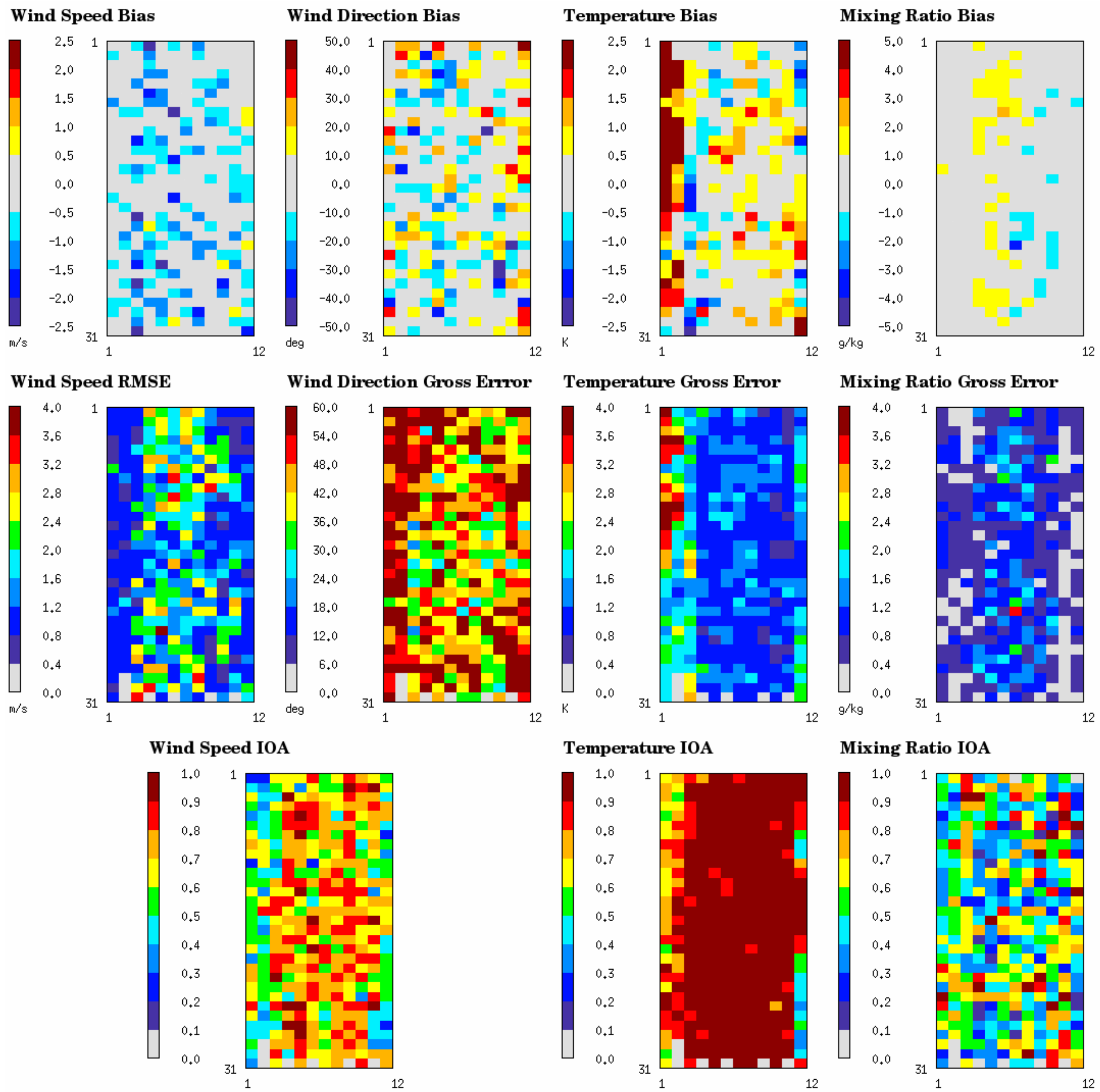


Figure 4-28 WRF 4-km Resolution Bakergrams Limited to the Uinta Basin Study Area

warm biases and larger errors in January and February. The temperature IOA indicates that WRF simulation performed well in the Uinta Basin during most seasons except for winter. Mixing ratio bias and error suggests that the ARMS WRF simulation performance is better in winter months than in summer months. Results for mixing ratio IOA generally show worse performance than the benchmark and no seasonal variation.

In general, model performance within the Uinta Basin Study Area was not as good as the performance over the full 4-km domain: this difference is particularly evident during winter. However, the WRF performance in the Uinta Basin is still within recommended benchmarks for most days and meteorological parameters, except for mixing ratio.

5.0 Conclusions

The purpose of the WRF modeling is to develop a gridded meteorological dataset that is appropriate to assess the potential air quality impacts for the ARMS Modeling Project. In support of this, the WRF configuration was tested extensively for the Uinta Basin Study Area (results are documented in **Appendix A**) to determine a preferred WRF configuration for conducting an annual WRF simulation. The result of these test led to two configurations: one for winter months and another for non-winter months. The primary differences between the two configurations are the PBL scheme (ACM2 scheme vs. MYJ), the microphysics scheme (Lin et al. scheme vs. WRF single moment 3-class), the short-wave radiation scheme (Dudhia vs. Goddard), and the land surface model (Pheim-Xiu versus Noah Land-Surface Model with Multi-Physics Options).

A qualitative and quantitative operational MPE was conducted on the annual WRF simulation. The operational MPE compared model results to observations using various graphical plots and statistical measures. The comparisons and statistical results were evaluated, over different temporal and spatial extents, to assess the WRF model's performance for accuracy, consistency, and reasonableness with respect to available observations. Statistical summaries were generated for the 4-km, 12-km, and 36-km model domains with a focus on the assessment of the 4-km results. In addition, the model performance was evaluated exclusively for the Uinta Basin Study Area to provide additional information about the area of interest for the ARMS study. The datasets used to evaluate WRF performance include TDL hourly observational datasets, NWS analysis maps, and vertical profiles of the atmosphere.

Based on the model performance evaluation, the 2010 annual ARMS WRF modeling simulation demonstrated good performance and is considered suitable for use as input to an air quality model for the ARMS Project. The WRF simulation was capable of reproducing the observed synoptic, precipitation, and snow cover patterns indicating proper placement of large-scale systems and overall good performance of model physics.

In general, the 2010 annual simulation passes the statistical benchmarks for both non-complex and complex terrain. In addition, the model results were slightly better for the 4-km domain than the 12-km domain likely due to both the finer resolution grid and the use of observation nudging in the 4-km domain. On an annual and seasonal basis, most meteorological parameters were within the traditional performance benchmarks. Moreover, when the results are evaluated relative to updated performance benchmarks for complex terrain, all results for the 4-km domain are within the accepted range.

The WRF 4-km domain results were further evaluated for each day of the year with annual Bakergrams. Bakergrams of bias, error, and IOA were plotted for each meteorological parameter in the 4-km domain. An analysis of the Bakergrams indicates:

- Little or no wind speed bias during most of the year with a slight low bias during winter.
- Little or no wind direction biases throughout the full annual simulation, but slightly higher gross errors in the summer.
- Temperature performed well in summer, but has a cold bias in spring and a warm bias in fall.
- Relative humidity performed well throughout the annual simulation, with slightly higher errors in summer.

In addition, model performance in the Uinta Basin Study Area was assessed. Analysis of model performance within the Uinta Basin shows:

- Wind speed performed well for most of the year. Results tend to have a slight low bias independent of season and slightly higher errors in the summer.

- Wind direction performed reasonably well during most of the year, but with larger errors during winter. Results tend to be under-predicted more frequently during summer months and over-predicted in winter months.

- Temperature bias, gross error and IOA results are within the benchmark for the majority of days, and suggest that WRF simulation performed better in summer than in winter.

- Mixing ratio bias and gross error show that the WRF simulation performance is within the model performance benchmark for most of the year. However, mixing ratio IOA is often below performance benchmarks.

Quantitative analysis of the meteorological fields demonstrated reasonable and generally consistent performance when compared to observations as analyzed over various levels of time and space. Potential areas of model error were identified, such as the tendency to under-predict wind speed and wind direction, and the biases in temperature during winter. These model tendencies and errors will be considered when evaluating the air quality impacts during the air quality MPE.

6.0 References

AECOM. 2012a. Utah Air Resource Management Strategy Air Quality Modeling and Assessment Protocol. Prepared for the Bureau of Land Management, Utah State Office. January 2012.

AECOM. 2012b. Memorandum on the Proposed WRF Modeling Approach for the ARMS Project from Zion Wang (AECOM) to Leonard Herr (Utah BLM). May 2012.

Energy Dynamics Lab. 2011. Uinta Basin Winter Ozone and Air Quality Study Final Report. June 2011.

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, by ENVIRON International Corporation, Novato, California.

Grell, G. A., J. Dudhia, and D. R. Stauffer. 1994. A Description of the Fifth Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Technical Note NCAR TN-398-STR. 138 pp.

Janjic, Z. I. and J. P. Gerrity, Jr. 2001. An Alternative Approach to Nonhydrostatic Modeling. Monthly Weather Review 129:1164-1178.

Kemball-Cook, S., Y. Jia, C. Emery, R. Morris, Z. Wang, G. Tonnesen. 2005. Alaska MM5 Modeling for the 2002 Annual Period to Support Visibility Modeling. Prepared for the Western Regional Partnership, Denver, Colorado. Prepared by ENVIRON International Corporation, Novato, California, and the University of California at Riverside, Center for Environmental Research and Technology. September 2005.

National Center for Atmospheric Research (NCAR). 2009. Mesoscale and Microscale Meteorology Division, ARW Version 3 Modeling System User's Guide. Internet website: http://www.mmm.ucar.edu/wrf/ users/docs/user_guide_V3.1/ARWUsersGuideV3.pdf. Accessed on July 2009.

National Oceanic and Atmospheric Administration (NOAA). 2011. NOAA/ESRL Radiosonde Database. Internet website: http://www.esrl.noaa.gov/raobs/. Accessed May 12, 2011.

National Oceanic and Atmospheric Administration (NOAA). 2013. Interactive Multisensor Snow and Ice Mapping System (IMS). United States National Ice Center. Internet website: http://www.natice.noaa.gov/ims/ims.html. Accessed January 24, 2013.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. Gill, D. Barker, M. G. Duda, X. Y. Huang, W. Wang, and J. G. Powers. 2008. A Description of the Advanced Research WRF Version 3. NCAR Technical Note NCAR/TN-475+STR. June 2008.

Tesche, T. W., D. E. McNally, and C. Tremback. 2002. Operational Evaluation of the MM5 Meteorological Model over the Continental United States: Protocol for Annual and Episodic Evaluation. Prepared for P. Dolwick USEPA Office of Air Quality Planning and Standards. July 2002.

University Corporation for Atmospheric Research (UCAR). 2012. WRF Model Downloads. Internet website: http://www.mmm.ucar.edu/wrf/users/downloads.html. Accessed May 2012.

Appendix A

Sensitivity Tests and Results

ARMS Weather Research and Forecasting Sensitivity Tests

This appendix summarizes the results of the Utah Bureau of Land Management (BLM) Weather Research and Forecasting (WRF) sensitivity tests, which were conducted to develop the final WRF configuration for the ARMS study. The Utah Department of Environmental Quality, Division of Air Quality (UDAQ) conducted WRF modeling for the State of Utah and this study leveraged extensively from their modeling efforts. The Base Case sensitivity test used the same modeling domains, vertical layers, physics options, and pre-processing methods as UDAQ. One exception is that a more recent version of the WRF model (version 3.4) was used for this study.[2]

The Base Case model results were compared to eight sensitivity tests designed to target improved model performance during winter months in the Uinta Basin. The configuration options for the Base Case and sensitivity tests are shown in **Tables A-1** and **A-2**, respectively. The different model configurations were tested for the months of February and July. February and July were selected for evaluation because the meteorological conditions during those months are markedly different and the optimal model configurations may differ accordingly. Conditions in February may include cold-pool stagnation events, while summer months are generally characterized by diurnal patterns of mountain-valley flow and afternoon thunderstorms.

The initial and lateral boundary conditions for the 36-km domain were identical to those used by UDAQ, which were extracted from the 12-km North American Model (NAM) archives (originally from the NOAA National Operational Model Archive and Distribution System [NOMADS] maintained by the National Climate Data Center). Analysis nudging was used on the 36-km, 12-km, and 4-km domains for temperature, winds, and humidity. No analysis nudging was applied in the boundary layer.

The Base Case and first five sensitivity simulations (Sens0 through Sens4) used observation nudging with the Meteorological Assimilation Data Ingest System (MADIS) datasets provided by Utah Department of Environmental Quality (UDEQ) and two observation sites located in the Uinta Basin (Ouray and Redwash) provided by the Utah BLM. Sensitivity Tests 5, 6, and 7 (Sens5 though Sens7) used an updated observation nudging dataset.

The results of the sensitivity tests were evaluated by comparing model-predicted values to measured meteorological conditions using the analyses described in Chapter 3.0, which included assessment of the model performance in the Uinta Basin Study Area. The best performing WRF configuration from all tests was used for the final annual simulation.

Based on the results of the sensitivity tests, it is found that the best overall model performance in the Uinta Basin Study Area was produced using:

- Different model physics for winter and non-winter months;
- Enhanced vertical resolution in the surface layer;
- A revised observation nudging dataset, which was processed without buddy checks; and
- Surface observation nudging was not used for mixing ratio.

[2] Note that UDAQ used WRF version 3.2. It is not anticipated that significant changes will result from using a more recent version of WRF.

Table A-1 WRF Base Case Model Configurations Developed by UDAQ

Physics, Dynamics, and Boundary Options	36-km Grid	12-km Grid	4-km Grid
Microphysics	Lin et al., scheme	Lin et al., scheme	Lin et al., scheme
Cumulus Parameterization	Kain-Fritsch scheme	Kain-Fritsch scheme	None
PBL	ACM2 scheme	ACM2 scheme	ACM2 scheme
Surface Layer	Pleim-Xiu scheme	Pleim-Xiu scheme	Pleim-Xiu scheme
Land Surface Model	Pleim-Xiu scheme	Pleim-Xiu scheme	Pleim-Xiu scheme
Long-wave Radiation	Rapid Radiative Transfer Model (RRTM)	RRTM	RRTM
Short-wave Radiation	Dudhia	Dudhia	Dudhia

Table A-2 WRF Model Configurations for Sensitivity Simulations[1]

Test Name	Number of Vertical Levels	Modeling Domains for 12 km & 4 km Nests[2]	Microphysics	Cumulus Scheme[3]	PBL Scheme	Surface Layer Scheme	Land Surface Model	Soil Nudging	Long Wave/Short Wave Schemes
Base Case	34	See Table A-3	Lin et al., scheme	Kain-Fritsch scheme	ACM2 scheme	Pleim-Xiu scheme	Pleim-Xiu scheme	N/A	Rapid Radiative Transfer Model (RRTM/ Dudhia)
Sens0	34	Same as Base Case	Single moment (6-class)	Grell-Devenyi Ensemble Scheme[3]	Mellor-Yamada-Janic (MYJ) scheme	Monin-Obukov (Janic) scheme	Noah-MP Land Surface Model	N/A	RRTM/ Goddard
Sens1	34	Same as Base Case	Single moment (6-class)	Grell-Devenyi Ensemble Scheme[3]	Mellor-Yamada-Janic (MYJ) scheme	Monin-Obukov (Janic) scheme	Unified Noah Land Surface Model	N/A	RRTM/ Goddard
Sens2	34	Same as Base Case	Morrison (2 moments)	Grell-Devenyi Ensemble Scheme[3]	QNSE-EDMF scheme	QNSE scheme	Unified Noah Land Surface Model	N/A	RRTM/ Goddard
Sens3	34	Expanded in the north direction (see Table A-3)	Physics configuration same as Base Case for February and Sens1 for July						
Sens4	42	Same as Base Case	Physics configuration same as Base Case for February and Sens1 for July						
Sens5	36 (see Table 2-1)	Revised (see Table 2-3)	Physics configuration same as Base Case for February, Revised Observation Nudging Dataset						

Table A-2 WRF Model Configurations for Sensitivity Simulations[1]

Test Name	Number of Vertical Levels	Modeling Domains for 12 km & 4 km Nests[2]	Microphysics	Cumulus Scheme[3]	PBL Scheme	Surface Layer Scheme	Land Surface Model	Soil Nudging	Long Wave/Short Wave Schemes
Sens6	36 (see Table 2-1)	Revised (see Table 2-3)	Physics configuration same as Base Case for February, Revised Observation Nudging Dataset and Buddy Checks Turned Off						
Sens7	36 (see Table 2-1)	Revised (see Table 2-3)	Physics configuration same as Base Case for February and Sens1 for July, Revised Observation Nudging Dataset, Buddy Checks and RH Nudging Turned Off						

1 Differences from the Base Case model configuration are shown in red text.

2 The 36-km modeling domain is identical for all tests.

3 Cumulus parameterization is used in the 36-km and 12-km domains, but not the 4-km domain.

Description of the Sensitivity Test Configurations

The first sensitivity test (referred to as Sens0) is based off of a set of physics used by the Polar Modeling Group at Ohio State University.[3] The Polar Modeling Group has conducted extensive tests of the WRF model in the Arctic. Based on this work, the Mellor-Yamada-Janic (MYJ) planetary boundary layer (PBL) scheme is coupled with the Monin-Obukov (Janic) similarity surface layer parameterization. The Rapid Radiative Transfer Model (RRTM) scheme and the Goddard shortwave scheme are used for longwave and shortwave radiation, respectively.

Importantly, the land surface model (LSM) is different between Sens0 and the preferred configuration of the Polar Modeling Group: Sens0 uses the newly-released Noah-MP LSM, rather than the Unified Noah LSM used in the Polar version 3.1.1 of WRF.[2] The Noah-MP LSM was selected for Sens0 due to the improved snow parameterizations; however, this configuration did not ultimately demonstrate improved performance for this study. The Noah-MP LSM parameterizes a multi-layer snow pack with liquid water storage and melt/refreeze capability, as well as a snow-interception model describing loading/unloading, melt/refreeze, and sublimation of the canopy-intercepted snow.

The second sensitivity test (referred to as Sens1) used the same set of physics options Polar Modeling Group. Sens1 is the same as the Sens0 run, except that the Noah-MP LSM was replaced by the Unified Noah LSM.

