Atmospheric and Environmental Research

131 Hartwell Avenue
Lexington, Massachusetts
02421-3126
USA
Tel: +1 781 761-2288
Fax: +1 781 761-2299

Final Report

High Resolution Meteorological Simulations of DISCOVER-AQ Houston

TCEQ Contract No. 582-15-50414
Work Order No. 582-15-54365-01

Submitted by:

Jennifer Hegarty, Rebecca Adams-Selin, and Thomas Nehrkorn
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02466

Correspondence to: jhegarty@aer.com

Prepared for:

Mark Estes
Texas Commission on Environmental Quality
Air Quality Division
Building E, Room 323
Austin, Texas 78711-3087

August 25, 2015

Document Change Record

Revision	Revision Date	Remarks
1.0	**10 August 2015**	**Draft Version of Final Report submitted to TCEQ as Deliverable 4.1**
2.0	**25 August 2015**	**Final Version of Final Report submitted to TCEQ as Deliverable 4.2**

Table of Contents

List Of Figures

List of Tables

Project Summary

The purpose of this project was to apply the Advanced Research Weather and Forecasting (WRF) model at high spatial and temporal resolution to investigate the fine-scale meteorological features influencing high ozone (O_3) levels in the Houston area during the DISCOVER-AQ (Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality) field campaign of September 2013. The WRF domain configuration included three nested grids of 36, 12, and 4 km centered over eastern Texas. An additional 1.33 km nested domain was centered over the Houston-Galveston Bay area. A series of WRF sensitivity runs were performed for the high ozone episode days of September 25th and 26th to determine the optimal WRF configuration for simulating the fine-scale meteorological features influencing these episodes. A key finding was that a WRF configuration with the combination of the Mellor Yamada Janjić (MYJ) planetary boundary layer (PBL) scheme and the Noah land surface model with an activated urban canopy model did a better job of simulating the development of the Galveston Bay and Gulf of Mexico sea breezes (which likely influenced the distribution of O_3 in the Houston area on September 25th) than a WRF configuration with the combination of the Asymmetric Convective Model (ACM2) PBL scheme coupled with the Pleim-Xiu LSM. Another finding was that imposing a realistic diurnal cycle based on buoy observations to the high-resolution sea surface temperature analysis input to WRF produced only minor improvements to the simulated winds along the Texas coastline during September 25th.

The optimal WRF configuration determined from the sensitivity runs was applied to simulate the meteorology for the 10-day period preceding and including the September 25th and 26th high O_3 episode days at a horizontal resolution of 1.33 km and at 10-minute output intervals. In addition the meteorology for the entire month of September 2013 was simulated at a horizontal resolution of 4 km and at 30-minute output intervals. The WRF output was validated with a quantitative statistical error analysis generatated by the WRF-MET software package and with qualitative comparisons of the 1.33 km WRF output with observations at locations impacted by the sea and bay breezes or high O_3 mixing ratios during the high O_3 episodes of September 25th and 26th.

Summary statistics of simulated vs. observed near-surface variables from the innermost 1.33 km domain indicated that the hourly domain-wide bias and root mean square error (rmse) in 2 m temperature and 10 m wind had considerable variation over the 10-day period, with the majority of times having a temperature bias/rmse below 1 to 2 K, wind speed bias between 0.5 and 1.5 m s^{-1}, and wind speed rmse below 2 m s^{-1}. Month-long WRF performance statistics for a sub-section of the 4 km WRF domain covering the Houston-Galveston area indicated a wind speed bias of 0.77 m s^{-1} and rmse of 1.92 m s^{-1} with both being greater over land than water. For temperature the bias and rmse were 0.34 K and 1.68 K, respectively, and were also greater over land. These statistics are consistent with those from other meteorological modeling studies in the Houston – Galveston area.

Comparisions of time series of WRF winds and temperature from the 1.33 km domain with observations at key locations indicated that WRF was generally able to capture the development of sea and bay breezes during the high O_3 episodes, but at some locations (i.e., Galveston Island on September 25th and 26th and the La Porte Airport on the 25th) the wind shift and temperature changes associated with the sea and bay breezes were delayed in the model by about 2-3 hours. However, observed and modeled winds were at times light (< 2.5 m s^{-1}) during these periods. At other locations on the western shore of Galveston Bay and at Conroe Airport northwest of Houston the WRF winds and temperatures tracked the observations well. Skies were mostly

clear during September 25th and 26th and WRF simulated solar radiation outputs were generally within 10 % of observed values.

The validated WRF output was processed to produce model-ready meteorological input fields for the CMAQ and CAMx chemical transport models. Both the raw WRF data and the processed data were delivered to the Texas Commission on Environmental Quality (TCEQ) for additional studies on the influence of fine-scale meteorological features on the air quality of the Houston area. In addition to the model data, all the plots and tables generated by WRF-MET for evaluating the WRF model performance are also being made available to TCEQ.

1. Introduction

The purpose of this work order was to develop an improved configuration of the Advanced Research Weather and Forecasting (WRF) model *[Skamarock and Klemp, 2008]* to provide validated, high spatial- and temporal-resolution meteorological fields for the September 2013 DISCOVER-AQ Houston campaign. The main objectives were to better simulate the fine-scale meteorology that contributed to the high ozone O_3 episodes in the Houston area during DISCOVER-AQ, and to use the high-resolution WRF output in conjunction with DISCOVER-AQ observations to gain a better understanding of the association between fine-scale meteorological features and high O_3 levels in the Houston area. This final report describes the work performed to achieve those objectives.

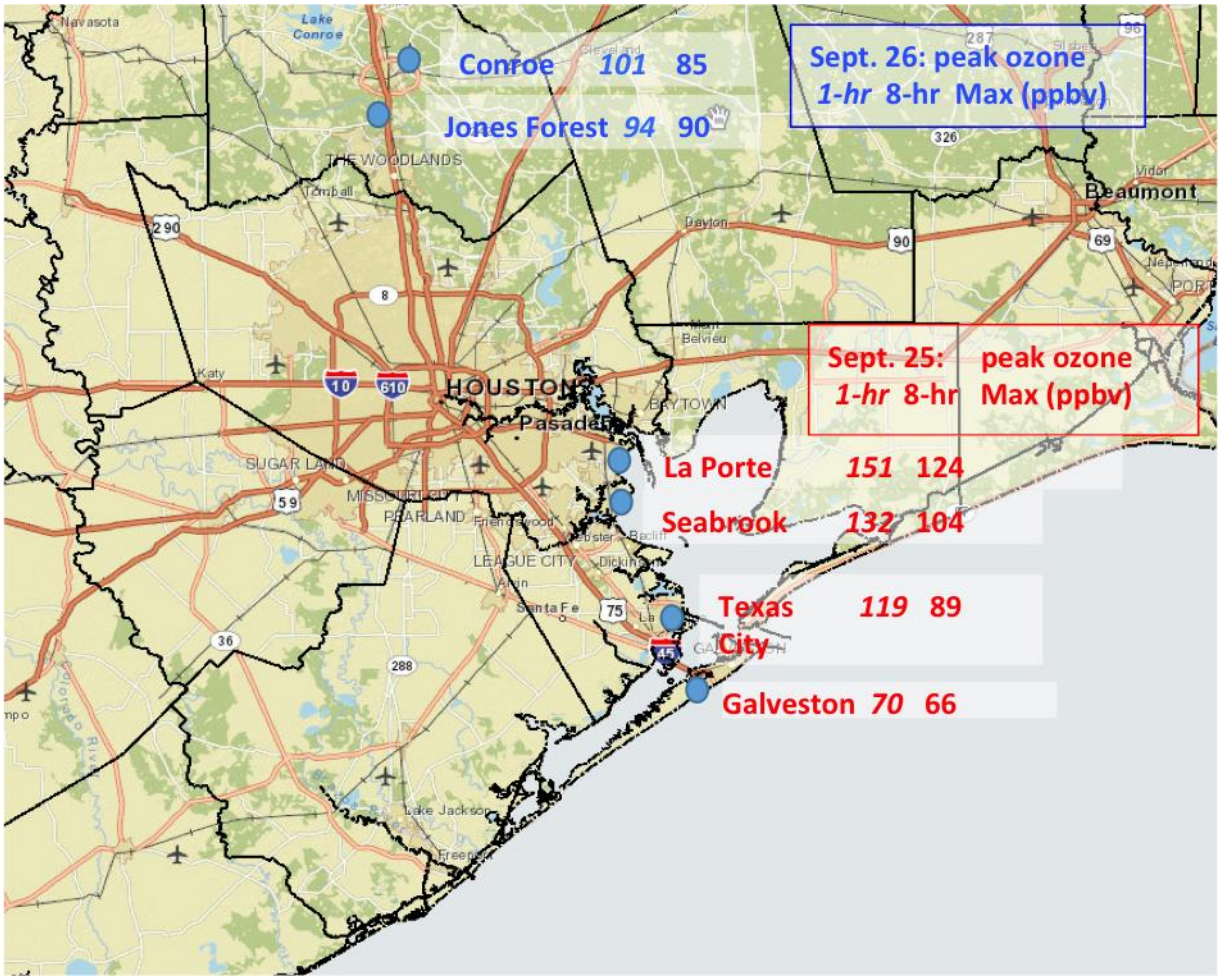

Figure 1. Maximum 1-hour (italics) and 8-hour average (MDA8) ozone mixing ratios (ppbv) for air monitoring stations during the high ozone episodes of September 25th (red) and 26th (blue), 2013. Ozone data obtained from TCEQ air monitoring website http://www.tceq.state.tx.us/cgi-bin/compliance/monops/site_info.pl and *Loughner et al. [2014]*.

During the Houston DISCOVER-AQ campaign, high O_3 episodes occurred on just two days. On September 25th, maximum daily average 8-hour ozone (MDA8) values above 120 ppbv were

recorded at monitoring stations near the western shore of Galveston Bay (Figure 1). On the 26th, an MDA8 of 90 ppbv was recorded northwest of Houston. Preliminary WRF-CMAQ simulations using a 4 km inner domain suggest that errors in WRF's representation of the bay breeze contribute to CMAQ incorrectly simulating the locations of these ozone maxima *[Loughner et al., 2014]*. Thus a focus of this work was to improve the ability of WRF to simulate mesoscale circulation features, such as the Galveston Bay breeze, that impacted these high ozone episodes. One of our refinements to the modeling study of *Loughner et al. [2014]* was the addition of a finer 1.33 km inner nest, as inadequate spatial resolution was cited as a potential cause of the errors in the preliminary simulations.