The third sensitivity test (referred to as Sens2) used the revised Quasi-Normal Scale Elimination Eddy-Diffusivity-Mass-Flux (named QNSE-EDMF) scheme for parameterizing the PBL. This PBL scheme was coupled with the QNSE surface layer scheme. The QNSE scheme is selected for Sens2 because the QNSE scheme is designed for stably stratified conditions, like those that occur in the winter in the Uinta Basin. During daylight hours the QNSE PBL scheme uses a mass flux method with shallow convection.

The results from the Base Case and Sens0 through Sens2 were used to inform the selection of the physics configuration before conducting the following sensitivity tests. It was found that the Base Case physics configuration performed best for February and Sens1 performed best for July.

The fourth sensitivity test (referred to as Sens3) tests a set of modeling domains that were expanded to account for complex terrain upwind of the domain boundaries. The different domains are shown in **Figure A-1** and the specifications are listed in **Table A-3**.

The fifth sensitivity test (referred to as Sens4) tested the effects of additional vertical resolution in the surface layer and in the upper troposphere/lower stratosphere. The Base Case and Sens0 through Sens4 were run with 34 vertical layers, while Sens4 tests 42 vertical layers.

After completion of these tests, the model still had notable temperature biases of +10K in the Uinta Basin during several days in February. Three additional sensitivity runs were performed to diagnosis the issue and improve model performance. The following three sensitivity tests focused exclusively on model performance in February. At the conclusion, July performance was tested to confirm the model performance had not degraded for summer months. For these three sensitivity tests, the WRF domains were adjusted slightly.

The sixth sensitivity simulation (referred to as Sens5) tests a revised observation nudging dataset.

[3] Wilson, A.B., Bromwich, D.H., and K.M. Hines. 2011. Evaluation of Polar WRF Forecasts on the Arctic System Reanalysis Domain: Surface and Upper Air Levels. Journal of Geophysical Research. 116. June 2011.

The sensitivity test 7 (referred to as Sens6) repeated the Sens5 configuration but turned off the "Buddy Check" QC Test in the OBSGRID processor. The Buddy Check QC Test criteria was rejecting observed temperature data, preventing the observations from being used in WRF's observation nudging processes.

Based on the results from Sens6, one final WRF sensitivity test (referred to as Sens7) was performed with Sens6 configuration but disabling the observational nudging of mixing ratio (relative humidity) within WRF 4-km domain.

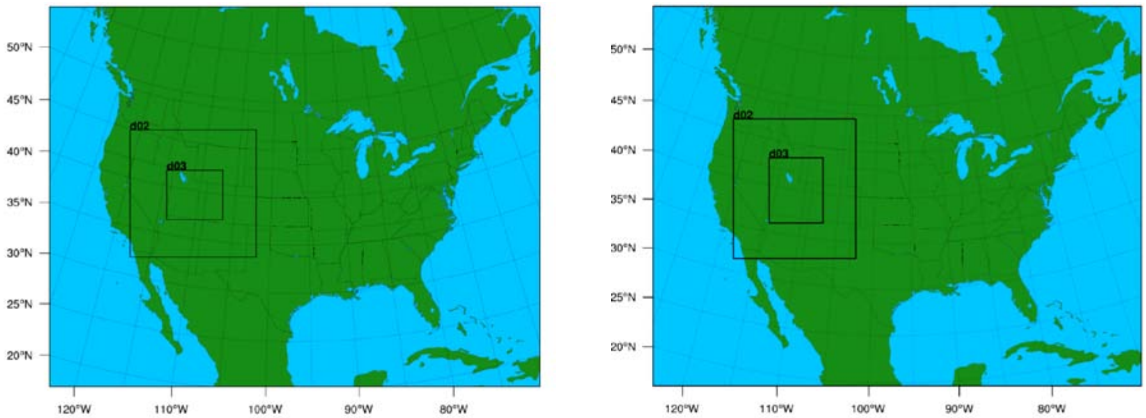


Figure A-1 Modeling Domains for the Base Case (left) versus Sens3 (right)

Table A-3 Model Specifications for Base Case and Sens3

Domain	Base Case, Sens0, Sens1, and Sens2 Number of Grid Cells	Sens3 Number of Grid Cells	Coordinates of southwestern corner of grid (km)
36-km domain	(165, 129)	(165, 129)	(-2952, -2304)
12-km domain	(124, 124)	(124, 142)	(-1980, -756)
4-km domain	(157, 139)	(157, 196)	(-1560, -324)

Since the objective of the sensitivity tests is to determine the preferred WRF model configuration for the ARMS Modeling Study, the results and associated discussions is focused exclusively on model performance statistics within the 4-km domain or the Uinta Basin Study Area.

Test Results: Base Case, Sensitivity Test 0, and Sensitivity Test 1

The model performance statistics for Base Case, Sens0, and Sens1 tests are shown in **Figures A-2** and **A-3**, for February and July, respectively. These figures compare the daily average biases for wind speed, wind direction, temperature, and relative humidity across 4-km domain. The daily biases for Base Case, Sens0, and Sens1 are shown in black, red, and blue, respectively.

It is clear from these bias comparisons that Sens0 performed better than the Base Case and Sens1 for wind speed and slightly better for wind direction. However, the temperature and mixing ratio performance for Sens0 was significantly degraded relative to Base Case and Sens1 results, particularly in July. Based on these results, it was decided use the Unified Noah LSM in subsequent sensitivity runs instead of the Noah-MP LSM; therefore, Sens0 was not tested further.

Test Results: Base Case, Sensitivity Test 1, and Sensitivity Test 2

The February model performance statistics for Base Case, Sens1, and Sens2 tests are shown in **Figures A-4** and **A-5**, for the full 4-km domain and Uinta Basin Study Area, respectively. The daily biases for Base Case, Sens1, and Sens2 are shown in black, red, and blue, respectively. **Figures A-6** and **A-7** are similar, but for July instead of February.

In February (**Figure A-4**), the daily wind speed bias suggests that Base Case performed the worst while Sens2 performed the best, albeit only slightly better than the Sens1. For the Uinta specific (**Figure A-5**), the daily wind speed bias suggests that Base Case performed the best while Sens1 and Sens2 performed relatively similarly. The daily wind direction tests are inconclusive with respect to the best performing configuration as was for Uinta specific. The daily temperature bias for Sens2 was better or close to the performance of Sens1; however, both simulations performed much better than the Base Case configuration. For Uinta (**Figure A-5**), all configurations show a strong warm bias. With respect to relative humidity, Sens2 was better or close to the performance of Sens1; however, both simulations performed much better than the Base Case configuration. For Uinta, all configurations show a strong wet bias with Base Case being the best of the three.

In July (**Figure A-6**), the daily wind speed biases are inconclusive: the Base Case has a negative bias and Sens1 and Sens2 have a positive bias. The results for wind speed are similar within the Uinta Basin. The daily wind direction tests are inconclusive with respect to the best performing configuration; whereas, within the Uinta Basin, Sens2 performed the best. The daily temperature bias shows a strong cold bias for Base Case configurations and strong warm bias for Sens1 configurations. Within the Uinta Basin (**Figure A-7**), results are inconclusive with respect to the best performing configuration for temperature. Sens2 performed better than the other two simulations for humidity in both geographic areas.

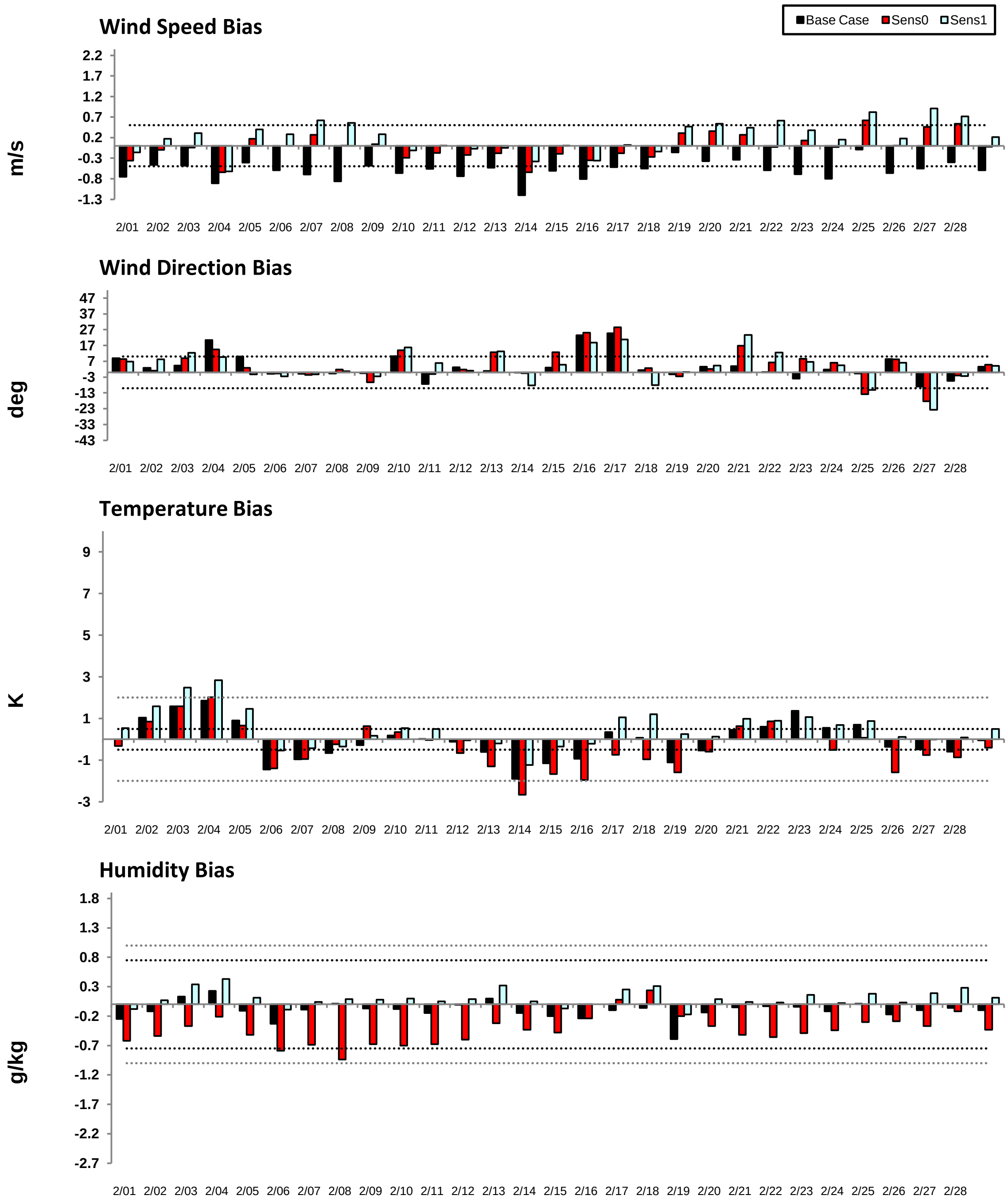


Figure A-2 4-km Domain Bias Bar Charts for Base Case, Sens0, and Sens1 during February, 2010

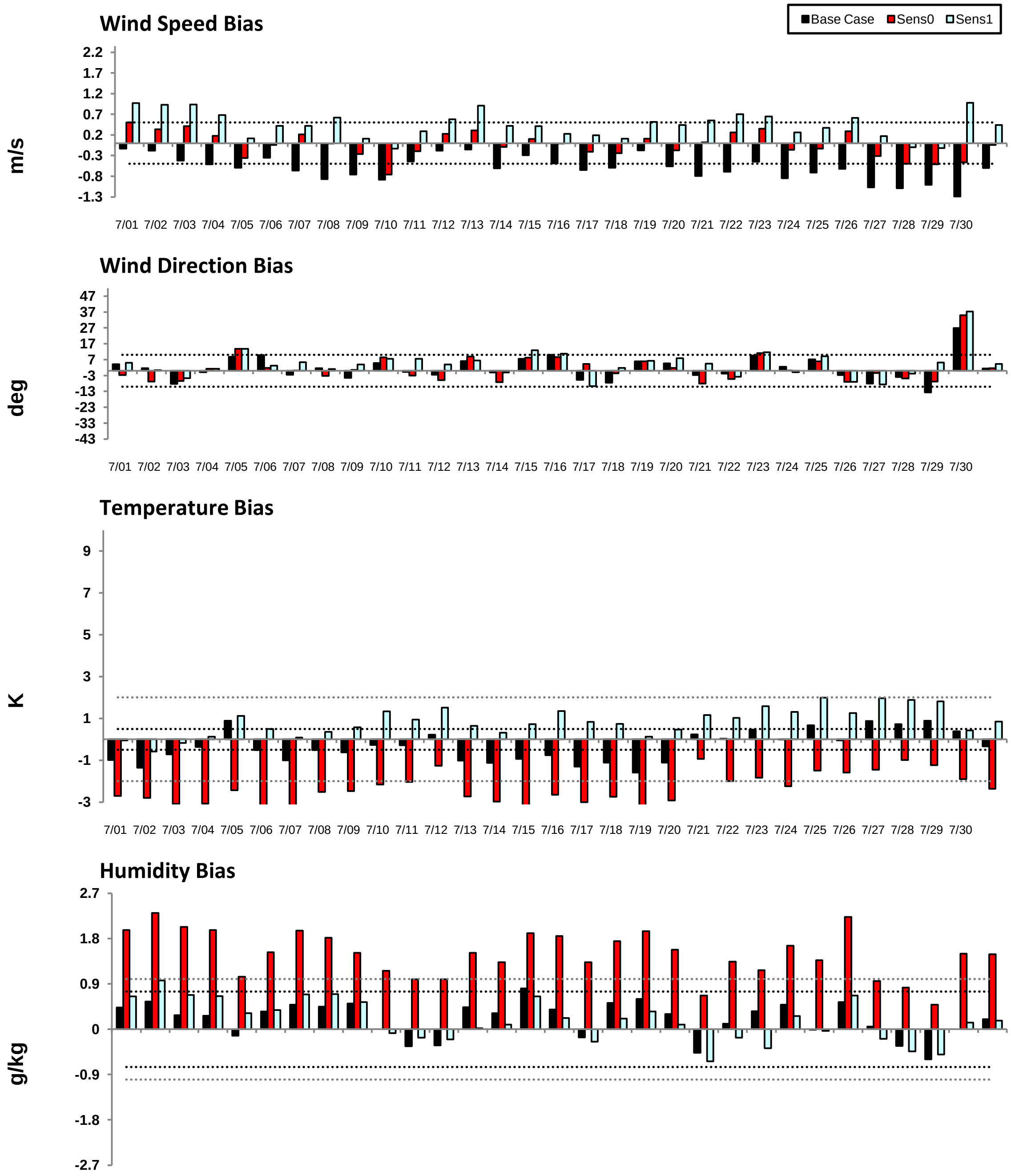


Figure A-3 4-km Domain Bias Bar Charts for Base Case, Sens0, and Sens1 during July, 2010

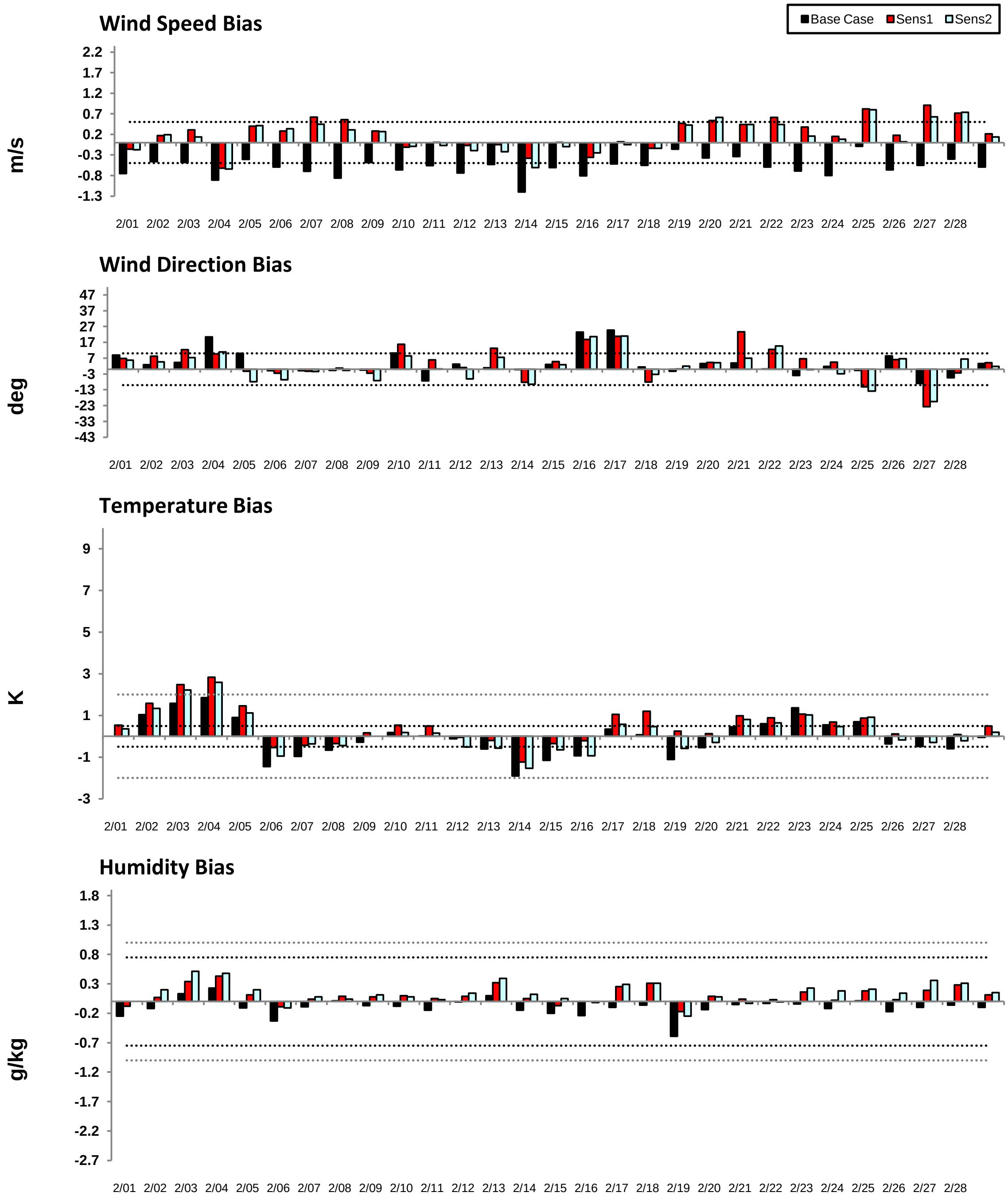


Figure A-4 4-km Domain Bias Bar Charts for Base Case, Sens1, and Sens2 during February, 2010