The ability of WRF to accurately simulate fine-scale meteorological features can be critically dependent on the specific physical parameterization schemes selected. Since the features of interest were at low altitude and precipitation was not a factor during the two-day high O_3 episode, we instead focused our attention on the planetary boundary layer (PBL) and land surface model (LSM) options and tested two different combinations as described in Section 2. One of the LSMs included the option of using an urban canopy model (UCM) *[Wang et al., 2012]* to parameterize the impacts of the dense urban landscape of Houston on the fine-scale circulation. We ran additional tests with and without this option to determine its impact.

The representation of the temperature gradient between land and water is critical to the accurate simulation of sea and bay breezes. Thus, the sea surface temperature (SST) analysis input to WRF for our Houston runs required high enough spatial resolution to adequately represent these strong gradients. Furthermore, since Galveston Bay and the Gulf of Mexico near Galveston Island are shallow, their SSTs could potentially have diurnal temperature variations large enough to influence the bay and sea breeze circulations. Therefore in this project we incorporated a high resolution (~1 km) SST analysis and investigated new methods to include the diurnal variation of SSTs in Galveston Bay and the adjacent Gulf of Mexico.

Each WRF simulation of the high O_3 episode days was evaluated using temperature and wind observations and mixing depths derived from Doppler wind profilers deployed during the DISCOVER-AQ field campaign, as described in Sections 2.2 to 2.4. The results of these evaluations formed the basis of the WRF configuration options used for simulating the entire DISCOVER-AQ period of September 2013 (see Section 3). The final WRF outputs generated from these runs were provided to TCEQ for use in chemical transport modeling studies of the processes leading to O_3 exceedances in Houston.

2. Optimizing the WRF Configuration

2.1. Baseline WRF Configuration with High-Resolution Nest

The WRF horizontal grid was configured with three nest levels of 36, 12, and 4 km that match the Rider 8 modeling domains, and an additional 1.33 km grid centered over the Houston-Galveston Bay area (Figure 2). The vertical grid also follows the Rider 8 configuration with a total of 43 layers, with 17 below 2 km above ground level (AGL). To better represent the sea surface temperature (SST), and thus the location, strength, orientation, and timing of the Galveston Bay breeze, we have incorporated into our WRF modeling system the Multi-Sensor Ultra-high Resolution (MUR) Level 4 (L4) SST analysis *[Chin et al., 2013]* available daily on a global grid of ~1 km resolution. The input SST data to the MUR L4 algorithm are generated from multiple infrared (IR) and microwave (MW) satellite measurements and are provided as Level 2 (L2) swath data by the Group for High Resolution Sea Surface Temperature (GHRSST, https://www.ghrsst.org).

The optimized WRF configuration was determined through a series of WRF sensitivity runs for September 25 and 26, 2013, during which the highest surface ozone levels of DISCOVER-AQ Houston were observed. The WRF configuration details for each sensitivity run are summarized in Table 1.

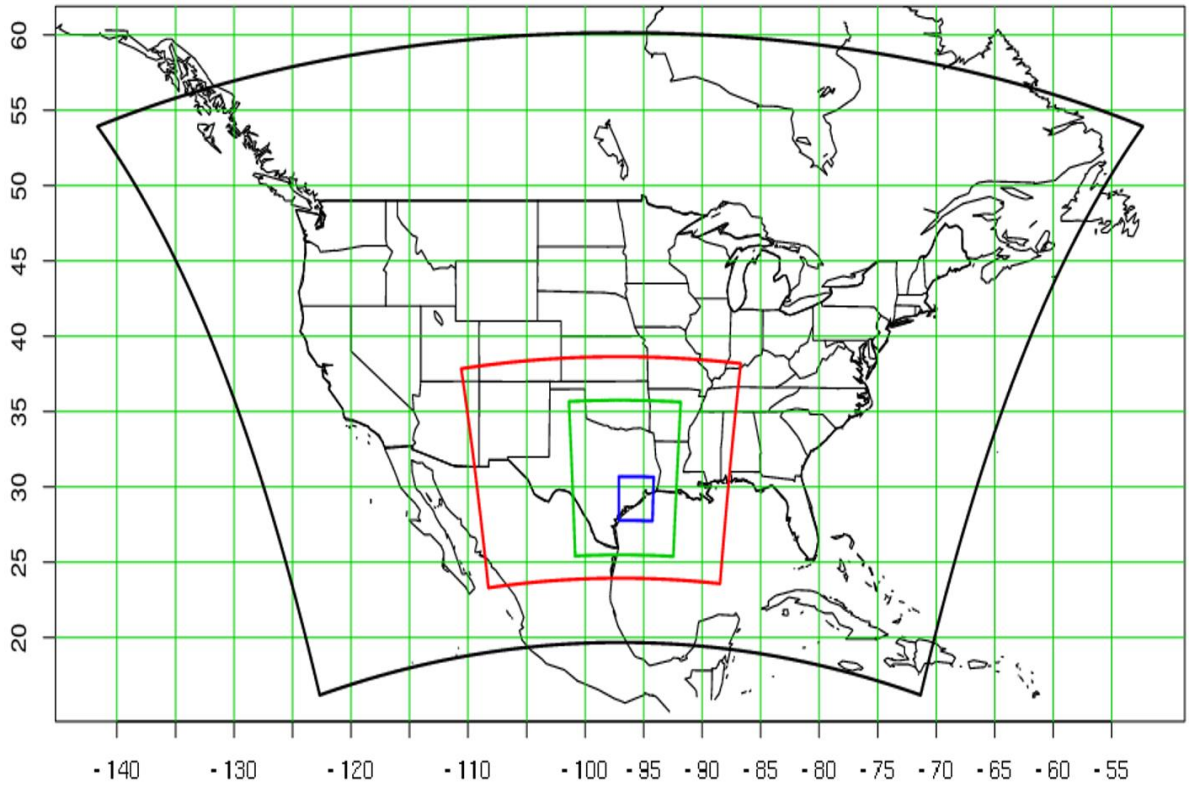

Figure 2. WRF domains; 36 km black, 12 km red, 4 km green, and 1.33 km blue.

2.2. WRF Sensitivity to PBL and LSM Parameterization Schemes

The first WRF configuration tested (CFG1) uses the Mellor Yamada Janjić planetary boundary layer (PBL) scheme *(*MYJ*, Mellor and Yamada [1982]; Janjić [2002])* with the Noah land surface model *[Chen and Dudhia, 2001; Mitchell et al., 2005]* that included an activated urban canopy model (UCM). The UCM [*Chen et al., 2011a*] accounts for the interaction of the urban landscape of the Houston (e.g., residential and commercial buildings of varying heights, paved surfaces) and the atmosphere. While the detailed flow around individual buildings cannot be represented in a mesoscale model, the UCM represents the bulk effects of the urban "canopy", an analogy to vegetative canopies such as forests. To take advantage of the UCM's capabilities, we incorporated into WRF a 30-meter resolution land use dataset based on the 2006 National Land Cover Database (NLCD), combined with detailed urban parameter databases compiled for 44 US cities (NUDAPT-44), which includes Houston. We have successfully used this high-resolution urban modeling approach in our previous WRF runs for Salt Lake City *[Nehrkorn et al., 2013]* and for the northeastern US urban corridor *[Nehrkorn et al., 2015]*.

The second configuration (CFG2) differed from CFG1 in that the Asymmetric Convective Model (ACM2, *Pleim [2007]*) planetary boundary layer (PBL) scheme coupled with the Pleim-Xiu Land Surface Model (PXLSM, *Pleim and Xiu [1995]; Xiu and Pleim [2001]*) were used instead. One motivation for choosing the ACM2 coupled with the PXLSM is that this combination is frequently used by EPA and other researchers in air quality simulations. This combination was used by *Loughner et al. [2014]* to simulate the DISCOVER-AQ high ozone episodes in Houston but at a lower resolution; thus including it in our higher-resolution (1.33 km) tests helped to clarify the relative importance of resolution versus physics. The PXLSM enables indirect nudging of soil moisture and deep soil temperature in WRF based on errors in modeled 2-m temperature and relative humidity and the gridded analyses of surface-based observations *[Pleim and Xiu, 2003]* and this can be helpful in addressing difficulties in the initialization and tracking of long-term evolution of the soil moisture fields. However, this option is likely to have little impact on our results since our simulations are re-initialized daily and our focus is on coastal and urban areas rather than highly vegetated areas. Therefore we have not included soil nudging in these runs. Another important difference with CFG1 is that the UCM is not used, as it is only available as an extension of the Noah LSM.

The CFG1 WRF simulations produced a notably more accurate representation of the sea breeze on September 25th (Figures 3, 4, and 5) than the preliminary result of *Loughner et al. [2014]*. However, our high-resolution simulations using CFG2 that used the same PBL scheme and LSM as *Loughner et al. [2014]* did not do as well in capturing the evolution of these features. This result suggests that the choice of phycical parameterizations rather than increased resolution was a more important factor in improving the simuations. On Galveston Island, the CFG2 sea breeze did not develop until about 5 hours after it was observed and about 3 hours later than in CFG1 (Figure 6). At the La Porte Airport on the western shore of Galveston Bay (where the highest MDA8 O_3 was recorded), there were no significant differences in the simulated winds between the simulations, but the CFG1 2 m temperatures during the afternoon were more closely matched to the observations (Figures 7 and 8). Furthermore, the MYJ scheme used in CFG1, though featuring more short-term fluctuations, does a better job capturing the time evolution of the PBL than the ACM2 used in CFG2 based on mixing heights estimated from the Doppler Wind profiler deployed at La Porte during DISCOVER-AQ (Figures 9 and 10). As shown in Figure

10, the ACM2 PBL peaks and collapses about 3 hours too early at this location, which could have potentially important impacts on the dilution of pollutants in an air quality simulation.

Table 1. WRF configuration options with differences highlighted.

ARW Version 3.6.1 Model Options				
	CFG1	**CFG2**	**CFG3**	**CFG4**
Initial and Boundary Conditions	NARR 32 km	NARR 32 km	NARR 32 km	NARR 32 km
Initialization / Spin-up	0000 UTC daily re-initialization at 30-hour runs, discard first 6.	0000 UTC daily re-initialization at 30-hour runs, discard first 6.	0000 UTC daily re-initialization at 30-hour runs, discard first 6.	0000 UTC daily re-initialization at 30-hour runs, discard first 6.
Radiation	RRTMG *[Iacono et al., 2008]*	RRTMG *[Iacono et al., 2008]*	RRTMG *[Iacono et al., 2008]*	RRTMG *[Iacono et al., 2008*
Surface Layer	Eta similarity based on Monin-Obukov *[Wang et al., 2012]*	Pleim-Xiu	Eta similarity based on Monin-Obukov *[Wang et al., 2012]*	Eta similarity based on Monin-Obukov *[Wang et al., 2012]*
Land Surface Model	Noah	Pleim-Xiu	Noah	Noah
Planetary Boundary Layer	MYJ	ACM2	MYJ	MYJ
Urban Surface Physics	Single layer UCM	None	Single layer UCM	None
Cumulus Parameterization	Grell-Freitas [*Grell and Freitas, 2014]*	Grell-Freitas *[Grell and Freitas, 2014]*	Grell-Freitas *[Grell and Freitas, 2014]*	Grell-Freitas *[Grell and Freitas, 2014]*
Cloud Microphysics	Lin et al. *[Lin et al., 2003; Chen and Sun, 2002]*	Lin et al. *[Lin et al., 2003; Chen and Sun, 2002]*	Lin et al. *[Lin et al., 2003; Chen and Sun, 2002]*	Lin et al. *[Lin et al., 2003; Chen and Sun, 2002]*
Nudging	Analysis nudging above PBL every 3 hours, 36 km domain only	Analysis nudging above PBL, every 3 hours, 36 km domain only	Analysis nudging above PBL, every 3 hours, 36 km domain only	Analysis nudging above PBL every 3 hours, 36 km domain only
Feedback	2-way	2-way	2-way	2-way
SST	MUR ~1km resolution SST, smoothed daily transition, no diurnal cycle	MUR ~1km resolution SST, smoothed daily transition, no diurnal cycle	MUR ~1km resolution SST, with superimposed diurnal cycle	MUR ~1km resolution SST, with superimposed diurnal cycle

On the 26th, the sea breeze developed at ~1700 UTC on Galveston Island, and both CFG1 and CFG2 seem to capture this timing, but the CFG1 winds are faster by ~1 m s^{-1} and are closer to the observed wind speeds of ~6 m s^{-1} (Figure 11 and 12). At the Conroe Airport northwest of

Houston (near the vicinity of the MDA8 O_3 maximum on the 26th) the CFG2 wind speeds are a little better match to the observed winds than those of CFG1 (Figures 13 and 14) and both configurations did a good job of matching the southeasterly wind directions. However the CFG1 temperatures are a notably better match to the observations, particularly during the morning and early afternoon hours when ozone production was likely high.

2.3. WRF Sensitivity to SST Diurnal Cycle

The third configuration (CFG3) is the same as CFG1 except for the addition of a diurnal cycle applied to the sea surface temperature (SST) inputs to WRF. We have determined from buoys that the SST in the shallow waters (~10 m in depth) adjacent to Galveston Island has a diurnal cycle with a phase corresponding to maximum values occurring near 2200 UTC, and an amplitude of ~2 K. This amplitude decreases to ~0.5 K by ~200 km offshore in the Gulf of Mexico where the water depth is over 100 m. For this test we modified the daily values from the high-resolution SST analysis by imposing a diurnal cycle with an amplitude based on the water depth as determined from the World Ocean Atlas 2013 database (https://www.nodc.noaa.gov/OC5/indprod.html).

Our analysis of the CFG3 simulations indicates that the inclusion of the SST diurnal cycle had only a minor impact. For example, on September 25th the winds at Galveston appeared to be of the same accuracy in CFG3 as in CFG1 while the CFG3 afternoon temperatures were about 1 K warmer and closer to the observed temperatures (see Figures 5 and 15). However, at the La Porte observation station there did not appear to be any noticeable differences between the simulations (see Figures 7 and 16). As the inclusion of the diurnal cycle did not produce any unexpected negative impacts on the simulations and produced some minor improvements in the surface temperatures, we retained this feature as part of the WRF configuration for the production runs. However, since the production runs span an entire month we made some modifications to account for the daily variability of the amplitude of the SST diurnal cycle as described in Section 3.2.

2.4. Impact of UCM

Given the results of the CFG1 and CFG2 simulations, it appears that CFG1 is better for simulating the fine-scale meteorological features contributing to the high ozone episodes observed during DISCOVER-AQ Houston. To determine if the cause of the simulation differences was that the UCM available in CFG1 produced a more realistic simulation of the winds and temperatures over the Houston urban area, thereby enabling a better simulation of the sea breeze and its inland influence, we performed an additional sensitivity simulation with CFG4 that used the same PBL and LSM as in CFG1 but without the UCM. This simulation also included the diurnal SST cycle, but as discussed in Section 2.3 this had a minimal impact on the development of the sea breeze.

Comparison of the CFG3 and CFG4 runs (Figure 17 and Figure 18, respectively) reveals that including the UCM is useful in capturing the evolution of the diurnal temperature cycle over Houston. The 2 m temperatures are approximately 2 K warmer in the Houston area in the CFG4 run at the start of peak heating. Stronger northerly/offshore flow is also evident over the Houston region. The temperature differences are particularly apparent at the La Porte station throughout the entire simulation time. For example, in Figure 16 (CFG3), nighttime temperatures reach a low of approximately 24°C between 1000 and 1300 UTC on September 25th. That is still somewhat warmer than the observed low of 21°C, but the CFG4 simulation

(Figure 19) only reaches a low of 25.5°C during that time period. The offshore flow from 1000 to 2000 UTC in the CFG4 simulation at La Porte was also stronger than both the CFG3 simulation and what was observed.

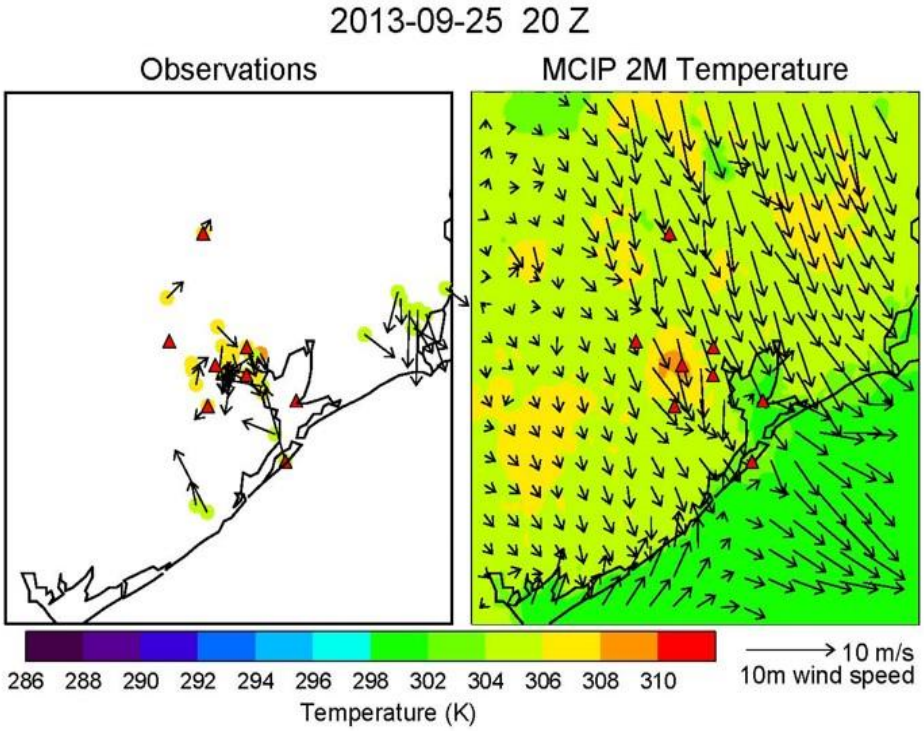

Figure 3. Observations (left) and WRF 4 km simulation (right) of 2 m temperature (color dots and fill) and 10 m winds (black arrows) for 2000 UTC September 25, 2013 from *Loughner et al. [2014]*.

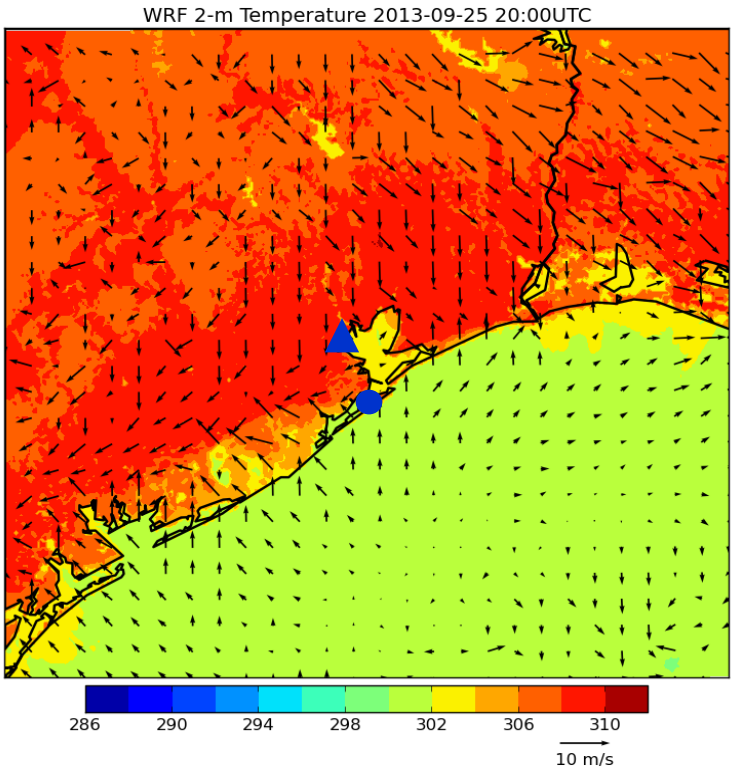

Figure 4. CFG1 WRF 1.33 km simulation of 2 m temperature (color shaded) and 10 m winds (black arrows) for 2000 UTC September 25, 2013. Blue dot indicates the location of Galveston 99th St. station shown in Figure 5, and blue triangle indicates the location of the La Porte Airport station shown in Figure 7.