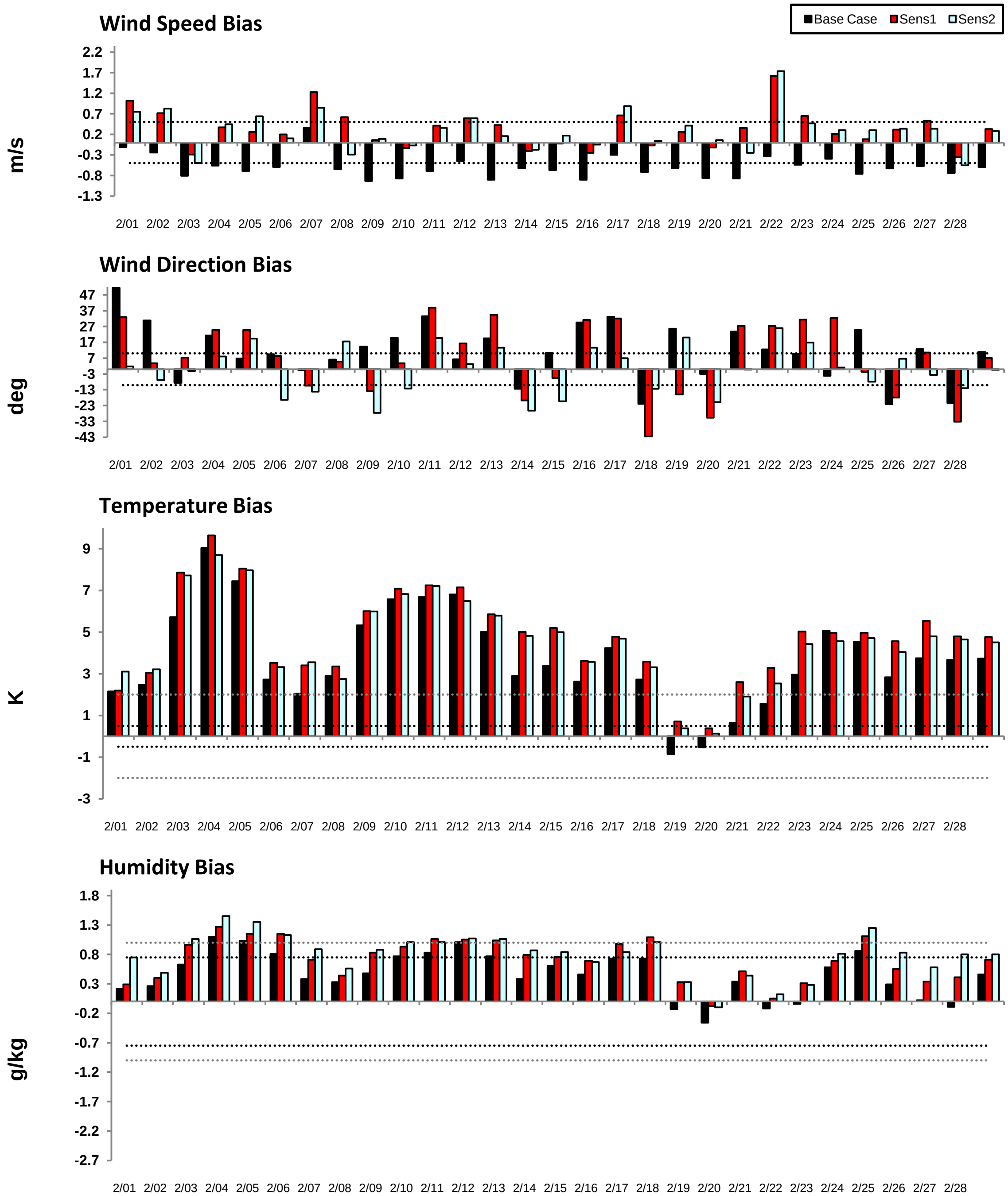


Figure A-5 Bias Bar Charts for Base Case, Sens1, and Sens2 during February, 2010 within Uinta Basin

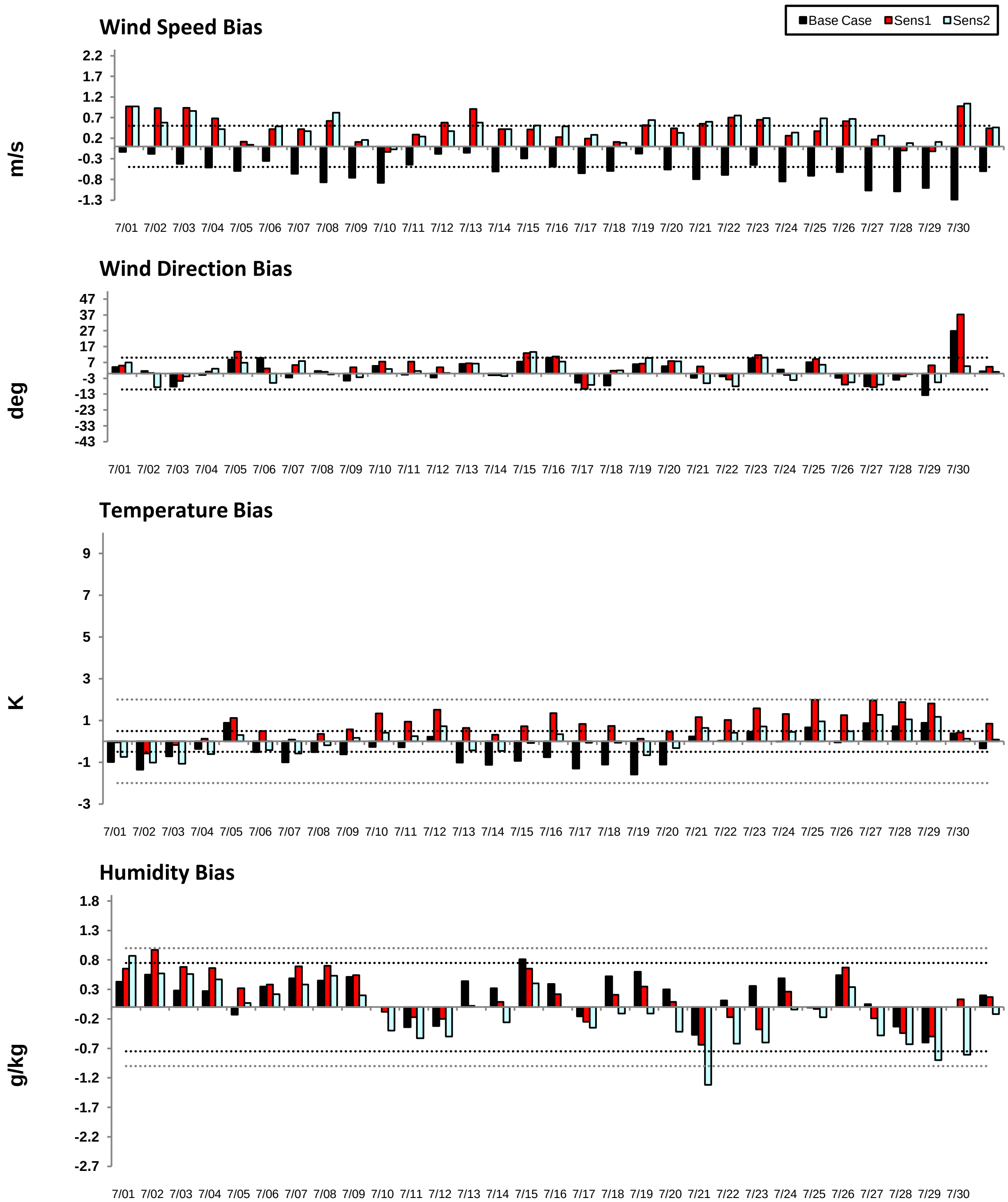


Figure A-6 4-km Domain Bias Bar Charts for Base Case, Sens1, and Sens2 during July, 2010

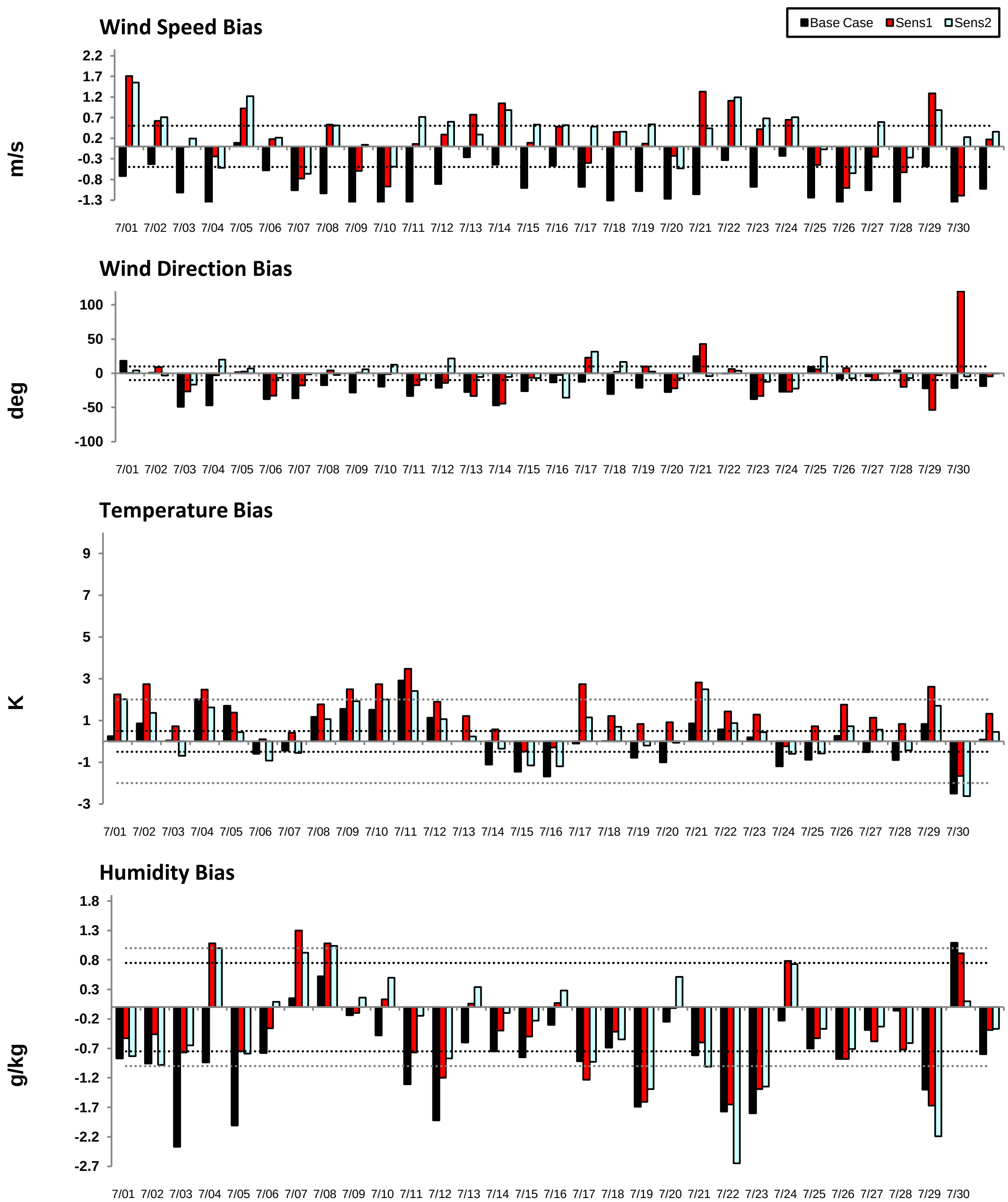


Figure A-7 Bias Bar Charts for Base Case, Sens0, and Sens1 during July, 2010 Within Uinta Basin

In addition to the bar charts presenting daily average statistical performance, monthly average statistics were calculated to provide insight into winter versus summer variations for the three simulations (Base Case, Sens1, and Sens2). The monthly average statistical results for root mean square error (RMSE), bias, gross error, and index of agreement (IOA) for wind speed, wind direction, temperature, and humidity are shown in **Table A-4** for February and **Table A-5** for July. The best performing configuration among the three runs for each statistical variable is shown in red.

Table A-4 Model Performance for Base Case, Sens1, and Sens2 in February

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values			Uinta Basin Average Values		
		(Tesche 2002)	Complex Terrain	Base	Sens1	Sens2	Base	Sens1	Sens2
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	***1.83***	*2.22*	*2.2*	***1.09***	***1.44***	***1.42***
	Bias	± 0.5		-0.59	**0.21**	**0.14**	-0.59	**0.33**	**0.28**
	IOA	≥ 0.6		0.56	0.55	0.54	0.42	0.37	0.38
Wind Direction (deg)	Bias	± 10		**3.58**	**4.09**	**1.78**	11	**7.19**	**-0.38**
	Gross Error	≤ 30	≤ 55	58.66	5.93	60.63	93.74	85.49	76.86
Temperature (K)	Bias	± 0.5	± 2	***-0.05***	***0.5***	***0.19***	3.73	4.77	4.51
	Gross Error	≤ 2	≤ 3.5	*2.71*	*2.87*	*2.88*	4.35	4.95	4.71
	IOA	≥ 0.8		**0.89**	**0.87**	**0.86**	0.58	0.54	0.56
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.1***	***0.11***	***0.15***	***0.46***	***0.71***	***0.8***
	Gross Error	≤ 2	≤ 2	***0.52***	***0.5***	***0.51***	***0.55***	***0.73***	***0.82***
	IOA	≥ 0.6		**0.8**	**0.82**	**0.8**	0.58	0.49	0.44

Values in **red** indicate best performance among the three runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Table A-5 Model Performance for Base Case, Sens1, and Sens2 in July

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values			Uinta Basin Average Values		
		(Tesche 2002)	Complex Terrain	Base	Sens1	Sens2	Base	Sens1	Sens2
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	*2.16*	*2.48*	2.54	*2.25*	*2.36*	*2.44*
	Bias	± 0.5		-0.6	**0.44**	**0.46**	-1.03	**0.17**	**0.36**
	IOA	≥ 0.6		0.59	**0.61**	**0.6**	0.57	0.59	**0.6**
Wind Direction (deg)	Bias	± 10		**1.48**	**4.29**	**1.04**	-18.63	**-4.45**	**-0.46**
	Gross Error	≤ 30	≤ 55	55.6	55.19	55.54	79.35	77.19	72.88
Temperature (K)	Bias	± 0.5	± 2	***-0.34***	*0.85*	***0.09***	***0.09***	*1.33*	***0.45***
	Gross Error	≤ 2	≤ 3.5	*2.4*	*2.6*	*2.49*	*2.12*	*2.4*	*2.11*
	IOA	≥ 0.8		**0.94**	**0.93**	**0.93**	**0.92**	**0.9**	**0.91**
Humidity (g/kg)	Bias	± 0.75	± 1	***0.2***	***0.17***	***-0.12***	*-0.8*	***-0.39***	***-0.37***
	Gross Error	≤ 2	≤ 2	***1.29***	***1.39***	***1.19***	***1.24***	***1.17***	***1.08***
	IOA	≥ 0.6		**0.75**	**0.74**	**0.79**	-32.77	-32.81	-32.75

Values in red indicate best performance among the three runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

These tables list statistical benchmarks, which are used only as a reference and should not be used as an acceptance and/or rejection criteria of the model simulations. Rather, the benchmarks should be used to put the model performances into perspective and allow for an initial identification of potential problems with the configurations.

Following the review of the daily and monthly average model performance statistics, there is not a single model configuration between the Base Case, Sens1, and Sens2 that consistently has better performance for all meteorological parameters. For February, the Base Case simulation tends to have better performance for most meteorological parameters both domain-wide and within the Uinta Basin and as result this configuration will be used for modeling winter months for all subsequent sensitivity tests. The magnitude of the gross error for wind direction and temperature errors in the Uinta Basin are a concern that are addressed with subsequent sensitivity tests.

For July, model performance summary statistics for both Sens1 and Sens2 are very similar and are slightly better than the Base Case. For Uinta Basin, Sens1 performed slightly better than Sens2 for wind speed, but slightly worse for wind direction and temperature. The overall differences between the summary statistics for Sens1 and Sens2 are relatively small. To determine which set of physics to select for the non-winter months in the annual simulation, we re-examined the physics used in both sensitivity runs. Since the QNSE-EDMF PBL scheme used in Sens2 is developed specifically for stable boundary layers, which are more prevalent in winter, and the results were very similar to Sens1, it was decided to use the MYJ PBL scheme (Sens1) for non-winter months for subsequent sensitivity tests.

Sensitivity Test 3 Test Results: Expanded Domains

The February model performance statistics for Base Case and Sens3 are shown in **Figures A-8** and **A-9**, for the full 4-km domain and Uinta Basin Study Area, respectively. The daily biases for Base Case and Sens3 are shown in black and red, respectively. **Figures A-10** and **A-11** are similar, but for July instead of February and show Sens1 in black and Sens3 in red. **Tables A-6** and **A-7** compare the monthly average statistics for the previous test to Sens3 results for February and July, respectively.

Importantly, there is not a notable improvement in the model performance for Sens3. As a result, it was concluded that the domains used in the Base Case were appropriate for the annual simulation.

Sensitivity Test 4 Test Results: Enhanced Vertical Layers

The vertical layer structure used for the Base Case, Sens1, Sens2, and Sens3 is shown in **Table A-8** and the enhanced vertical resolution tested in Sens4 is shown in **Table A-9.** Two changes to the Base Case vertical layer structure were tested in Sens4. The first was an enhanced vertical resolution at the surface level. The bottom two vertical layers in the Base Case model configuration (layer numbers 1 and 2 shown in **Table A-8**) were replaced by four vertical layers of 20 m depth (layer numbers 1 through 4 shown in **Table A-9**). The second change was to the upper layers of the WRF model. WRF developers do not recommended using model vertical layers thicker than 1,000 m.[4] Therefore, the top six vertical layers in the Base Case model configuration (layer numbers 29 through 34 shown in **Table A-8**) were replaced by twelve vertical layers of 1,000 m depth (layer numbers 31 through 42 shown in **Table A-9**). Altogether, these changes to the vertical layer structure resulted in 42 vertical layers in Sens4.

---

[4] January 2012 tutorial presentation "Considerations for Designing an Numerical Experiment" (http://www.mmm.ucar.edu/wrf/users/tutorial/201201/WRF_expt-design.ppt.pdf)

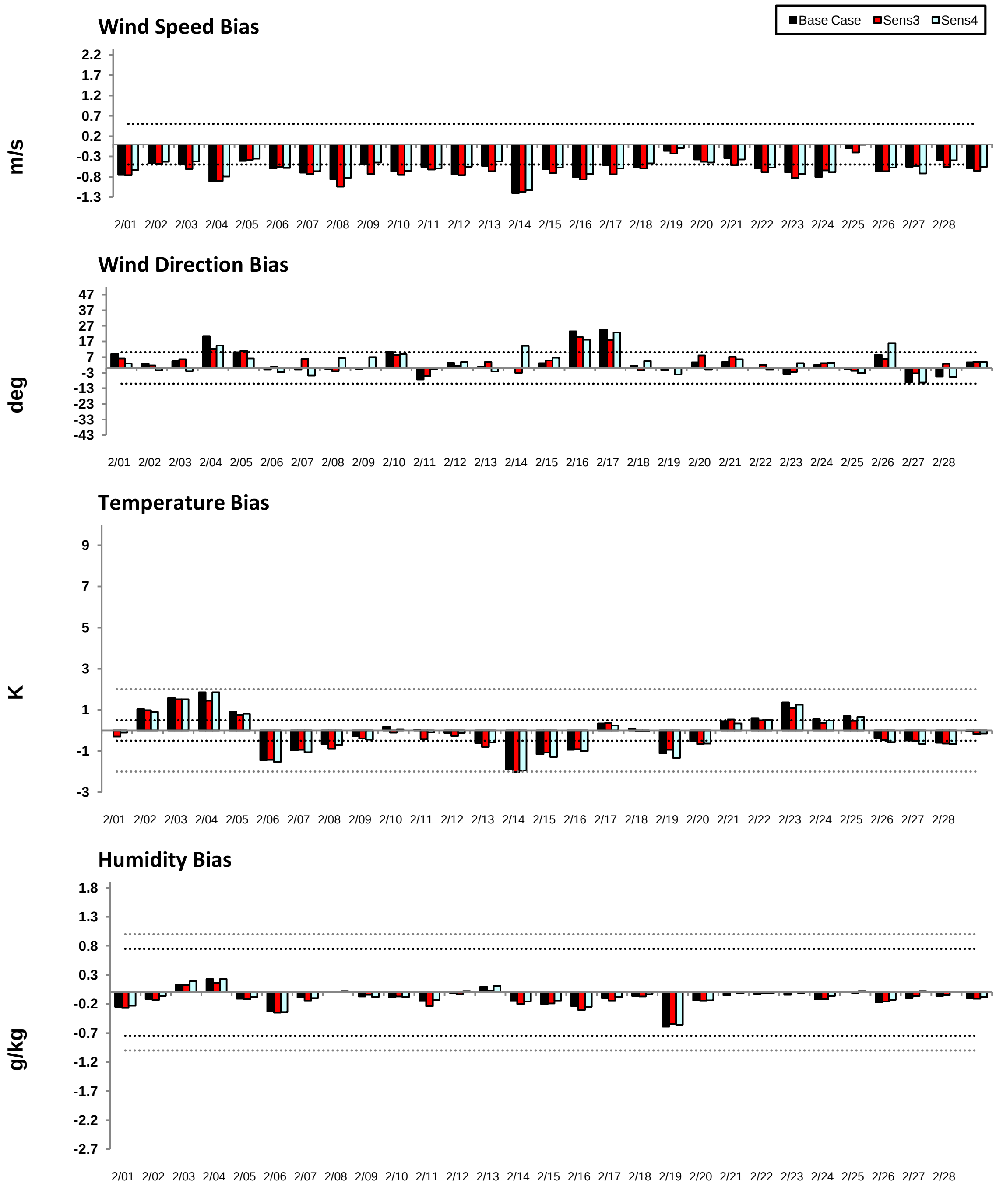


Figure A-8 4-km Domain Bias Bar Charts for Base Case, Sens3, and Sens4 during February 2010

Figure A-9 Bias Bar Charts for Base Case, Sens3, and Sens4 during February 2010 within the Uinta Basin

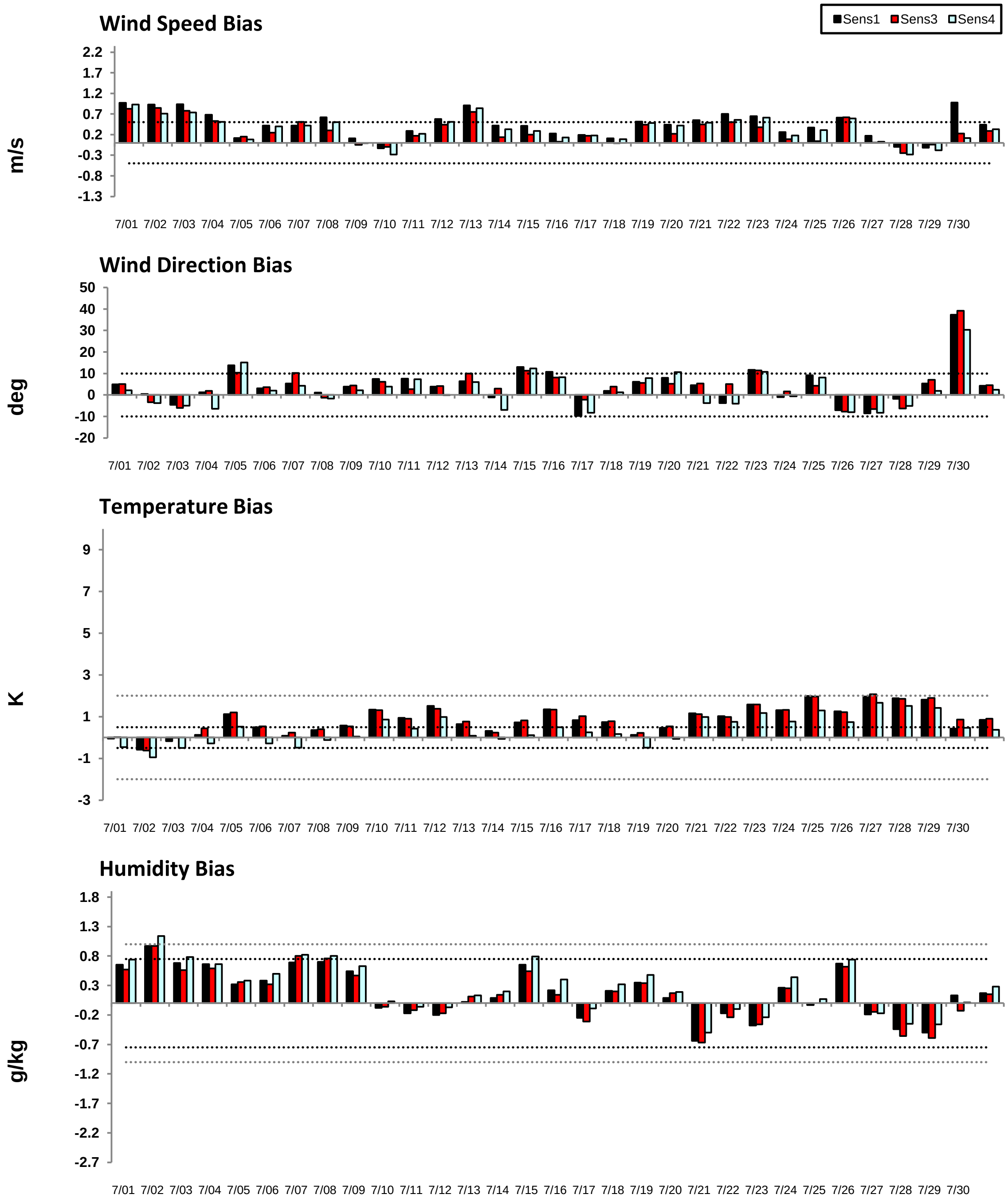


Figure A-10 4-km Domain Bias Bar Charts for Sens1, Sens3, and Sens4 during July 2010

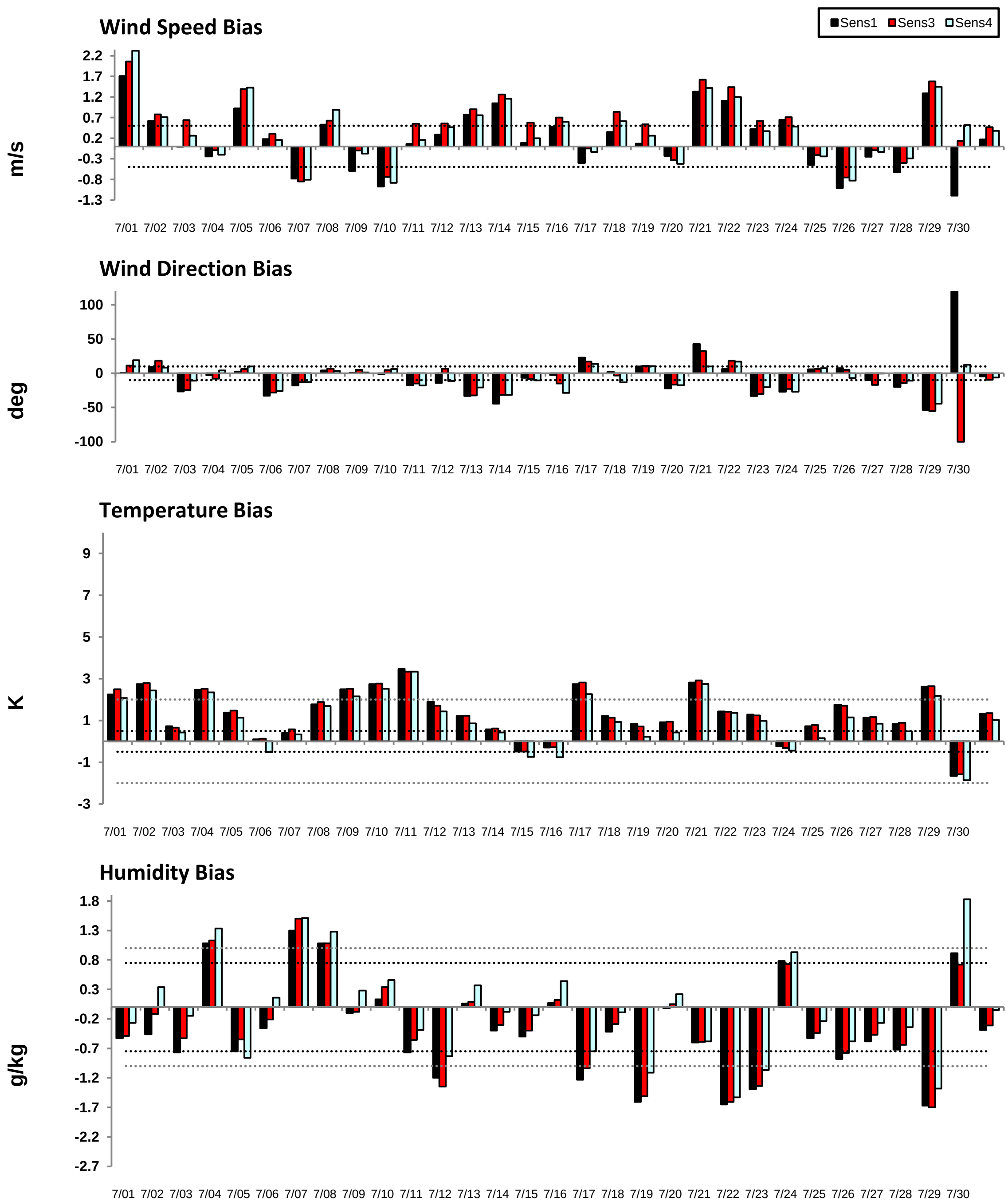


Figure A-11 Bias Bar Charts for Sens1, Sens3, and Sens4 during July 2010 within Uinta Basin

Table A-6 Model Performance for Base Case and Sens3 in February

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Basin Average Values	
		(Tesche 2002)	Complex Terrain	Base	Sens3	Base	Sens3
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	***1.83***	***1.82***	***1.09***	***1.24***
	Bias	± 0.5		-0.59	-0.65	-0.59	**-0.15**
	IOA	≥ 0.6		0.56	0.58	0.42	0.39
Wind Direction (deg)	Bias	± 10		**3.58**	**3.84**	11	10.18
	Gross Error	≤ 30	≤ 55	58.66	56.73	93.74	82.22
Temperature (K)	Bias	± 0.5	± 2	***-0.05***	***-0.17***	3.73	4.26
	Gross Error	≤ 2	≤ 3.5	*2.71*	*2.59*	4.35	4.75
	IOA	≥ 0.8		**0.89**	**0.9**	0.58	0.59
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.1***	***-0.11***	***0.46***	***0.53***
	Gross Error	≤ 2	≤ 2	***0.52***	***0.5***	***0.55***	***0.62***
	IOA	≥ 0.6		**0.8**	**0.81**	0.58	0.5

Values in red indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Table A-7 Model Performance for Sens1 and Sens3 in July

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Bain Average Values	
		(Tesche 2002)	Complex Terrain	Sens1	Sens3	Sens1	Sens3
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	*2.48*	*2.4*	*2.36*	*2.44*
	Bias	± 0.5		**0.44**	**0.29**	**0.17**	**0.47**
	IOA	≥ 0.6		**0.61**	**0.63**	0.59	0.59
Wind Direction (deg)	Bias	± 10		**4.29**	**4.53**	**-4.45**	**-9.62**
	Gross Error	≤ 30	≤ 55	55.19	*53.93*	77.19	75.37
Temperature (K)	Bias	± 0.5	± 2	*0.85*	*0.9*	*1.33*	*1.35*
	Gross Error	≤ 2	≤ 3.5	*2.6*	*2.5*	*2.4*	*2.42*
	IOA	≥ 0.8		**0.93**	**0.93**	**0.9**	**0.89**
Humidity (g/kg)	Bias	± 0.75	± 1	***0.17***	***0.15***	***-0.39***	***-0.31***
	Gross Error	≤ 2	≤ 2	***1.39***	***1.34***	***1.17***	***1.18***
	IOA	≥ 0.6		**0.74**	**0.75**	-32.81	-32.82

Values in **red** indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Table A-8 Vertical Layer Structure For Base Case, Sens0, Sen1, Sens2, and Sens3

Model Layer	Sigma	Pressure (mb)	Height (meters)	Depth (meters)
34 – top	0.000	50	20,559	4,262
33	0.050	98	16,297	2,527
32	0.100	145	13,770	1,805
31	0.150	193	11,965	1,407
30	0.200	240	10,559	1,185
39	0.250	288	9,374	1,035
28	0.300	335	8,339	931
27	0.350	383	7,408	832
26	0.400	430	6,576	760
25	0.450	478	5,816	701
24	0.500	525	5,115	652
23	0.550	573	4,463	609
22	0.600	620	3,854	572
21	0.650	668	3,282	540
20	0.700	715	2,741	412
19	0.740	753	2,329	298
18	0.770	782	2,032	290
17	0.800	810	1,742	188
16	0.820	829	1,554	185
15	0.840	848	1,369	182
14	0.860	867	1,188	178
13	0.880	886	1,009	175
12	0.900	905	834	87
11	0.910	915	747	85
10	0.920	924	662	85
9	0.930	934	577	85
8	0.940	943	492	83
7	0.950	953	409	83
6	0.960	962	326	83
5	0.970	972	243	81
4	0.980	981	162	41
3	0.985	986	121	41
2	0.990	991	80	40
1	0.995	995	40	40
0 – ground	1.000	1,000	0	0

Table A-9 Vertical Layer Structure for Sens4

Model Layer	Sigma	Pressure (mb)	Height (meters)	Depth (meters)
42 – Top	0.000	50	20,559	1,220
41	0.011	61	19,339	1,000
40	0.022	71	18,339	1,000
39	0.035	83	17,339	1,000
38	0.050	97	16,339	1,000
37	0.067	114	15,339	1,000
36	0.087	133	14,339	1,000
35	0.111	155	13,339	1,000
34	0.139	182	12,339	1,000
33	0.171	213	11,339	1,000
32	0.209	249	10,339	1,000
31	0.252	289	9,339	1,000
30	0.300	335	8,339	931
29	0.350	383	7,408	832
28	0.400	430	6,576	760
27	0.450	478	5,816	701
26	0.500	525	5,115	652
25	0.550	573	4,463	609
24	0.600	620	3,854	572
23	0.650	668	3,282	540
22	0.700	715	2,741	412
21	0.740	753	2,329	298
20	0.770	782	2,032	290
19	0.800	810	1,742	188
18	0.820	829	1,554	185
17	0.840	848	1,369	182
16	0.860	867	1,188	178
15	0.880	886	1,009	175
14	0.900	905	834	87
13	0.910	915	747	85
12	0.920	924	662	85
11	0.930	934	577	85
10	0.940	943	492	83
9	0.950	953	409	83
8	0.960	962	326	83
7	0.970	972	243	81
6	0.980	981	162	41

Table A-9 Vertical Layer Structure for Sens4

Model Layer	Sigma	Pressure (mb)	Height (meters)	Depth (meters)
5	0.985	986	121	41
4	0.990	991	80	20
3	0.9929	993	60	20
2	0.995	995	40	20
1	0.9976	998	20	20
0 –ground	1.000	1,000	0	0

As described in the prior section, the WRF configuration chosen for February and July are Base Case and Sens1, respectively. Therefore, for Sens4, 42 vertical layers were tested with the Base Case configuration in February and with Sens1 configuration in July. The model domains from Base Case are used in Sens4.

The February model performance statistics for Base Case and Sens4 are shown in **Figures A-8** and **A-9**, for the full 4-km domain and Uinta Basin Study Area, respectively. The daily biases for Base Case and Sens4 are shown in black and blue, respectively. **Figures A-10** and **A-11** are similar, but for July instead of February and show Sens1 in black and Sens4 in blue. **Tables A-10** and **A-11** compare the monthly average statistics for the previous test to Sens4 results for February and July, respectively.

Based on these results, the model performance for Sens4 in February was slightly better for wind speed and humidity, and was slightly better for wind direction, temperature, and humidity in July.

To analyze WRF performance for conditions aloft, the upper-air data collected by NWS twice per day at 0000 and 1200 GMT (or 5:00 p.m. and 5:00 a.m. MST, respectively) from the Salt Lake City, Utah (SLC) station were compared with the WRF model simulated vertical structure of the atmosphere. The observed soundings for the SLC station were plotted with the vertical profile of the Base Case and Sens4 WRF output at the grid cell that contains the SLC station. **Figures A-12** and **A-13** show the Base and Sens4 comparison profiles at 1200 GMT for SLC on February 3, 2010 and July 4, 2010, respectively. The observed sounding is shown in black and the WRF simulation in red. Due to the limit in vertical resolution in the WRF simulation, the WRF vertical profiles are smoother than those from observations.