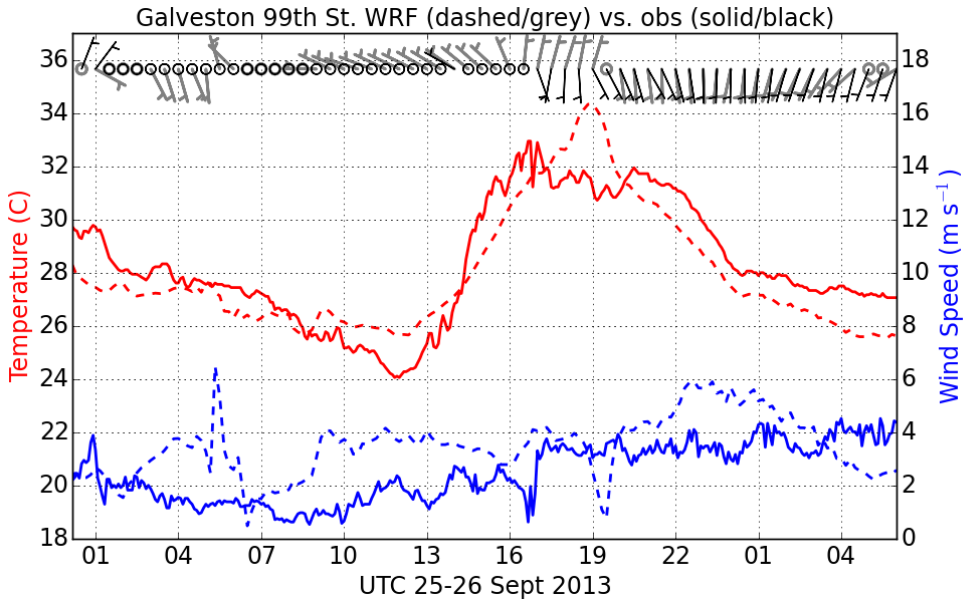

Figure 5. Observed (solid/black) and WRF CFG1 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at 99th Street on Galveston Island for starting 0000 UTC September 25, 2013.

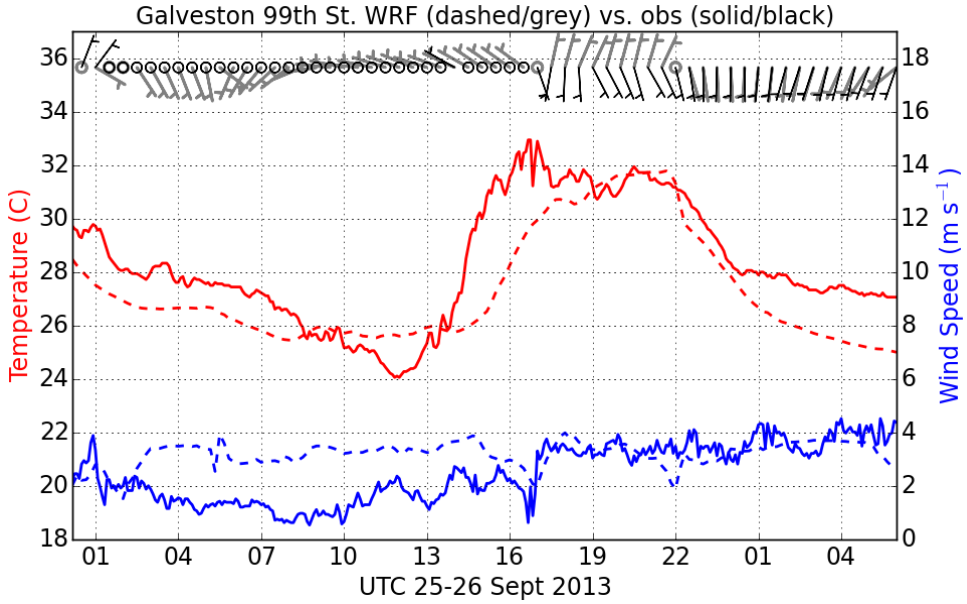

Figure 6. Same as Figure 5 but for WRF CFG2.

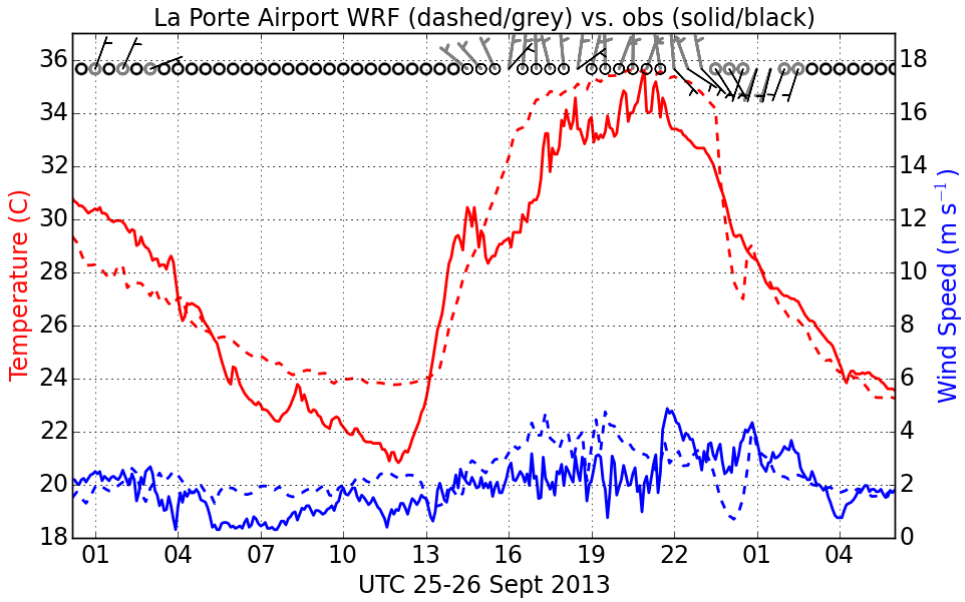

Figure 7. Observed (solid /black) and WRF CFG1 1.33 km (dashed /grey) time series of 10 m winds and 2 m temperature for the observation station at La Porte Airport starting 0000 UTC September 25, 2013.

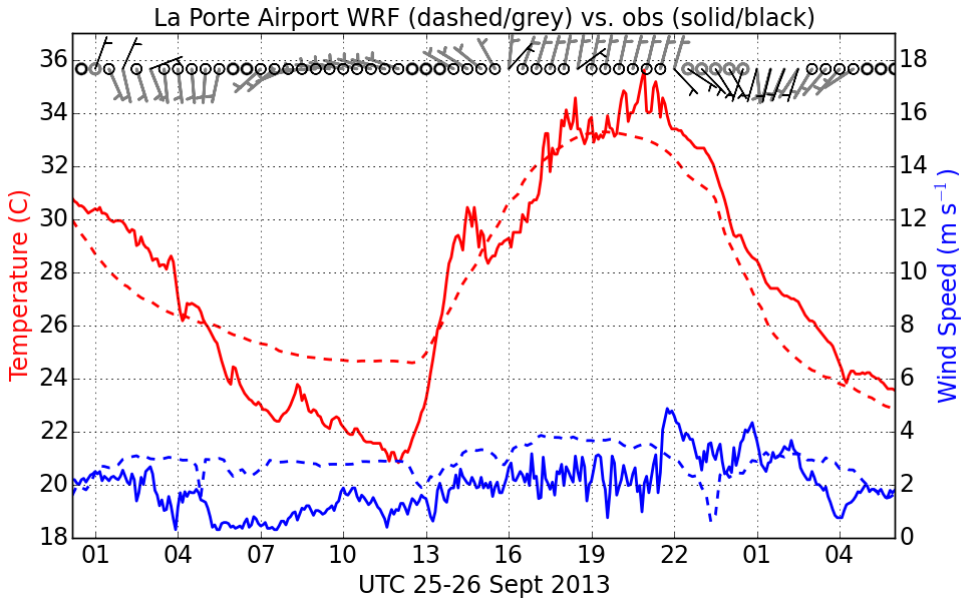

Figure 8. Same as Figure 7 but for WRF CFG2.

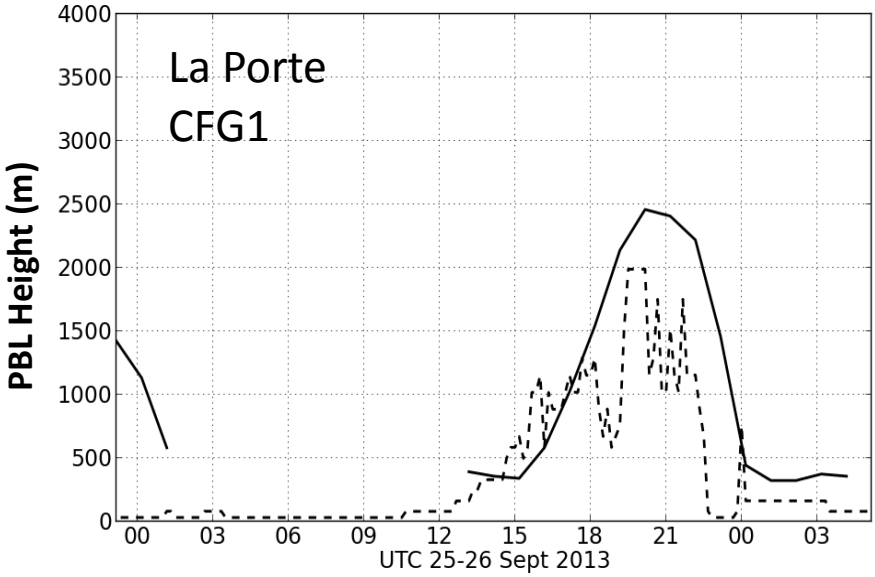

Figure 9. Doppler profiler mixing height (solid) and WRF CFG1 PBL height (dashed) at La Porte starting 0000 UTC September 25, 2013.

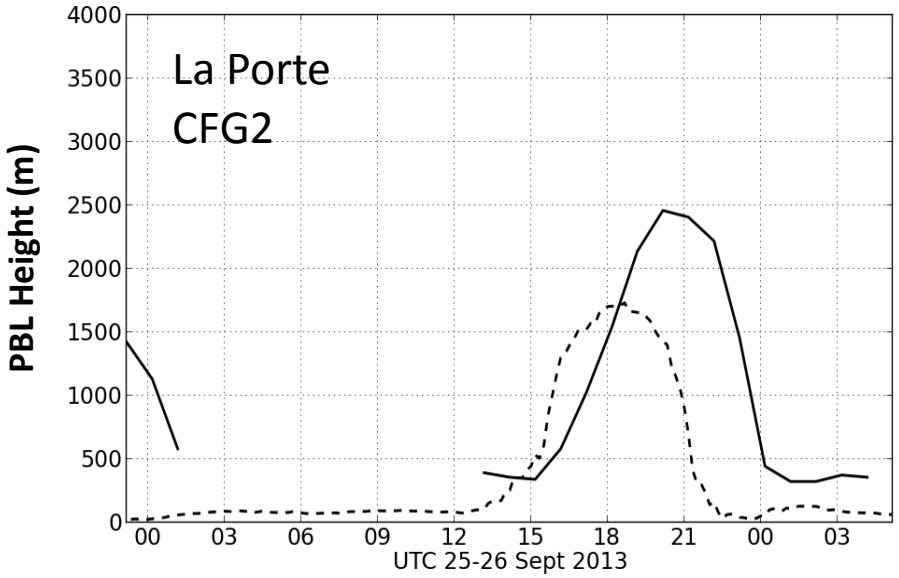

Figure 10. Same as Figure 9 but for WRF CFG2.