In general, both the temperature (solid lines) and dew point temperature (dashed lines) profiles from the Base Case and Sens4 simulations followed the profiles from the observed soundings. The temperature profile performed better than the dew point temperature profile. Even with the expanded vertical layers, the model had difficulty replicating the sharp changes in the dew point temperature that is related to the water vapor mixing ratio. The model generally was able to simulate the vertical variability in wind direction and speed. Importantly, on February 3, Sens4 reproduced the height of the inversions as well as the strengths.

Based upon these analyses, the enhanced vertical resolution in the surface layers improves model performance at the surface. The skew-T plots also suggest that little is gained by using the enhanced vertical levels in the upper troposphere and stratosphere. Therefore, the final vertical model layers (shown in **Table 2-1**) used the refined surface layers from Sen4, but maintained the rest of the vertical layers from the Base Case.

Sensitivity Test 5 Test Results: Revised Observation Nudging Dataset

A revised observation nudging dataset was tested for February performance. The observation nudging dataset was processed to cover the full WRF 4-km domain for year 2010 and include monitored meteorological data downloaded from MADIS and processed to include observations at Ouray and Redwash, which are located in the Uinta Basin.

Table A-10 Model Performance for Base Case and Sens4 in February

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Basin Average Values	
		(Tesche 2002)	Complex Terrain	Base	Sens4	Base	Sens4
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	***1.83***	***1.81***	***1.09***	***1.17***
	Bias	± 0.5		-0.59	-0.55	-0.59	**-0.05**
	IOA	≥ 0.6		0.56	0.56	0.42	0.41
Wind Direction (deg)	Bias	± 10		**3.58**	**3.76**	11	**2.75**
	Gross Error	≤ 30	≤ 55	58.66	56.94	93.74	78.45
Temperature (K)	Bias	± 0.5	± 2	***-0.05***	***-0.15***	3.73	4.02
	Gross Error	≤ 2	≤ 3.5	*2.71*	*2.69*	4.35	4.62
	IOA	≥ 0.8		**0.89**	**0.89**	0.58	0.61
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.1***	***-0.08***	***0.46***	***0.53***
	Gross Error	≤ 2	≤ 2	***0.52***	***0.52***	***0.55***	***0.62***
	IOA	≥ 0.6		**0.8**	**0.8**	0.58	0.51

Values in **red** indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Table A-11 Model Performance for Sens1 and Sens4 in July

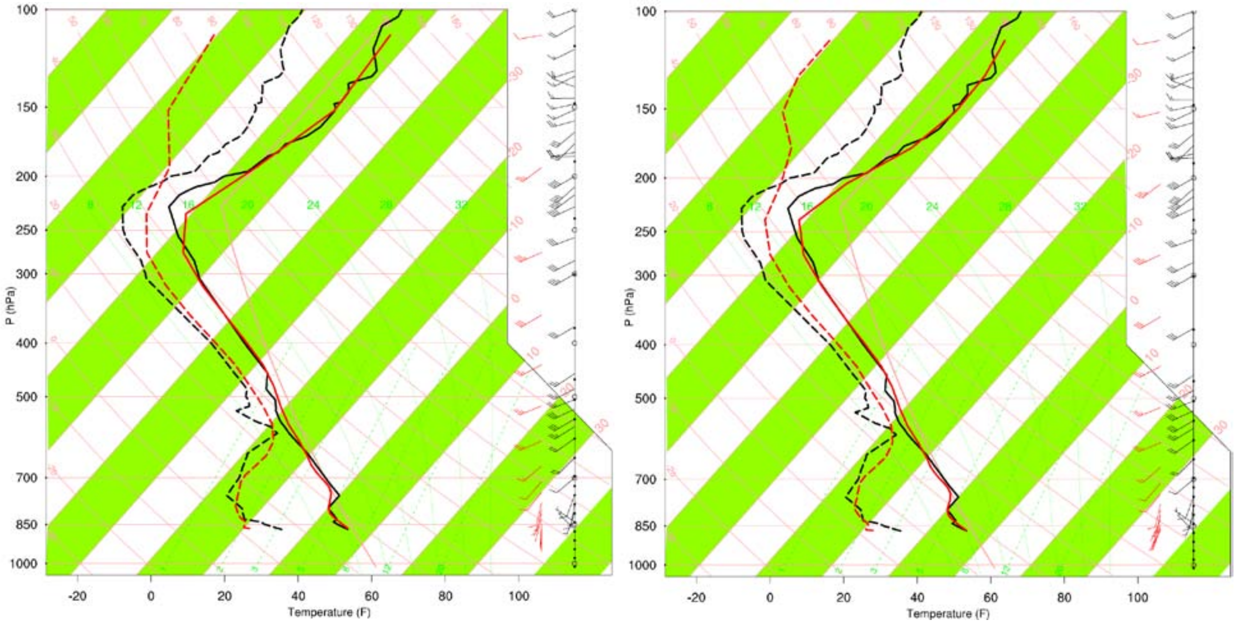
Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Basin Average Values	
		(Tesche 2002)	Complex Terrain	Sens1	Sens4	Sens1	Sens4
Wind Speed (m/s)	RSME	≤ 2	≤ 2.5	*2.48*	*2.44*	*2.36*	*2.45*
	Bias	± 0.5		**0.44**	**0.33**	**0.17**	**0.38**
	IOA	≥ 0.6		**0.61**	**0.62**	0.59	0.59
Wind Direction (deg)	Bias	± 10		**4.29**	**2.4**	**-4.45**	**-6.29**
	Gross Error	≤ 30	≤ 55	55.19	55.44	77.19	74.81
Temperature (K)	Bias	± 0.5	± 2	*0.85*	***0.37***	*1.33*	*1.02*
	Gross Error	≤ 2	≤ 3.5	*2.6*	*2.54*	*2.4*	*2.34*
	IOA	≥ 0.8		**0.93**	**0.93**	**0.9**	**0.9**
Humidity (g/kg)	Bias	± 0.75	± 1	***0.17***	***0.28***	***-0.39***	***-0.05***
	Gross Error	≤ 2	≤ 2	***1.39***	***1.37***	***1.17***	***1.12***
	IOA	≥ 0.6		**0.74**	**0.75**	-32.81	-32.79

Values in red indicate best performance among the five runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Base Case Sens4

a) February 3, 2010 at 0000GMT



b) February 3, 2010 at 1200GMT

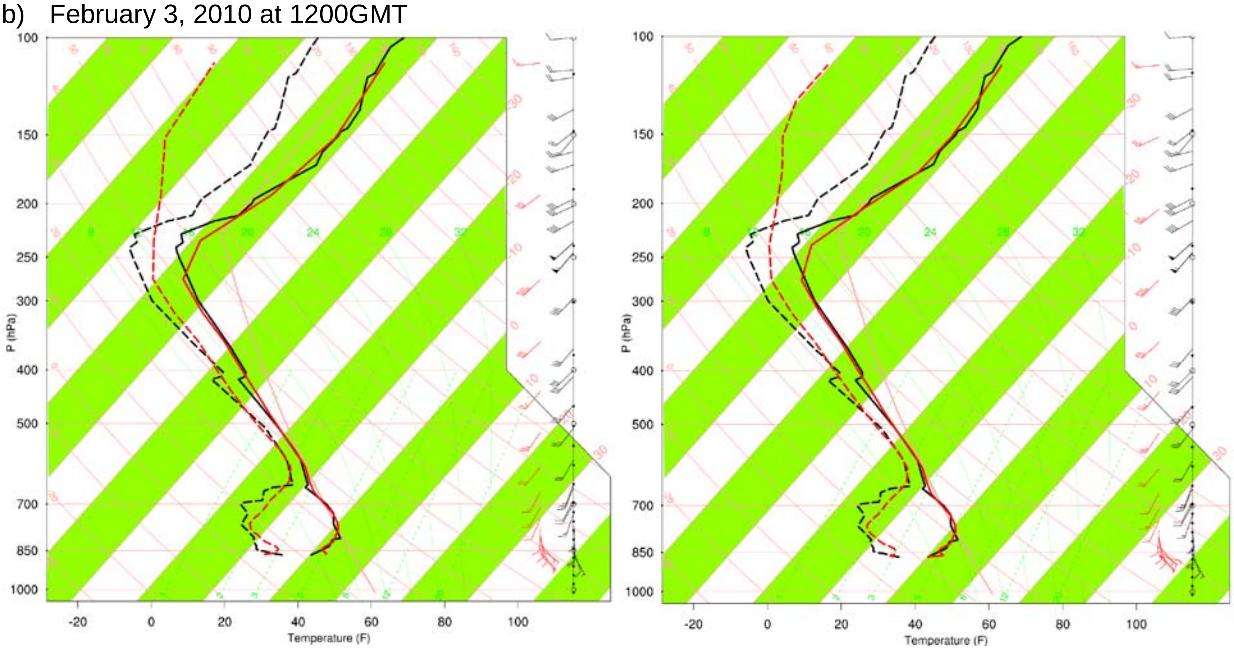
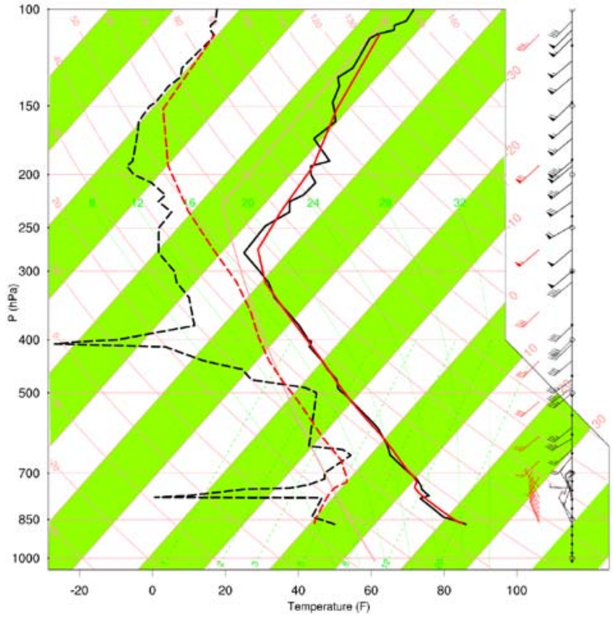


Figure A-12 Comparison of SLC Vertical Profiles between the Base Case and Sens4

Sens1 Sens4

a) July 4, 2010 at 0000GMT



b) July 4, 2010 at 1200GMT

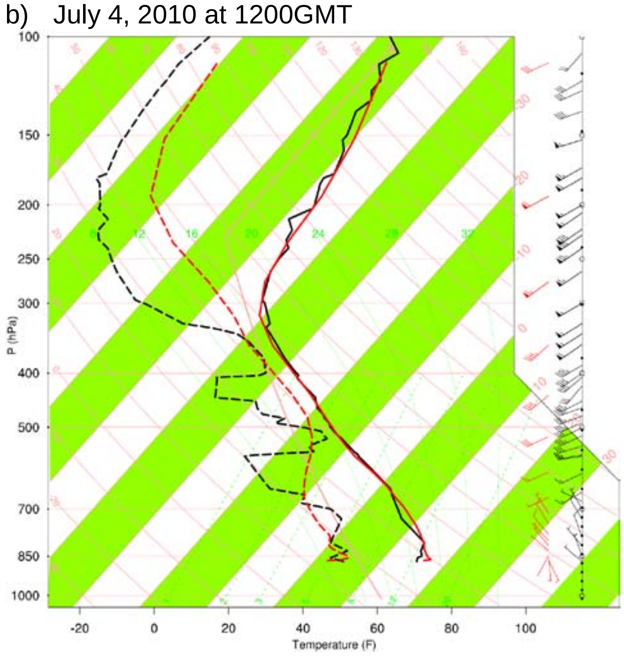


Figure A-13 Comparison of SLC Vertical Profiles between the Sens1 and Sens4

The February time series plots for Sens5 are shown relative to observations in **Figures A-14** and **A-15**, for the full 4-km domain and Uinta Basin Study Area, respectively. Overall, the 4-km domain-wide time series plots show that simulation for temperature and mixing ratio followed the synoptic and diurnal pattern of the observations reasonably well. However, the modeled wind speeds showed a consistent under-prediction of the wind peak speed and the modeled wind directions were unable to replicate the change of the wind in some days. In the Uinta Basin, modeled wind speeds followed the observations for most days, but WRF failed to simulate the sharp increases in wind speed on some days. The modeled wind directions did not replicate the variable nature of the wind caused by the complex terrain in the Uinta Basin. Importantly, large errors in the modeled temperature and mixing ratio are still present in Uinta Basin.

Table A-12 compares the monthly average statistics from the Base Case to Sens5 results for February. Based on these results, domain-wide performance is improved for Sens5. However, within the Uinta Basin, Sens5 showed slight improvement over the Base Case simulation for wind speed and direction, but temperature and mixing ratio performances degraded slightly.

Sensitivity Test 6 Test Results: Buddy Checks Turned Off

Although Sens5 yielded improvements in the 4-km domain-wide model performance, the magnitude of the gross error for temperature and mixing ratio within the Uinta Basin is problematic. The summary statistics and the time-series plots suggest that not all available surface observations were incorporated into observation nudging.

We re-examined the options that were used in OBSGRID, a preprocessor that improves first-guess gridded analyses used in WRF by incorporating additional observational information. All the sensitivity simulations performed up to this point applied all four available quality control (QC) tests in OBSGRID. The four tests are: 1) Error Max Test: checks the difference between the first-guess and the observation and keeps the observation only if differences do not exceed allowable thresholds; 2) Buddy Check Test: checks the difference between a single observation and neighboring observations for; 3) Spike Removal: removes vertical spikes in the data; and 4) Removal of Super-adiabatic Lapse Rates: removes any super-adiabatic lapse rate in a sounding by conservation of dry static energy.

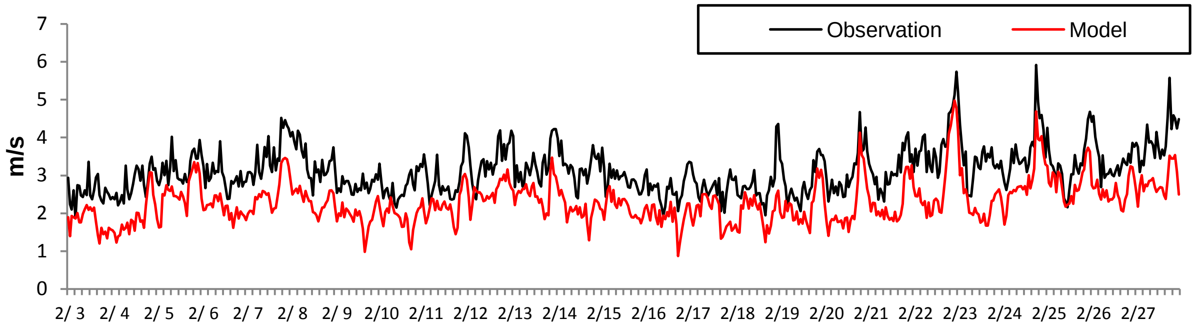
For sensitivity test 6, the Buddy Check Test was disabled in OBSGRID.

The February time series plots for Sens6 are shown relative to observations in **Figures A-16** and **A-17**, for the full 4-km domain and Uinta Basin Study Area, respectively. Overall, these plots show that WRF wind speed, wind direction, and temperature followed the synoptic and diurnal pattern of the observations. However, the modeled mixing ratio showed a consistent negative bias for the entire period. Comparison of **Figures A-16** and **Figure A-14** suggested that Sens6 performed slightly better for wind speed and wind direction. For temperature, Sens6 performed comparable to Sens5 for daily maximum but is slightly worse for daily minimum. The Sens6 mixing ratio performance is worse than Sens5.

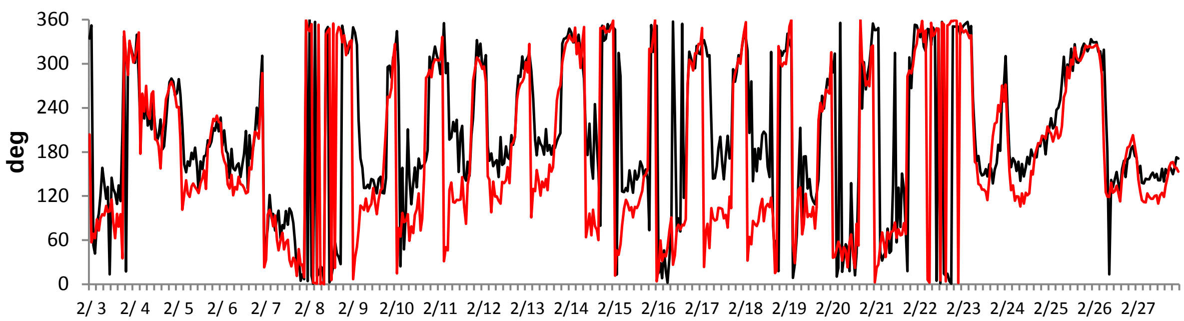
The Sens6 wind speed and direction performance is similar to those in Sens5 (comparison of **Figure A-17** and **Figure A-15**). Importantly, both temperature and mixing ratio performed much better in the Uinta Basin in Sens6 than in Sens5. In Sens6, with the exception of a few days, model simulation of temperature and mixing ratio followed the diurnal pattern of the observations reasonably well.

Table A-13 compares the monthly average statistics from the Sens5 results to Sens6 results for February. Based on these results, the model performance degraded slightly in the 4-km domain but improved within the Uinta Basin.