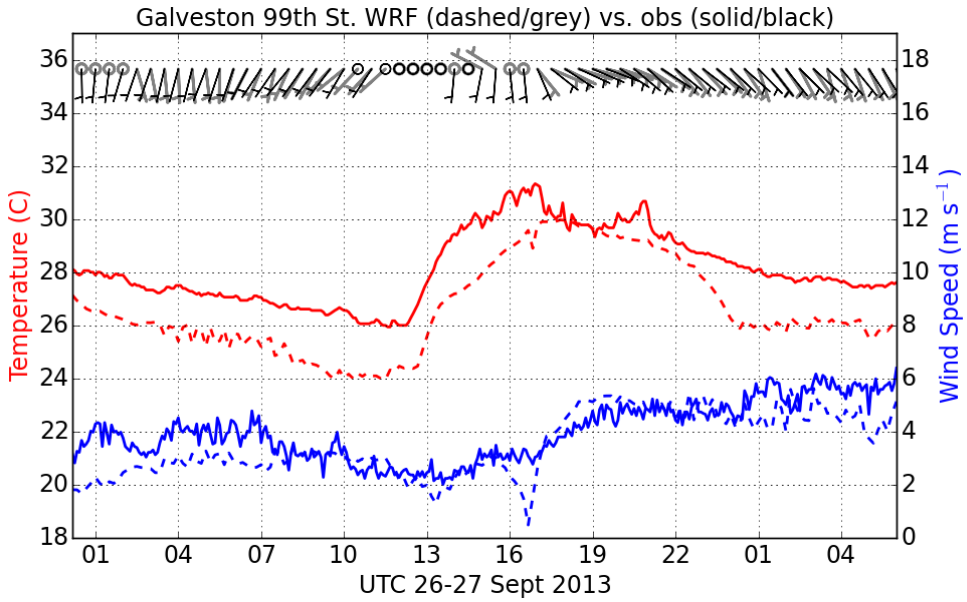

Figure 11. Observed (solid/black) and WRF CFG1 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at 99th Street on Galveston Island for starting 0000 UTC September 26, 2013.

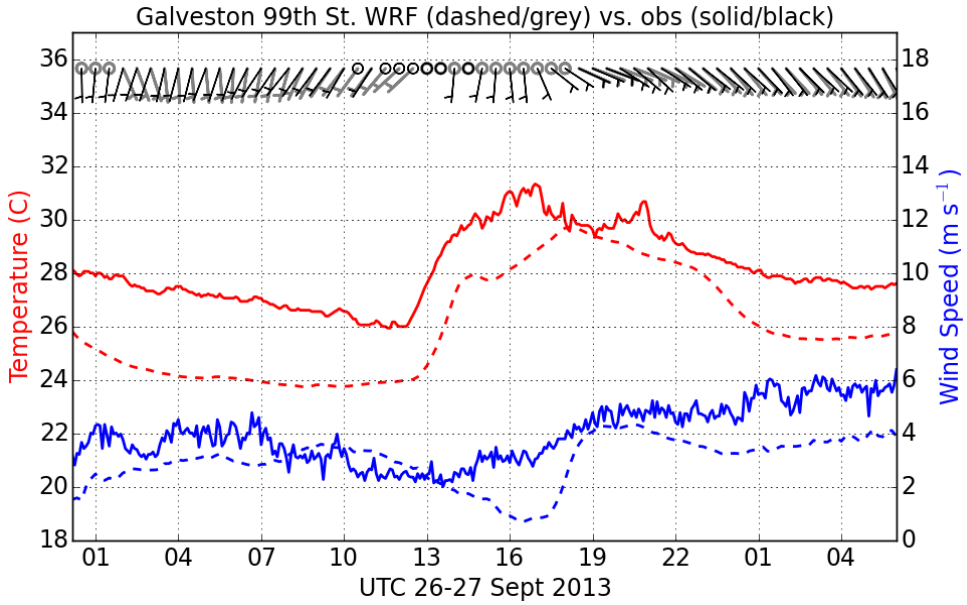

Figure 12. Same as Figure 11 but for WRF CFG2.

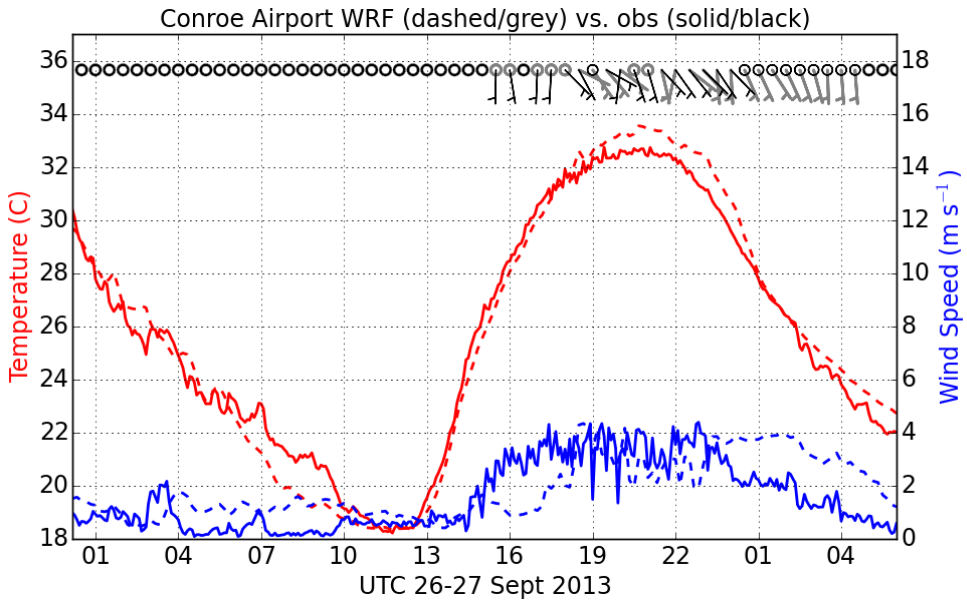

Figure 13. Observed (solid/black) and WRF CFG1 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Conroe Airport starting 0000 UTC September 26, 2013.

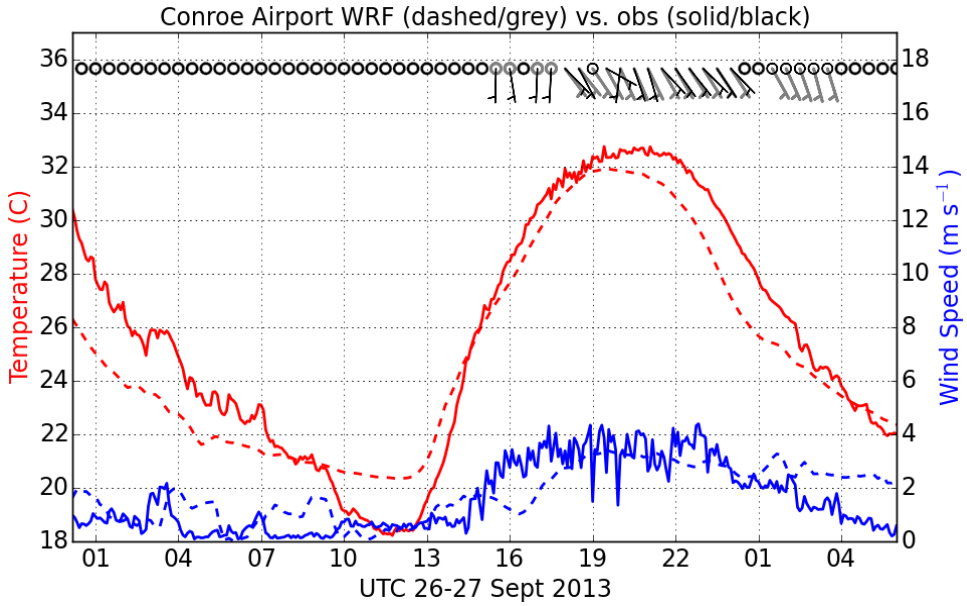

Figure 14. Same as Figure 12 but for WRF CFG2.

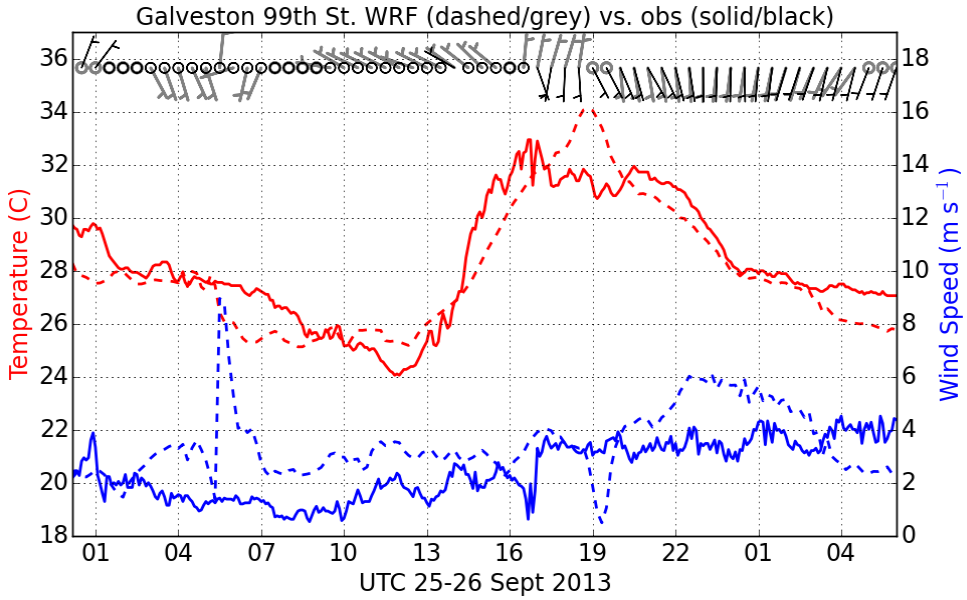

Figure 15. Observed (solid /black) and WRF CFG3 1.33 km (dashed /grey) time series of 10 m winds and 2 m temperature for the observation station at 99th Street on Galveston Island starting 0000 UTC September 25, 2013.