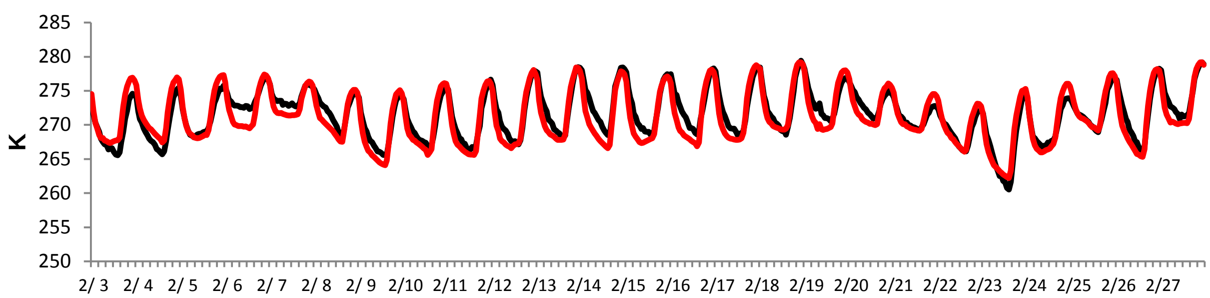
Wind Speed



Wind Direction



Temperature



Humidity

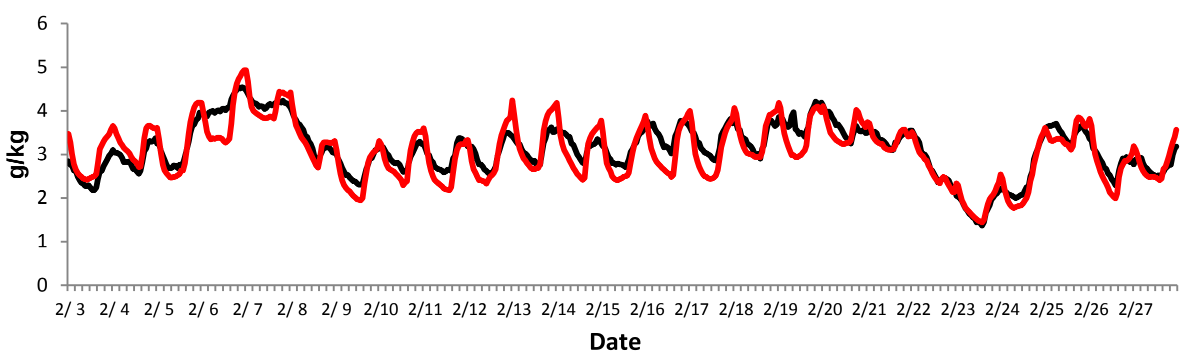
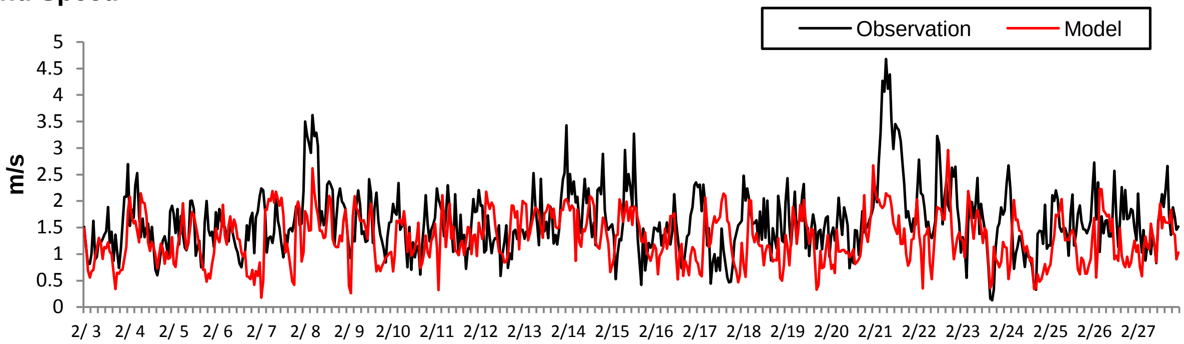
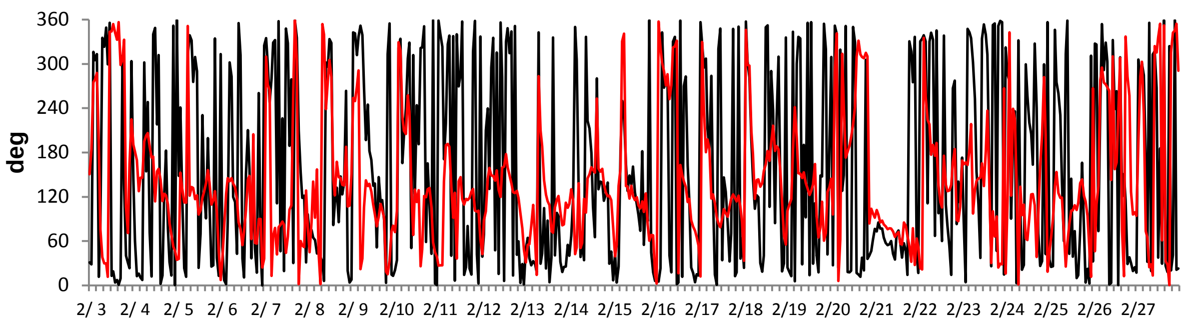


Figure A-14 4-km Domain Time Series for Sens5 during February

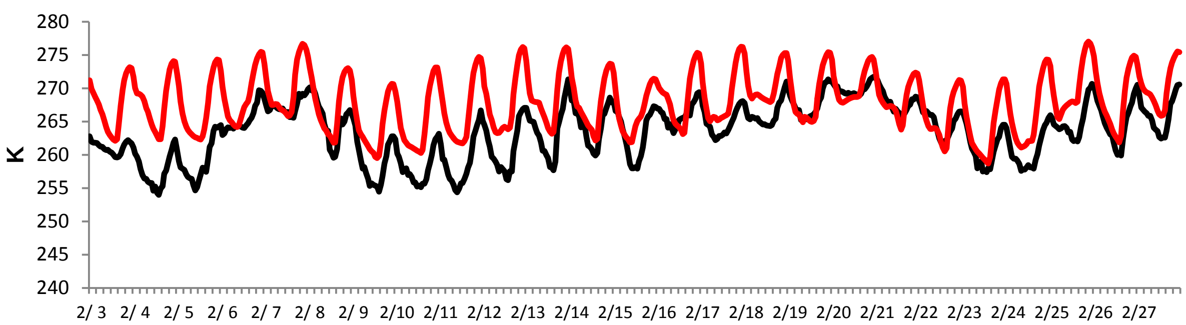
Wind Speed



Wind Direction



Temperature



Humidity

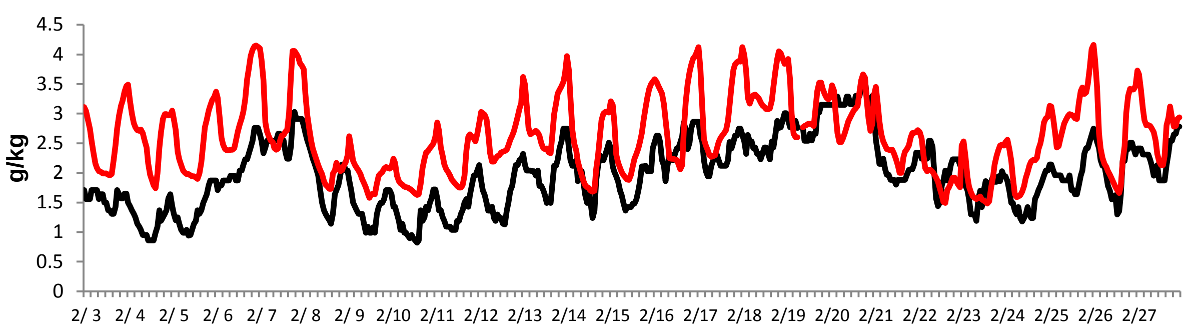


Figure A-15 Uinta Basin Time Series for Sens5 during February

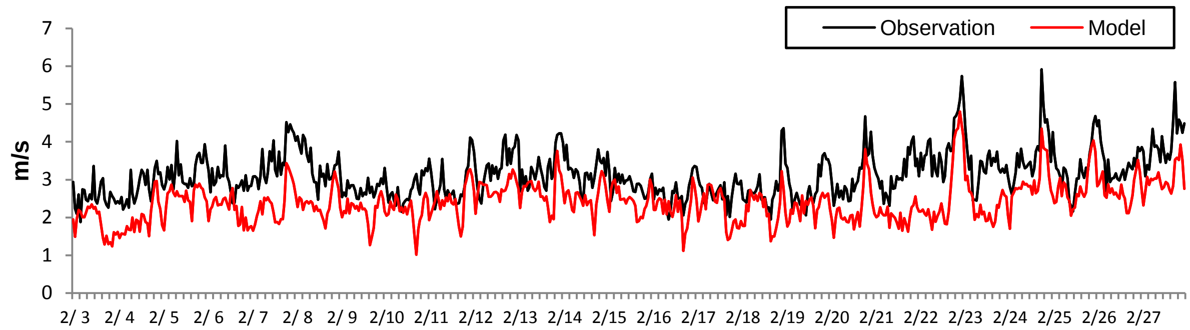
Table A-12 Model Performance for Base Case and Sens5 in February

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Basin Average Values	
		(Tesche 2002)	Complex Terrain	Base	Sens5	Base	Sens5
Wind Speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.84***	***1.58***	***1.11***	***0.96***
	Bias	± 0.5		-0.59	-0.83	-0.63	**-0.32**
	IOA	≥ 0.6		0.57	**0.66**	0.42	0.52
Wind Direction (deg)	Bias	± 10		**3.40**	**2.52**	8.67	**5.61**
	Gross Error	≤ 30	≤ 55	58.51	*40.33*	92.90	59.42
Temperature (K)	Bias	± 0.5	± 2	***-0.10***	***-0.39***	3.84	4□40
	G□oss Error	≤ 2	≤ 3.5	*2.65*	*2.41*	4.45	4.81
	IOA	≥ 0.8		**0.90**	**0.91**	0.59	0.62
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.09***	***-0.06***	***0.48***	***0.64***
	Gross Error	≤ 2	≤ 2	***0.52***	***0.51***	***0.57***	***0.70***
	IOA	≥ 0.6		**0.80**	**0.81**	0.58	0.51

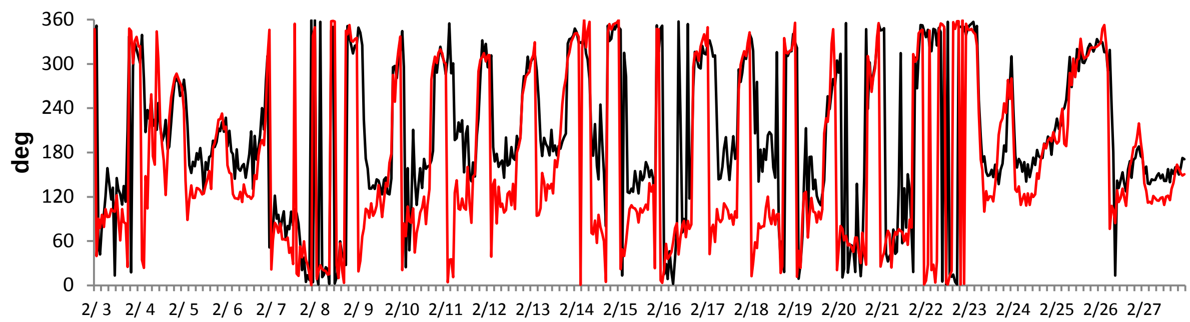
Values in red indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

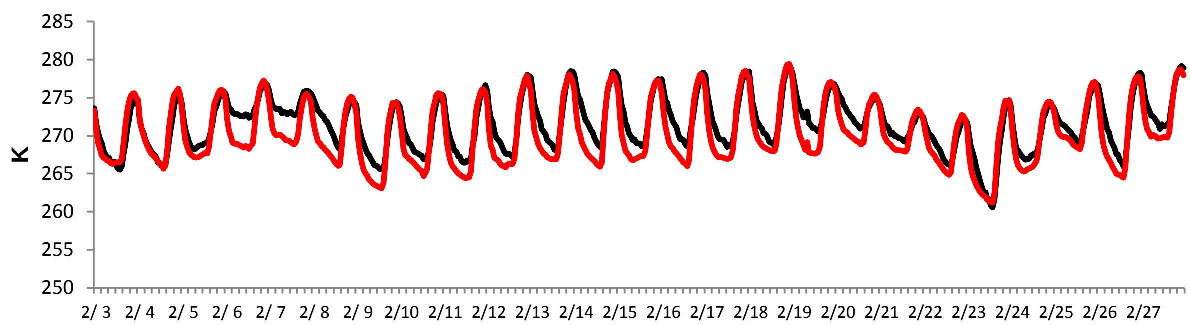
Wind Speed



Wind Direction



Temperature



Humidity

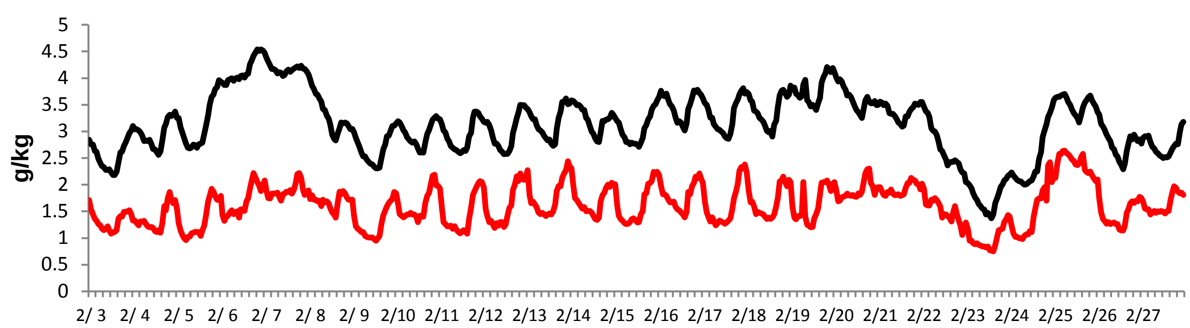
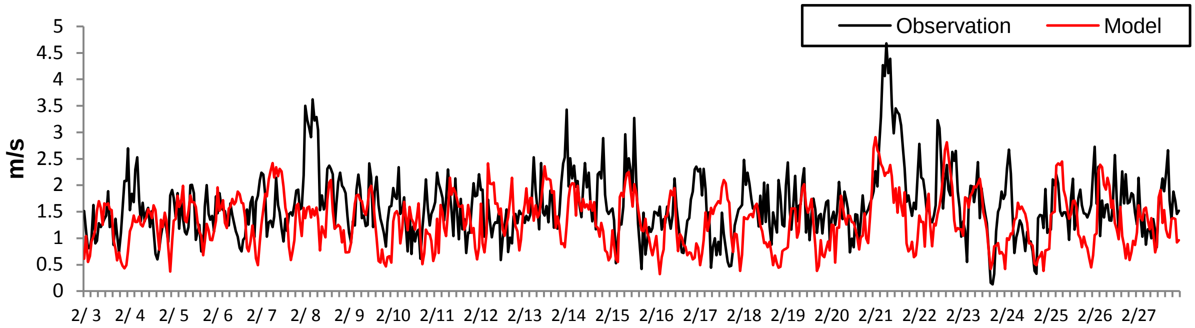
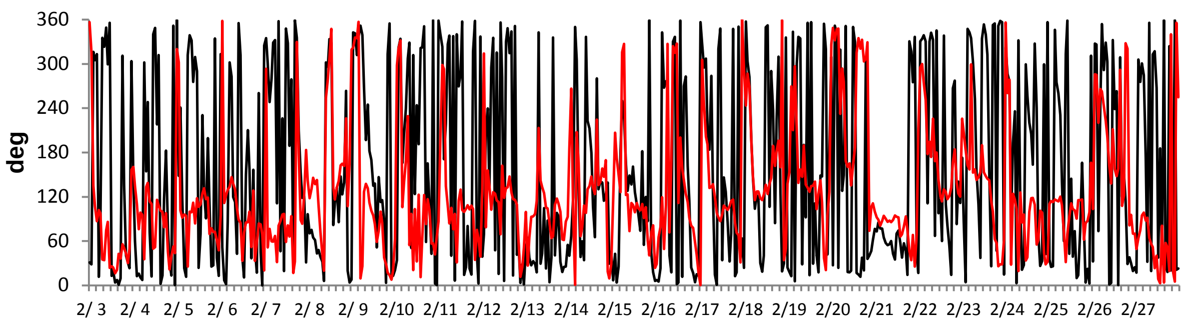


Figure A-16 4-km Domain Time Series for Sens6 during February

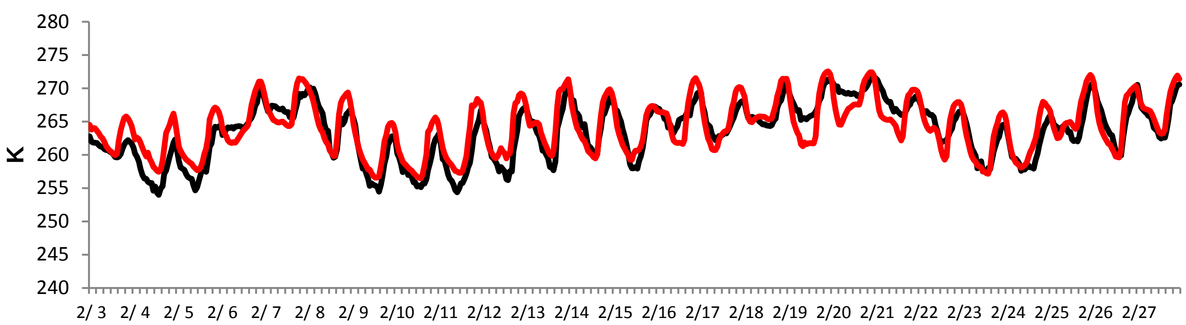
Wind Speed



Wind Direction



Temperature



Humidity

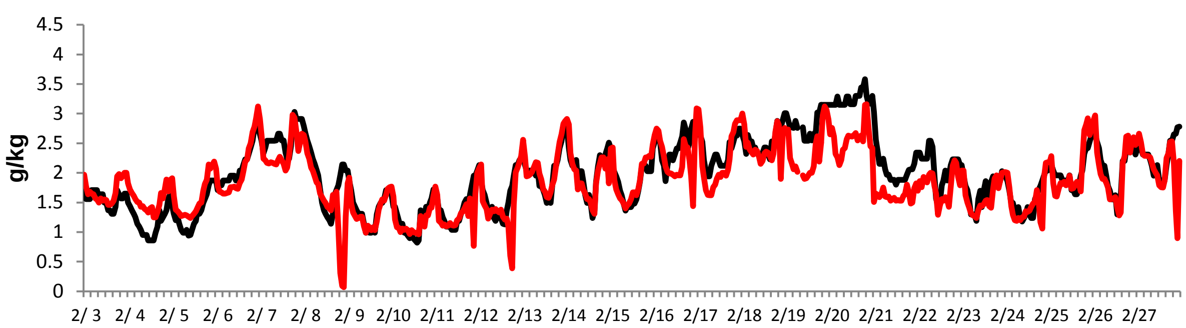


Figure A-17 Uinta Basin Time Series for Sens6 during February

Table A-13 Model Performance for Sens5 and Sens6 in February

Parameter	Statistics	Statistical Benchmark (Tesche 2002)	Complex Terrain	4-km Domain Average Values Sens5	Sens6	Uinta Basin Average Values Sens5	Sens6
Wind Speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.58***	***1.63***	***1.09***	***1.01***
	Bias	± 0.5		-0.83	-0.72	-0.59	**-0.29**
	IOA	≥ 0.6		0.66	0.64	0.42	0.48
Wind Direction (deg)	Bias	± 10		**2.52**	**3.65**	11	**2.64**
	Gross Error	≤ 30	≤ 55	40.33	43.39	93.74	58.47
Temperature (K)	Bias	± 0.5	± 2	***-0.39***	*-1.31*	3.73	*0.72*
	Gross Error	≤ 2	≤ 3.5	2.41	2.36	4.35	*2.08*
	IOA	≥ 0.8		**0.91**	**0.92**	0.58	**0.84**
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.06***	-1.51	***0.46***	***-0.11***
	Gross Error	≤ 2	≤ 2	***0.51***	***1.55***	***0.55***	***0.24***
	IOA	≥ 0.6		**0.81**	0.42	0.58	**0.74**

Values in red indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

By not including the Buddy Check Test in OBSGRID, the domain-wide model performance decreased slightly for most of the wind speed, wind direction, and temperature bias statistics. The biggest degradation of model in performance is in the mixing ratio. However, within the Uinta Basin, all statistic variables improved by not using the Buddy Check Test. The most significant improvement of model performance in the Uinta Basin is in temperature.

Based upon the findings in Sens6, applying observation nudging in WRF for water vapor mixing ratio (RH) degraded model performance significantly for mixing ratio.

Sensitivity Test 7 Test Results: No Surface Nudging of Relative Humidity

The February model performance statistics for Base Case, Sens1, and Sens2 tests are shown in **Figures A-18** and **A-19**, for the full 4-km domain and Uinta Basin Study Area, respectively. The daily biases for Base Case, Sens6, and Sens7 are shown in black, blue, and red, respectively. In **Figure A-18**, the daily wind speed bias and wind direction bias suggest that the three sensitivity runs performed comparably. For temperature and mixing ratio, Sens6 performed worst while Base and Sens7 showed similar performance. Within the Uinta Basin (**Figure A-19**), the daily wind speed bias suggests that all three sensitivity runs performed relatively similarly. The daily wind direction tests are inconclusive with respect to the best performing configuration as was for Uinta specific. The daily temperature bias for Sens7 was better than the performance of Sens6; however, both simulations performed much better than the Base configuration. The daily humidity performance for Sens7 is better than the Base and Sens6 configurations.

Table A-14 compares the monthly average statistics for the Base Case, Sens6, and Sens7 results for February. Based on these results, the model performance for Sens7 showed improved performance over the Base Case and Sens6. For the Uinta Basin specific summary statistics, Sens6 performed the best with Sens7 a close second. Both simulations (Sens6 and Sens7) show a significant in model performance over the Base Case. The biggest improvement in performance is in temperature.

Table A-15 compares the monthly average statistics for Sens1 to Sens7 for July. The Sens7 configuration for July uses Sens1 (non-winter) configuration turning off Buddy Check Test in OBSGRID and RH observational nudging. The summary statistics for the entire domain show that Sens7 performance is improved relative to Sens1 for most parameters. Within the Uinta Basin, Sens7 performance is improved significantly relative to Sens1 for wind direction and temperature.

The February time series plots for Sens7 are shown relative to observations in **Figures A-20** and **A-21**, for the full 4-km domain and Uinta Basin Study Area, respectively. Overall, the 4-km domain is able to reproduce the synoptic and diurnal pattern of the observations. Although the model produces slightly slower wind speed for all days and is drier on a number of days (shown with a negative RH bias). Comparison of **Figures A-16** and **Figure A-20** showed that Sens6 and Sens7 performed comparably for wind speed and wind direction. For temperature, Sens7 maintained the good performance in Sens6 and improved the under-prediction in early evening hours. The Sens7 mixing ratio performance for mixing ratio improved significantly over Sens6. In the Uinta Basin, temperature and mixing ratio in Sens7 simulated the diurnal pattern of the observation better than those in Sens6.

The July time series plots for Sens7 are shown relative to observations in **Figures A-22** and **A-23**, for the full 4-km domain and Uinta Basin Study Area, respectively. Overall, the 4-km domain-wide time series plots (**Figure A-22**) show that simulation for wind speed, wind direction, and temperature followed the synoptic and diurnal pattern of the observations. The model over-predicted daily maximum wind speed and under-predicted minimum wind speed, and performs slightly worse in following the diurnal pattern for mixing ratio. **Figure A-23** showed that the model was able to simulate the diurnal and synoptic patter of the observations located within the Uinta Basin. The model has some difficulty matching the hourly fluctuation of mixing ratio from observations, but the model simulation generally follows the diurnal pattern.

Based on the results from Sens5, Sens6, and Sens7, the annual simulation was run with the following configuration:

- The physics options from the Base Case for the months of January, February March, and December (winter months);
- The physics options from Sens1 for all other non-winter months; and
- The revised observation nudging dataset was used throughout the simulation with the Buddy Check Test turned off in OBSGRID and RH was not nudged with observation data.

Table A-14 Model Performance for Base Case, Sens6, and Sens7 in February

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values			Uinta Basin Average Values		
		(Tesche 2002)	Complex Terrain	Base	Sens6	Sens7	Base	Sens6	Sens7
Wind Speed (m/s)	RMSE	≤ 2	≤ 2.5	***1.84***	***1.63***	***1.57***	***1.11***	***1.01***	***0.98***
	Bias	± 0.5		-0.59	-0.72	-0.73	-0.63	**-0.29**	**-0.39**
	IOA	≥ 0.6		0.57	**0.64**	**0.67**	0.42	0.48	0.5
Wind Direction (deg)	Bias	± 10		**3.40**	**3.65**	**3.8**	8.67	**2.64**	**-3.48**
	Gross Error	≤ 30	≤ 55	58.51	43.39	*39.94*	92.90	58.47	59.□3
Temperature (K)	Bias	± 0.5	± 2	***-0.10***	*-1.31*	***-0.43***	3.84	*0.72*	*1.7*
	Gross Error	≤ 2	≤ 3.5	*2.65*	*2.36*	***1.87***	4.45	*2.08*	*2.26*
	IOA	≥ 0.8		**0.90**	**0.92**	**0.94**	0.59	**0.84**	**0.83**
Humidity (g/kg)	Bias	± 0.75	± 1	***-0.09***	-1.51	***-0.06***	***0.48***	***-0.11***	***0.53***
	Gross Error	≤ 2	≤ 2	***0.52***	*1.55*	***0.5***	***0.57***	***0.24***	***0.59***
	IOA	≥ 0.6		**0.80**	0.42	**0.82**	0.58	**0.74**	0.56

Values in red indicate best performance among the three runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

Table A-15 Model Performance for Sens1 and Sens7 in July

Parameter	Statistics	Statistical Benchmark		4-km Domain Average Values		Uinta Basin Average Values	
		(Tesche 2002)	Complex Terrain	Sens1	Sens7	Sens1	Sens7
Wind Speed (m/s)	RMSE	≤ 2	≤ 2.5	*2.45*	*2.32*	*2.30*	***1.88***
	Bias	± 0.5		**-0.43**	**-0.01**	**0.20**	**-0.40**
	IOA	≥ 0.6		**0.62**	**0.68**	**0.61**	**0.73**
Wind Direction (deg)	Bias	± 10		**4.27**	**1.63**	**-8.11**	**-1.76**
	Gross Error	≤ 30	≤ 55	54.41	*44.68*	73.48	*43.96*
Temperature (K)	Bias	± 0.5	± 2	**0.81**	*-0.03*	1.31	*0.53*
	Gross Error	≤ 2	≤ 3.5	2.63	***1.79***	2.41	***1.26***
	IOA	≥ 0.8		**0.93**	**0.97**	**0.91**	**0.98**
Humidity (g/kg)	Bias	± 0.75	± 1	0.17	***0.04***	-0.34	***-0.01***
	Gross Error	≤ 2	≤ 2	***1.33***	***1.32***	**1.12**	***1.24***
	IOA	≥ 0.6		**0.75**	**0.76**	0.54	0.56

Values in red indicate better performance between the two runs.

Bold indicates passing benchmark for Tesche 2002. Italics indicates passing benchmark for complex terrain.

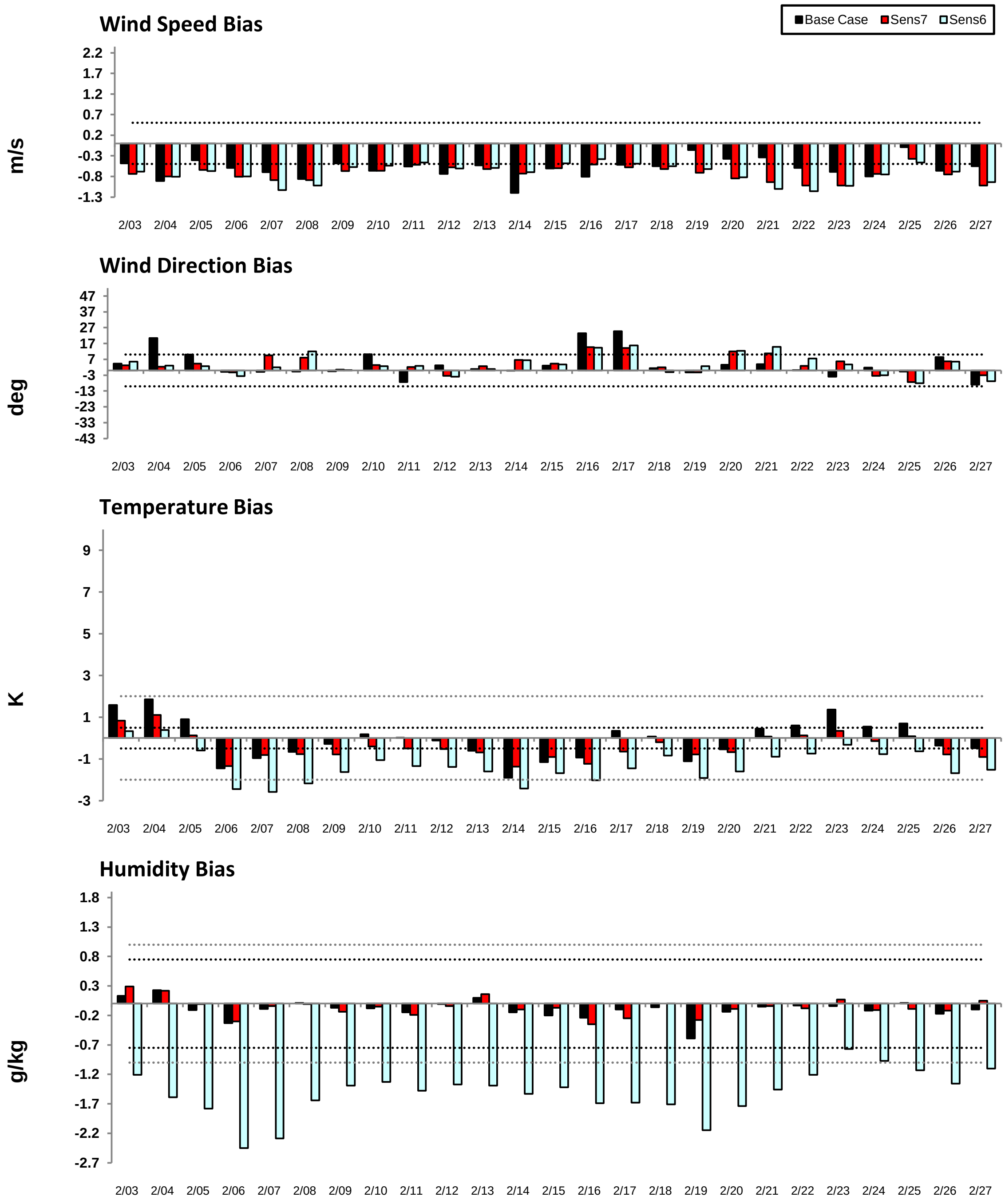


Figure A-18 4-km Domain Bias Bar Charts for Base Case, Sens6, and Sens7 during February, 2010

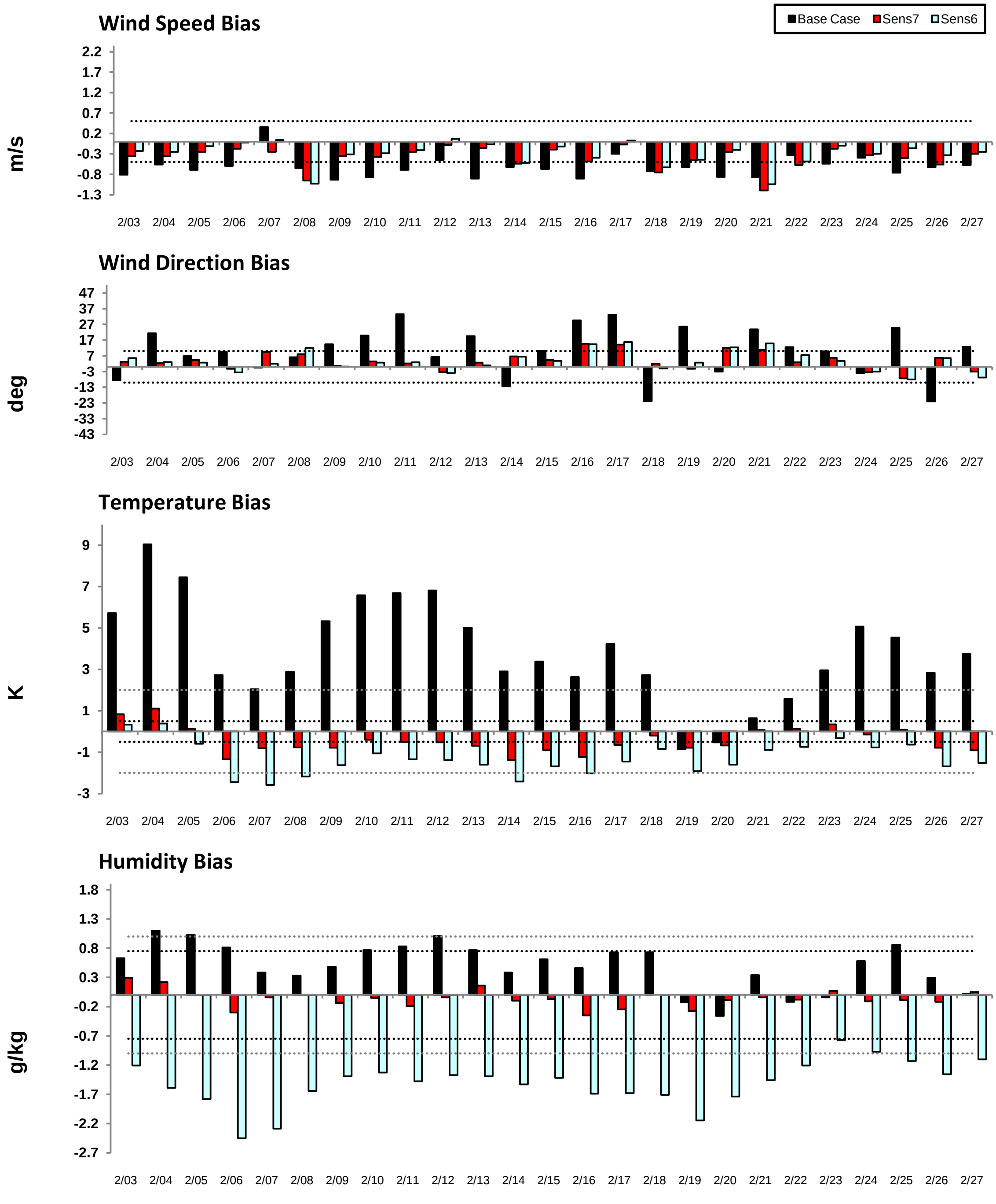
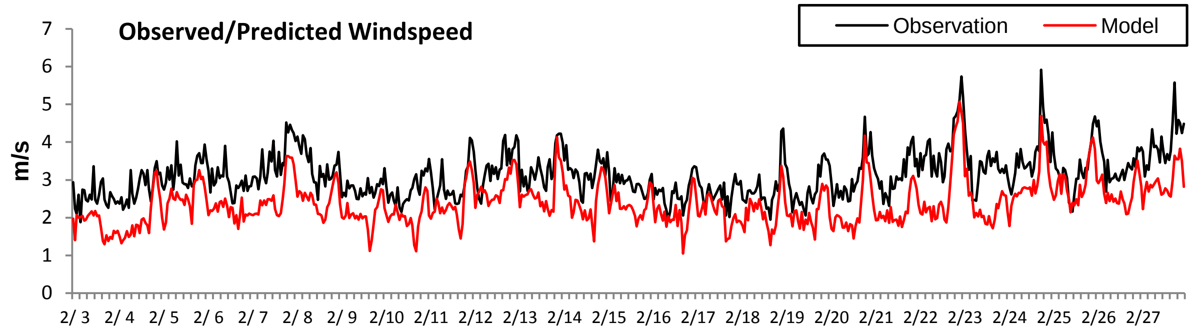
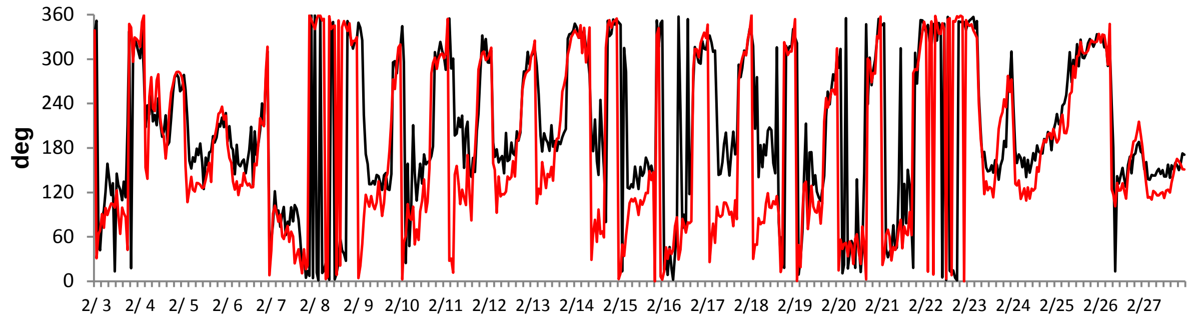