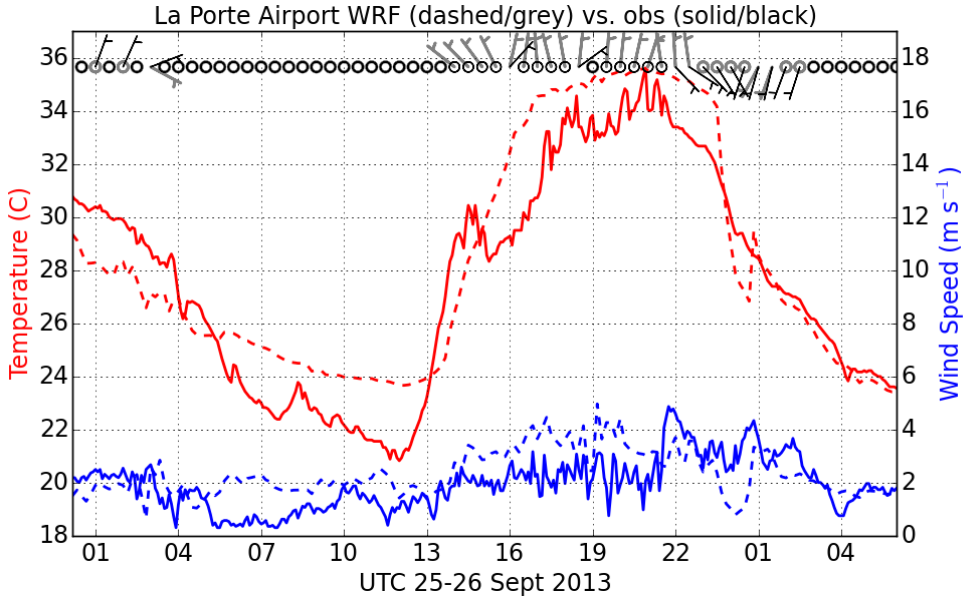

Figure 16. Same as Figure 15 but for the observation station at La Porte Airport.

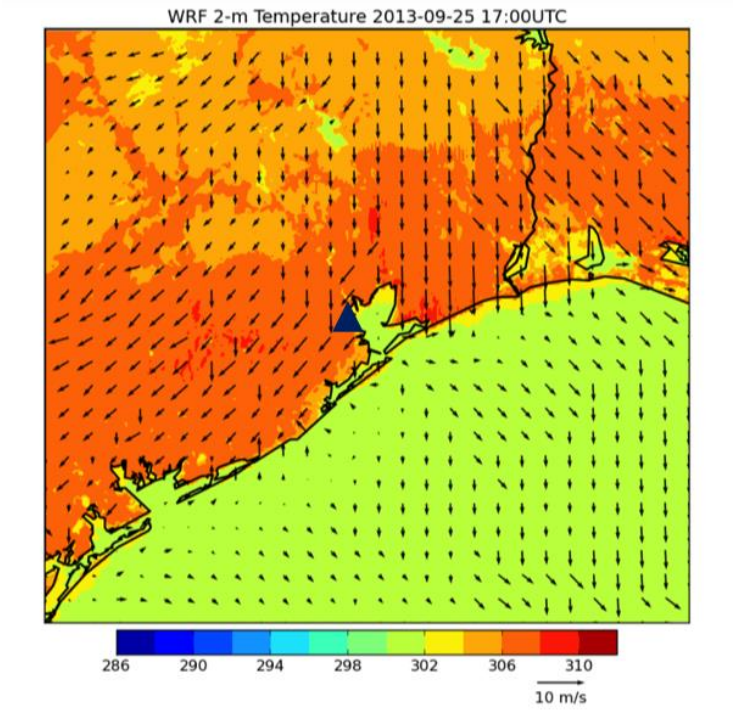

Figure 17. CFG3 WRF 1.33 km simulation of 2 m temperature (color shaded) and 10 m winds (black arrows) at 1700 UTC September 25, 2013. Blue triangle shows the location of the La Porte Airport station, shown in Figures 16 and 19.

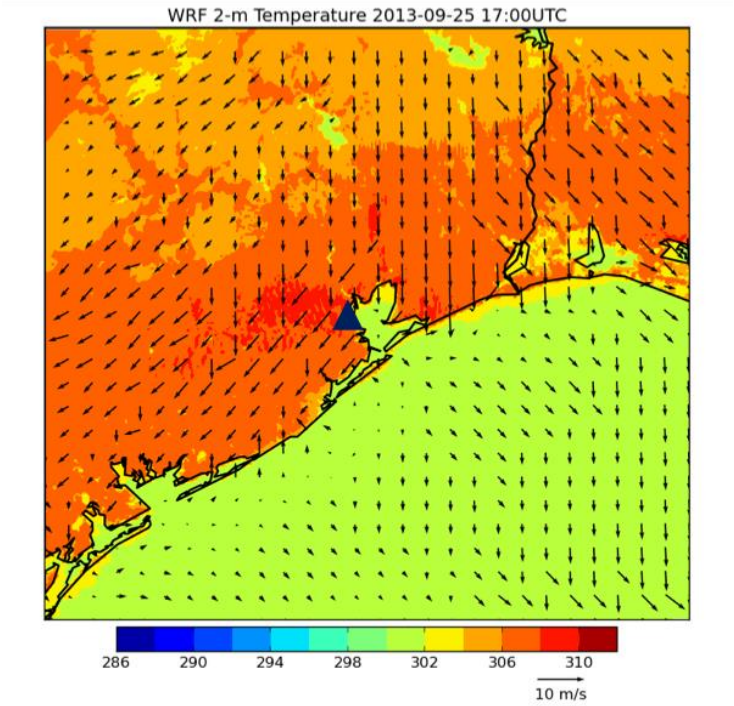

Figure 18. As in Figure 17, but for CFG4 at 1700 UTC September 25, 2013.

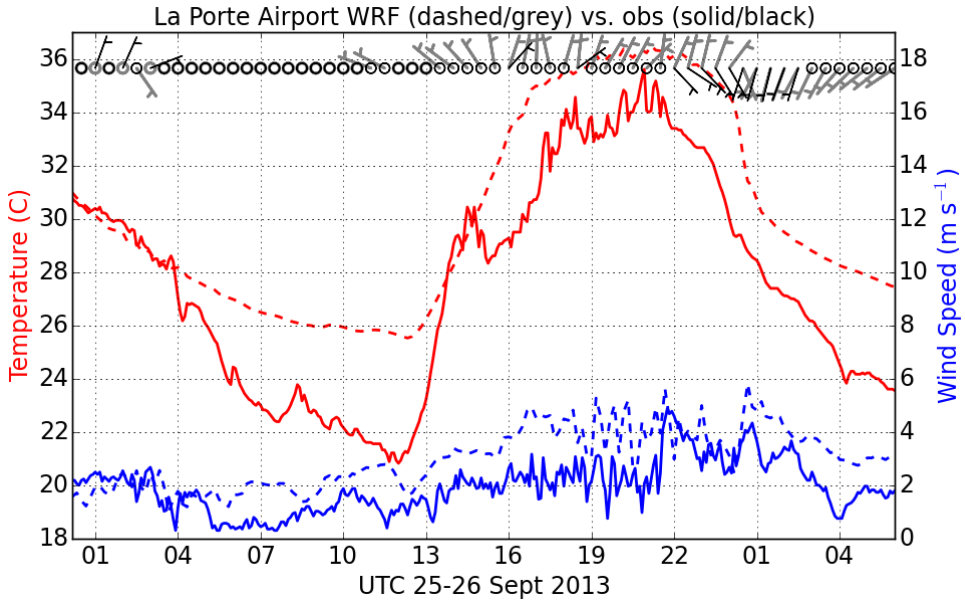

Figure 19. Observed (solid /black) and WRF CFG4 1.33 km (dashed /grey) time series of 10 m winds and 2 m temperature for the observation station at La Porte Airport starting 0000 UTC September 25, 2013.

3. WRF Production Runs

3.1. Data Coverage and Format

AER ran WRF for the month of September 2013 to provide high-resolution meteorological data covering the time period of the DISCOVER-AQ Houston field campaign of 2013. The WRF outputs were provided to TCEQ on the 36, 12, and 4 km domains described in Section 2.1 for the entire month at 30-minute intervals for the inner 4 km nest and at 1-hourly intervals for the outer 36 and 12 km nests. Data from an additional inner 1.33 km nest was provided at 10-minute intervals for the 10-day period starting September 17, 2013 that precedes and includes the 2-day high O_3 episodes of September 25th and 26th. In addition to the raw WRF output, model ready meteorological inputs for the CMAQ and CAMx models derived from the WRF data were provided on the 4 km nest for the entire month and on both the 4 km and 1.33 km nests for the 10-day period starting September 17th. All the outputs were provided on a portable hard drive that was shipped to the Project Manager Mark Estes of TCEQ. All the output files are in self-describing netCDF format.

The WRF output files have the following name convention

wrfout_d<domain>_yyyy-mm-dd:hh:mn:ss

where <domain> is the domain number with increasing number for each nest refinement level (e.g. d01 is the 36 km domain, d02 is the 12 km domain), yyyy, mm, dd,hh,mn, and ss are the year, month, day, time, hour, minute, and second of the simulation valid time.

The CMAQ meteorological input files were created by Version 4.1 of the Meteorology-Chemistry Interface Processor (MCIP). A set of the following files was created for each simulation day.

GRIDBDY2D_DISAQ13_D<domain>_yyyymmdd
GRIDCRO2D_DISAQ13_D<domain>_yyyymmdd
GRIDDOT2D_DISAQ13_D<domain>_yyyymmdd
METBDY3D_DISAQ13_D<domain>_yyyymmdd
METCRO2D_DISAQ13_D<domain>_yyyymmdd
METCRO3D_DISAQ13_D<domain>_yyyymmdd
METDOT3D_DISAQ13_D<domain>_yyyymmdd

The file names include <domain>, which is the domain number of the WRF data used, D03 for the 4km nest and D04 for the 1.33 km nest. The string yyyymmdd indicates the year, month, and day of the data. The data in each file includes times from 0000 UTC of the data indicated to 0600 UTC of the next calendar day as the first 6-hours of each WRF run are considered model spin-up time and are not used.

MCIP also creates a text file called GRIDDESC containing the grid definition parameters, and this was included in the files delivered to TCEQ. For more detailed information on the MCIP output files see the CMAQ documentation available at https://www.cmascenter.org.