Figure A-19 Bias Bar Charts for Base Case, Sens6, and Sens7 during February, 2010 within the Uinta Basin

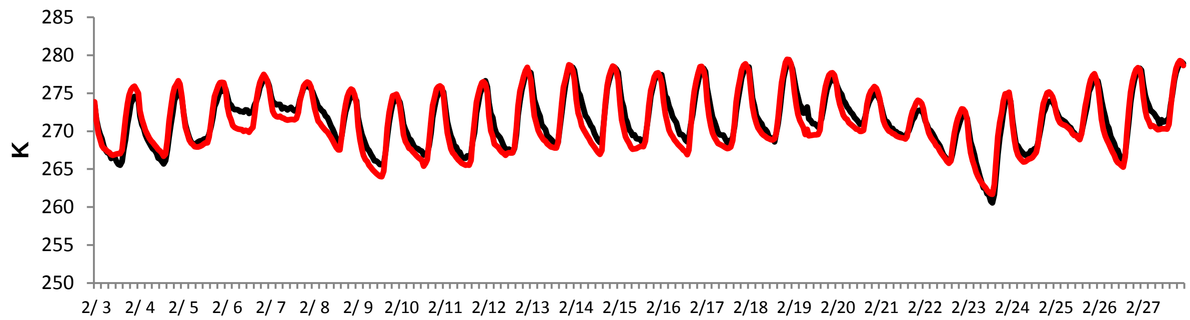
Wind Speed



Wind Direction



Temperature



Humidity

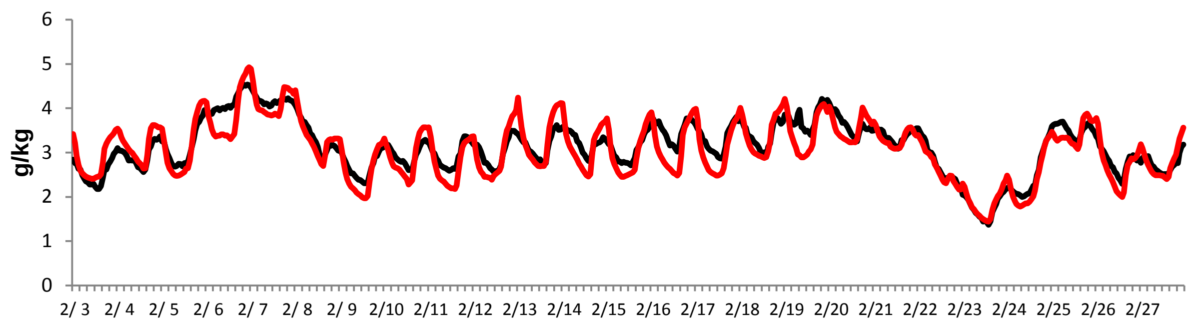
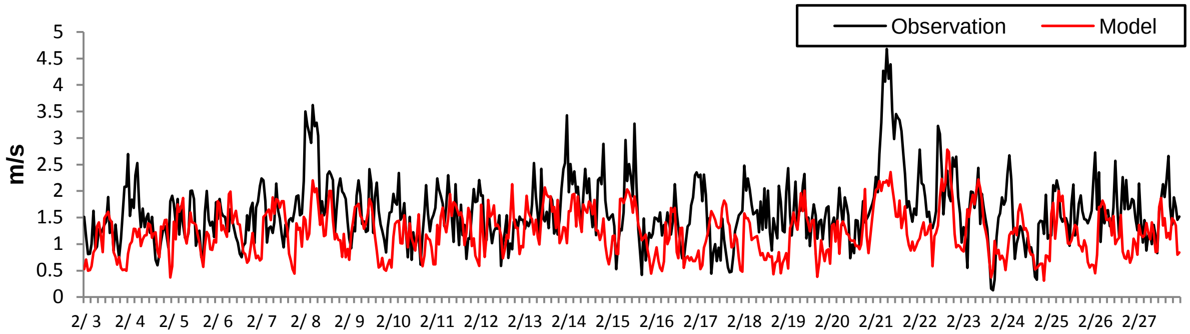
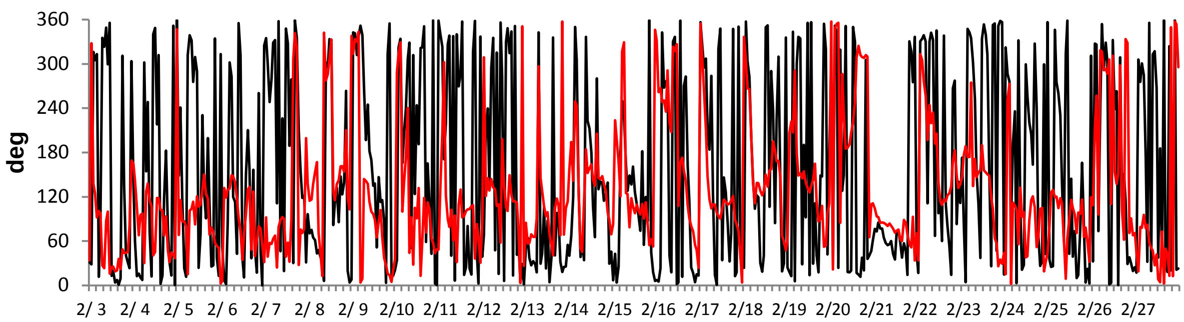


Figure A-20 4-km Domain Time Series for Sens7 during February

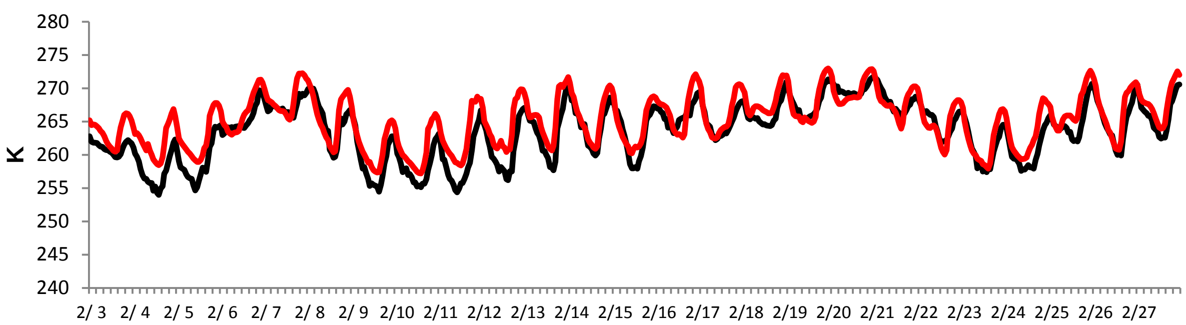
Wind Speed



Wind Direction



Temperature



Humidity

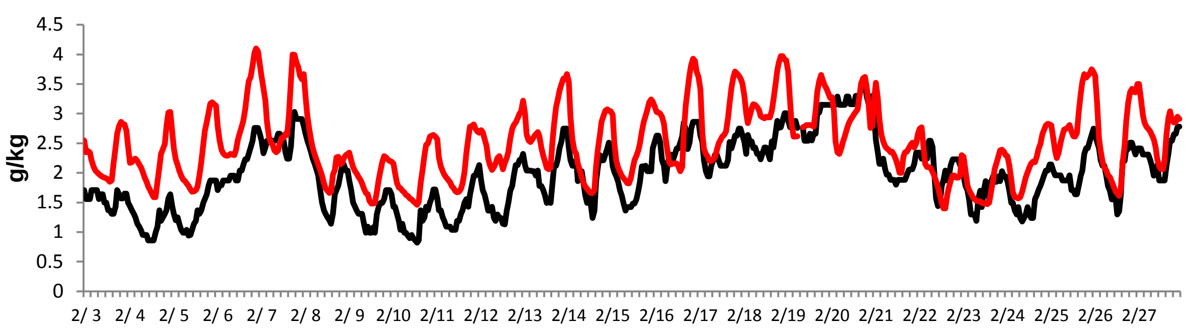
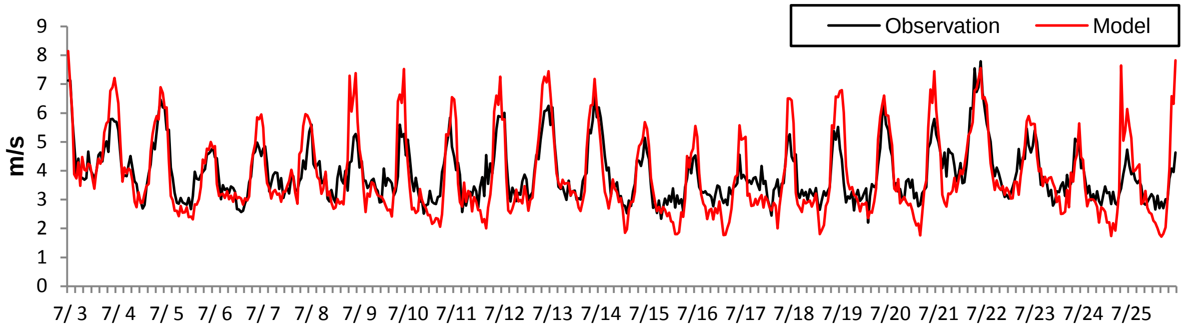
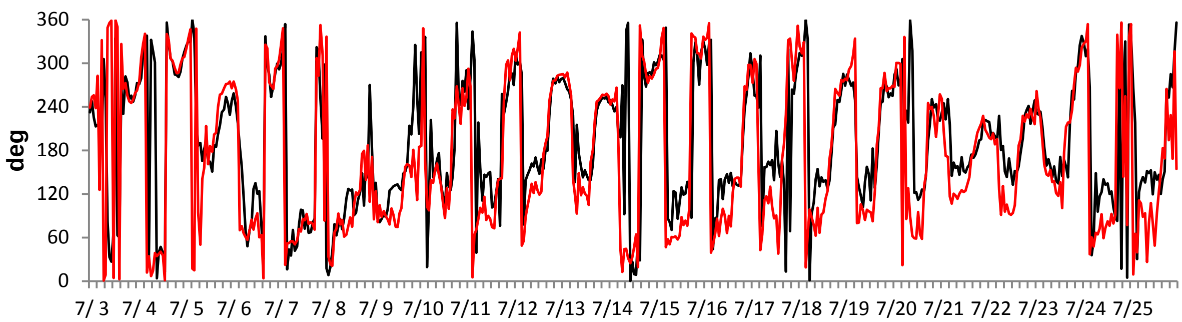


Figure A-21 Uinta Basin Time Series for Sens7 during February

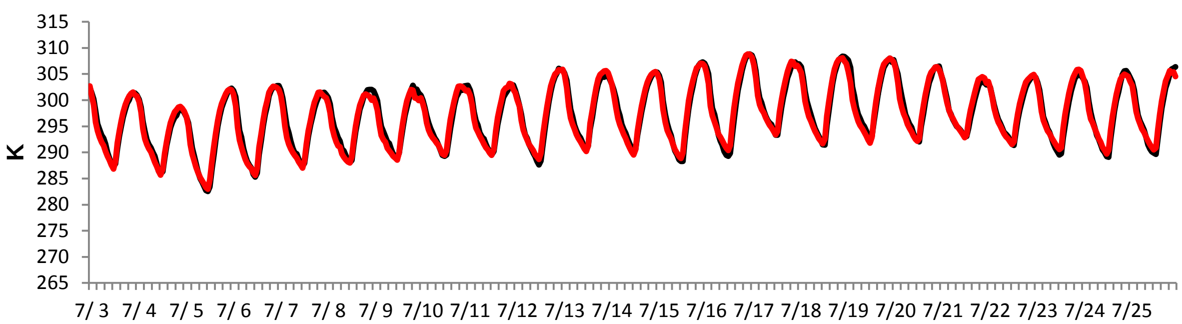
Wind Speed



Wind Direction



Temperature



Humidity

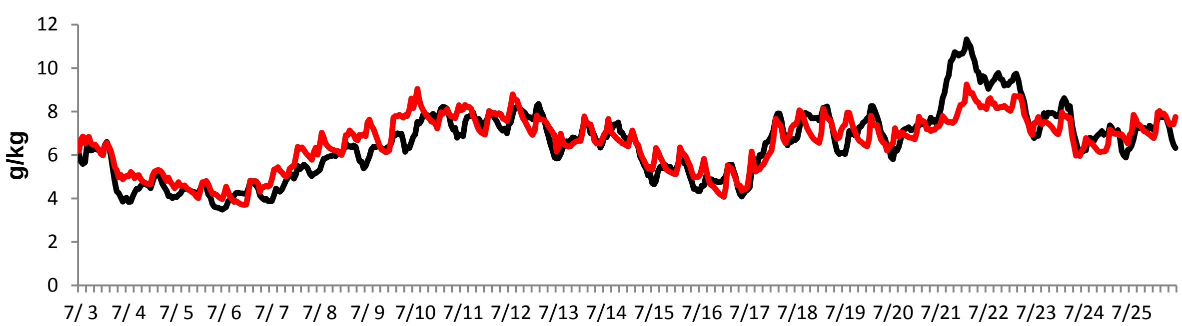
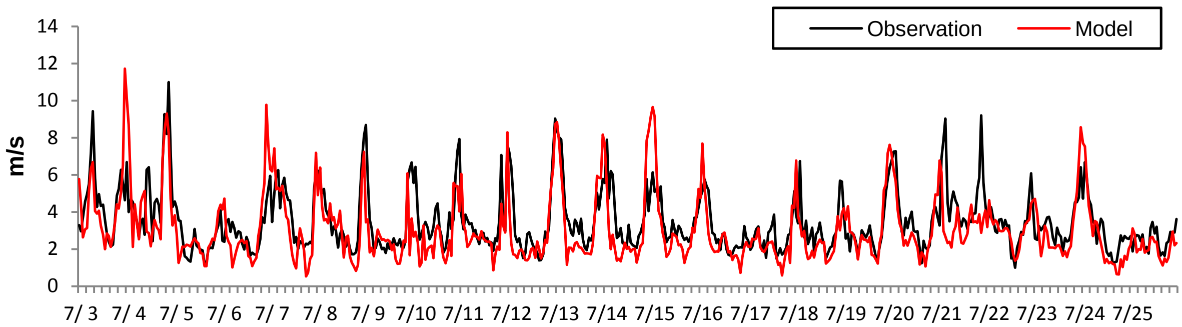
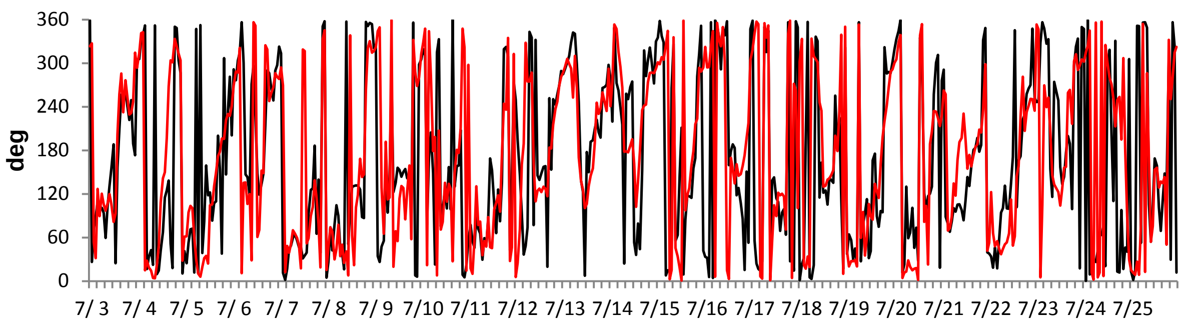


Figure A-22 4-km Domain Time Series for Sens7 during July

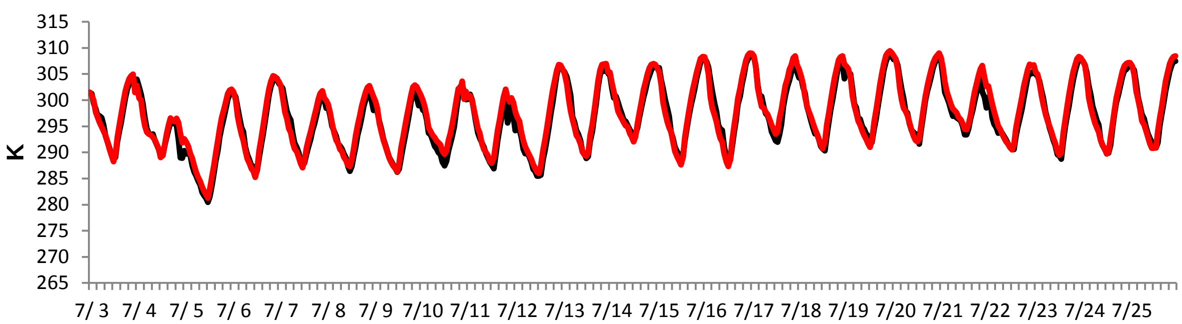
Wind Speed



Wind Direction



Temperature



Humidity

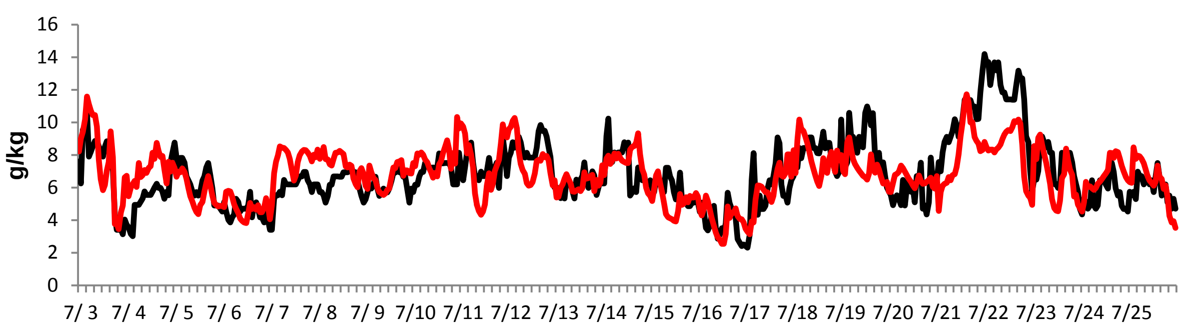


Figure A-23 Uinta Basin Time Series for Sens7 during July