The CAMx input files are created using the WRF-CAMx pre-processor that generates meteorological input files for CAMx v6+. The input files are in CAMx Fortran binary format. Five files for each simulation day are generated:

camx.2d.D<domain>.yyyymmdd.bin	2-d meteorological inputs
camx.3d.D<domain>.yyyymmdd.bin	3-d meteorological inputs
camx.cr.D<domain>.yyyymmdd.bin	cloud/precipitation field
camx.kv.D<domain>.yyyymmdd.bin	vertical diffusivity
camx.lu.D<domain>.bin	land use

Each file (except for land use) contains 25 hours of meteorology at hourly intervals, as recommended by the WRF-CAMx user guide. The 25 hours run from 0600 UTC to 0600 UTC the next day (or local midnight to local midnight). The input files are on the same projection and horizontal and vertical grid structure as the 1.33 km (for domain 4) or 4.00 km (for domain 3) WRF domain. The MYJ approach to diagnosing vertical diffusivities *[Janjić, 1994]* was chosen in agreement with the boundary layer parameterization selected in the WRF configuration (see Section 3.2). No sub-grid clouds were diagnosed due to the relatively high horizontal resolution of the inner domains. Additional information about CAMx is available in the online user guide: http://www.camx.com/files/camxusersguide_v6-20.pdf.

3.2. WRF Configuration

Based on the WRF evaluation described in Section 2, we have determined that the best configuration for simulating the fine-scale meteorological features influencing the high ozone episodes in the Houston area on September 25-26, 2013 is CFG3 as summarized in Table 1. The namelist inputs to run WRF with these options are included in Appendix A. CFG3 differs from the configurations currently being used at TCEQ to model June 2012 in two important ways.

1. CFG3 uses the Noah LSM instead of PXLSM so that the UCM can be implemented.
2. CFG3 includes an inner 1.33 km domain.

In addition, CFG3 includes the MYJ PBL instead of the Yonsei University (YSU, *Hong et al. [2006]*) PBL as our previous work has indicated that the MYJ performs slightly better simulating the PBL diurnal cycle in urban environments *[Nehrkorn et al., 2013]*. Similarly, since no precipitation events occurred over the Houston area during our testing period, we were not able to adequately test the impact of microphysics or cumulus parameterizations and so we chose schemes that we have used successfully in other urban environments. Here we used the *Lin et al. [2003]* microphysics scheme *[Chen and Sun, 2002; Lin et al., 2003]* and Grell-Freitas *[Grell and Freitas, 2014]* cumulus parameterization, which is a newer, improved version of the Grell-Devenyi scheme *[Grell and Devenyi, 2002]*.

CFG3 includes an SST diurnal cycle imposed on the high-resolution MUR SST with an amplitude that is based on the water depth. For the WRF sensitivity run described in Section 2.3 we used a constant maximum amplitude of 1.96 K in the shallow waters near the Texas coast and a constant slope factor to decrease the amplitude with ocean depth for both September 25th and 26th based on the average water temperatures measured from buoys on those days. However, for the production runs that cover the entire month of September we refined this procedure by enabling the maximum amplitude and slope factor to vary daily based on buoy observations from

the National Data Buoy Center (http://www.ndbc.noaa.gov/). The buoys used here are 42035 (water depth of 13m) and 42046 (water depth of 105m) (Figure 20). The procedure is illustrated in Figure 21 for September 24th to 27th. To conform to the usage of the gridded MUR SST data (daily values valid at 0000 UTC are linearly interpolated in time), we first compute 24-hour averages centered at 0000 UTC for each day (marked by the vertical lines in Figure 21) from the raw buoy data (plotted in black), and use these to define the long-term (greater than diurnal) trend by linear interpolation (f1, dashed green lines). We then estimate the diurnal amplitude by the range of the detrended data for each day. The results for the entire month of September 2013 are shown for both buoys in Figure 22. For each day, the derived diurnal amplitude values for the two buoys (and their water depth) are fitted to a straight line, resulting in a daily value of maximum diurnal amplitude (corresponding to depth=0), and a ramp depth values (beyond which the amplitude is zero). These values are shown in Figure 23. Finally, we assume a uniform time of day at which the diurnal cycle peaks (2100 UTC), and interpolate the gridded ocean depths shown in Figure 24 to the WRF grid to apply this model of the diurnal cycle (no diurnal SST cycle is imposed for grid points beyond the latitude-longitude window of Figure 24).

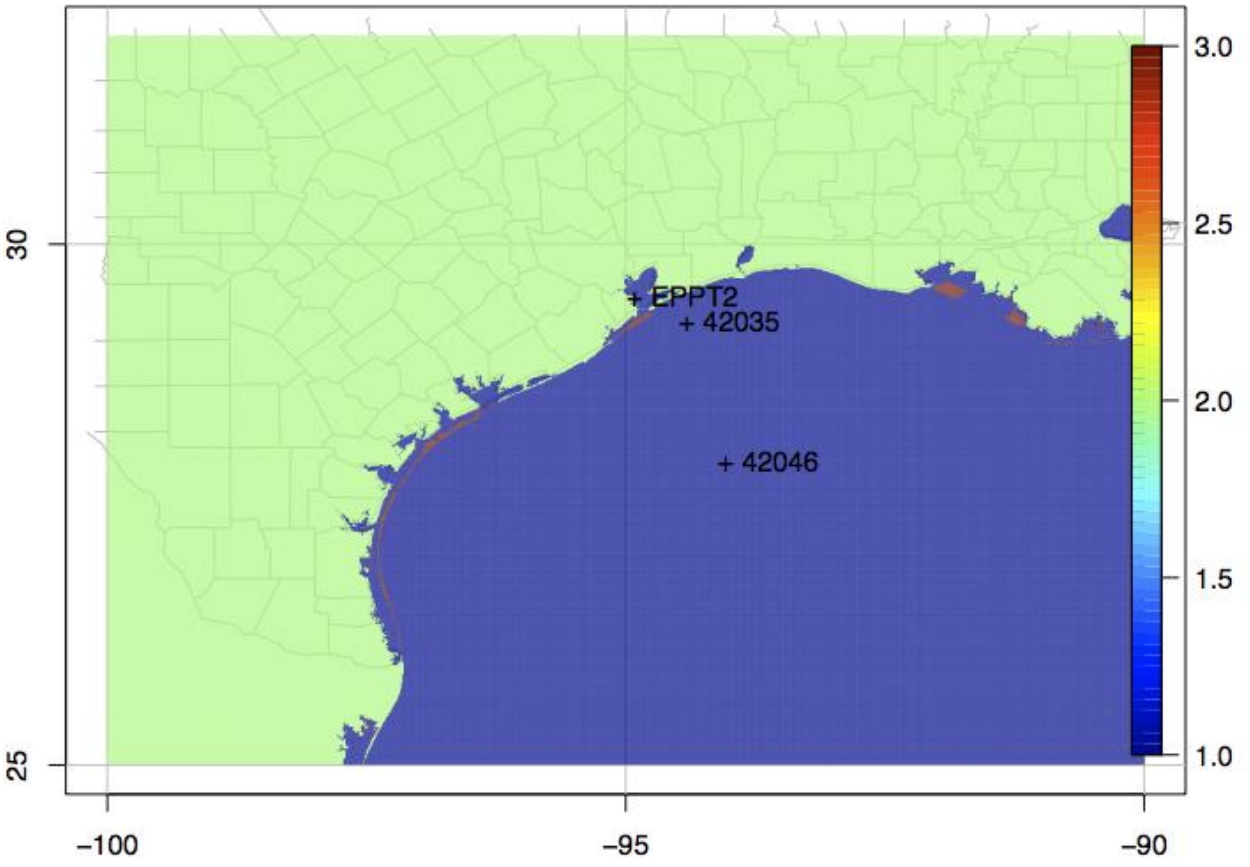

Figure 20. Location of the buoys 42035 and 42046. The background shows the land-water mask of the high-resolution gridded daily SST data. (Buoy EPPT2 is not used in our procedure, since it is not representative of ocean conditions.)

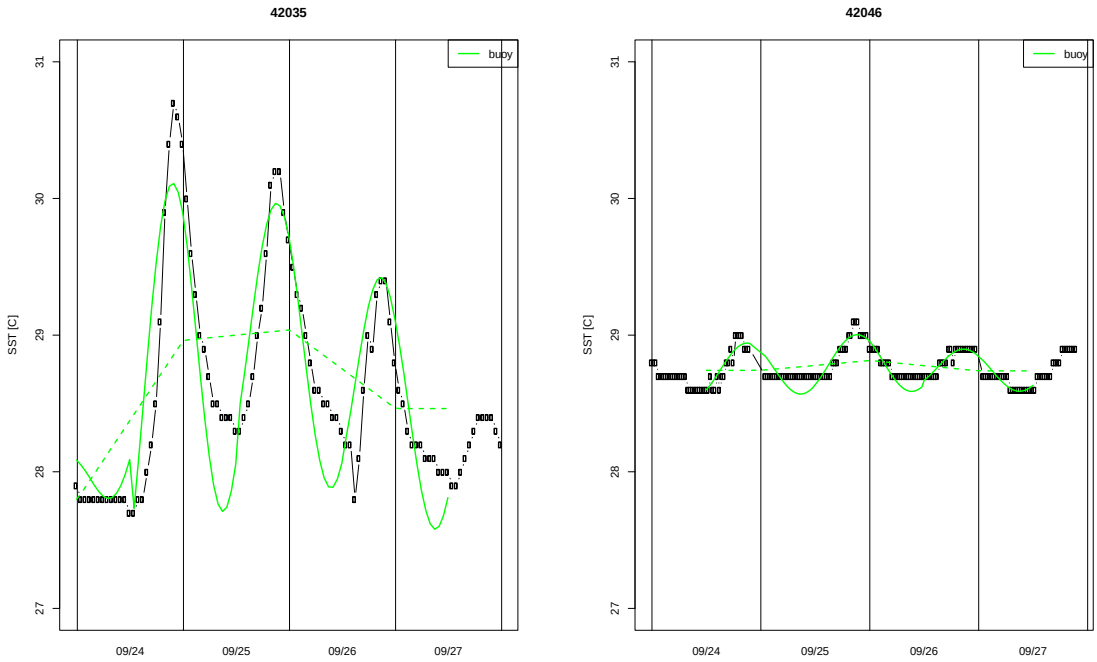

Figure 21. Illustration of diurnal range computation for buoys 42035 (left) and 42046 (right). See text for details.

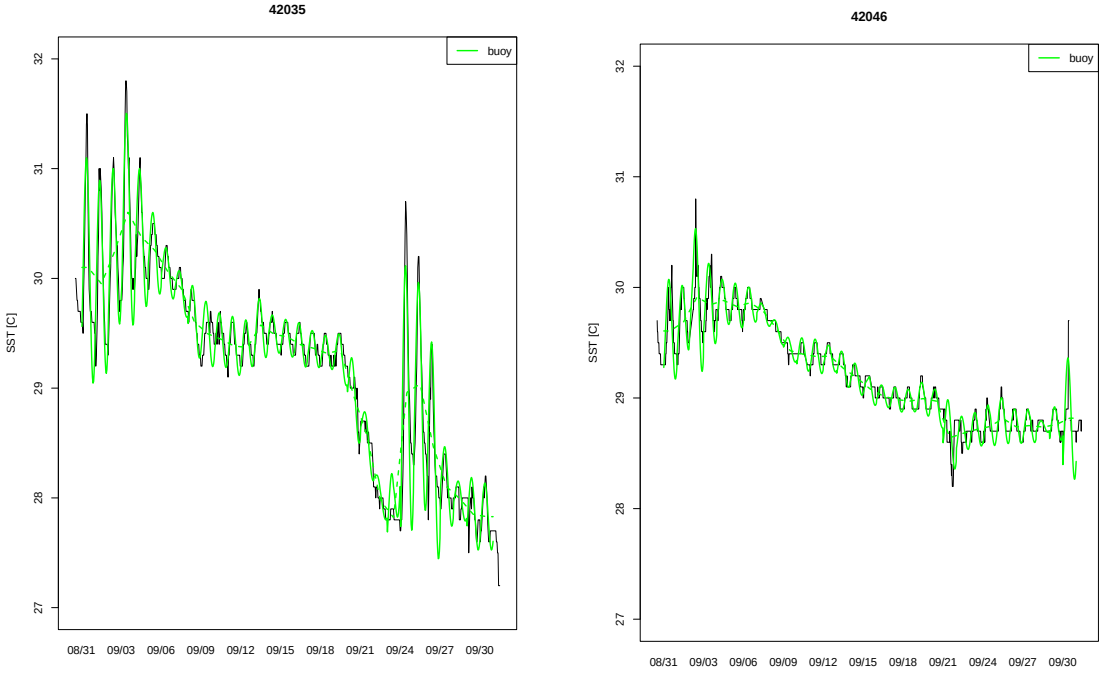

Figure 22. As in Figure 21, but for the entire month of September 2013. For clarity of presentation, vertical lines at midnight have been omitted, and observed buoy values are plotted as black lines without symbols (they are partially obscured by the over-plotted modeled values in green).

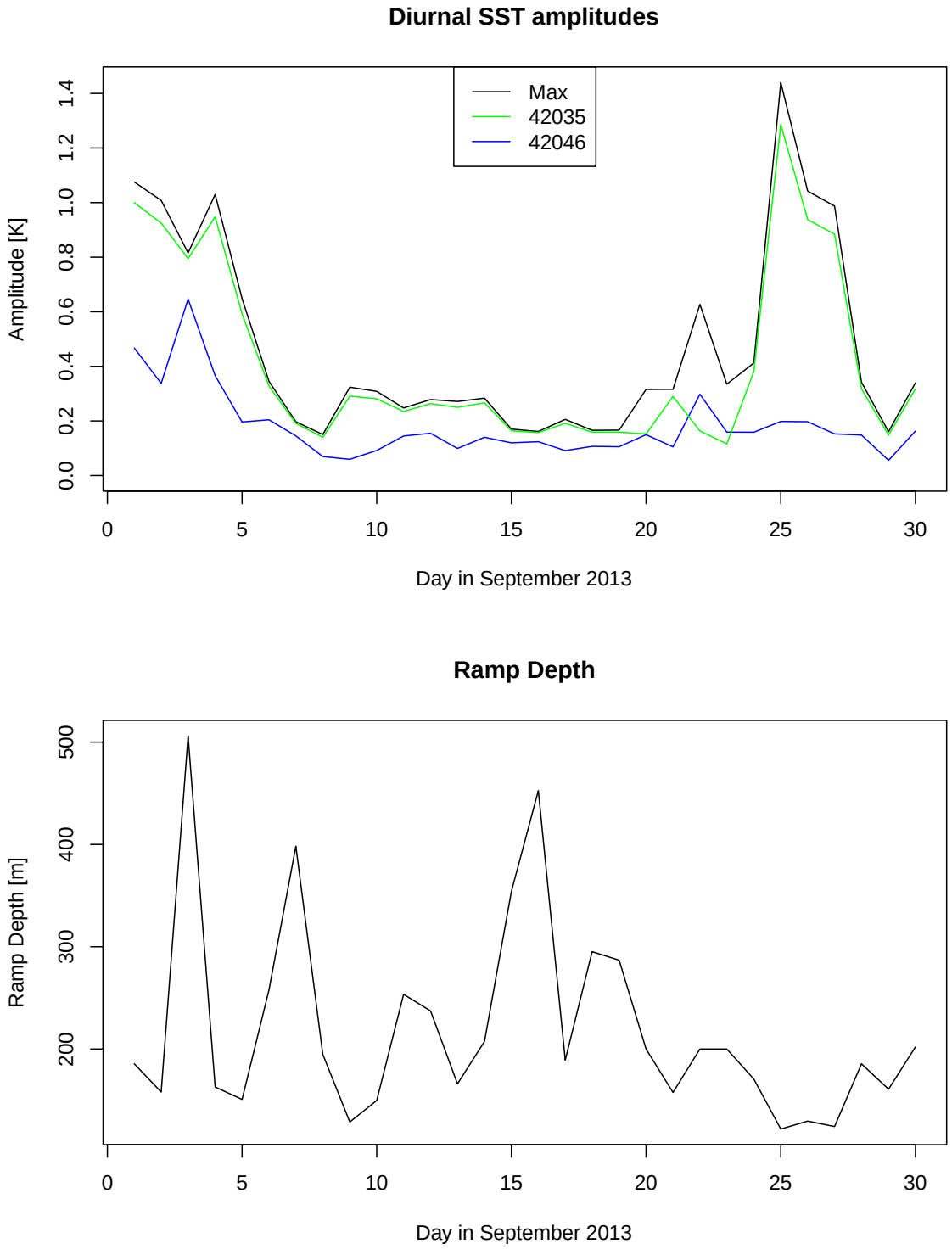

Figure 23. Derived values of daily amplitude (top) and ramp depth (bottom). Amplitude values are those for the individual buoys, and the corresponding derived maximum amplitude value.

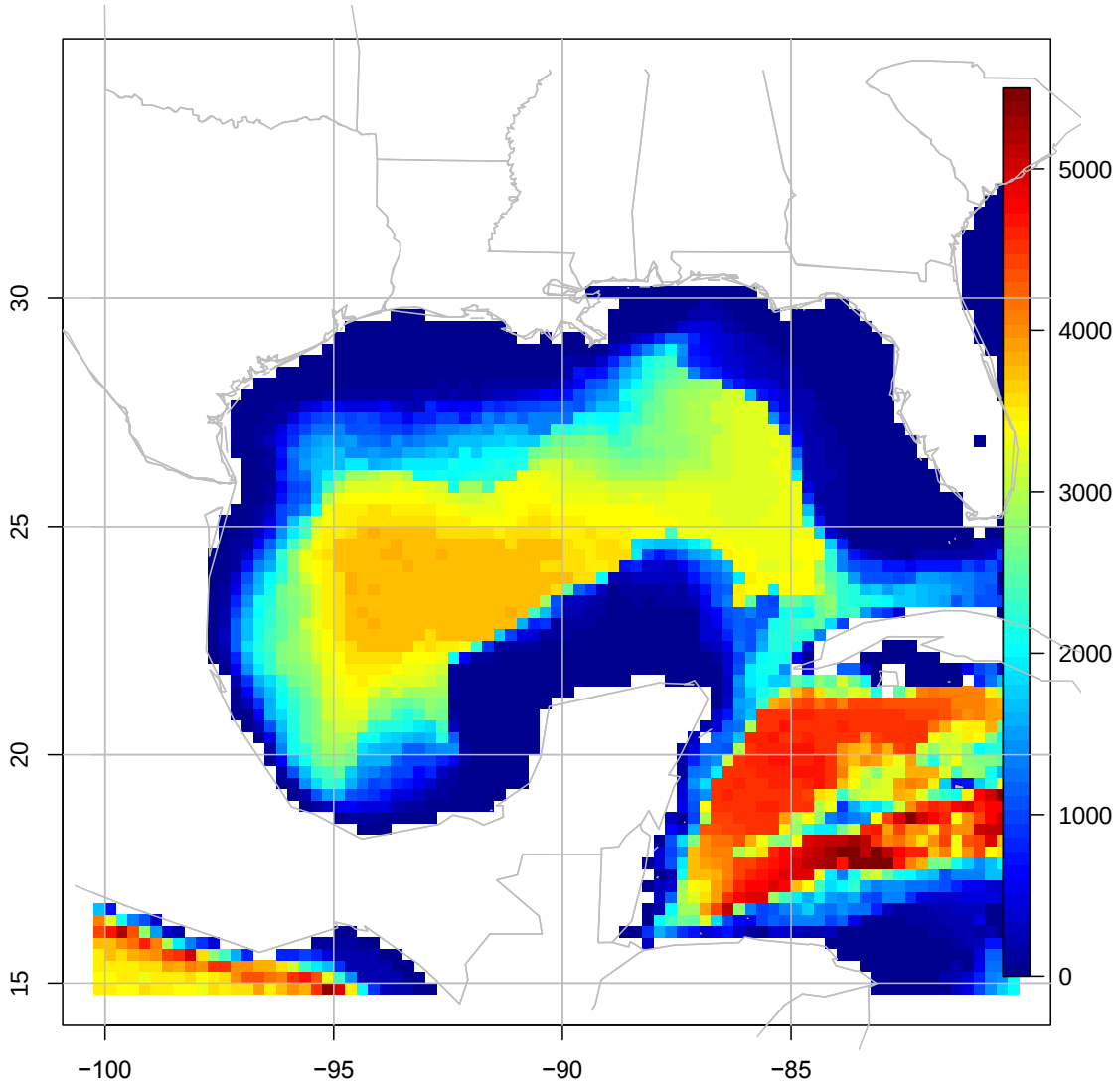

Figure 24. Gridded ocean depth values (m) derived from the World Ocean Atlas data used for interpolation to the WRF grids.

3.3. Validation of WRF Production Runs

3.3.1 Statistical Evaluation for September 2013

As described in the Quality Assurance Application Plan (QAPP), WRF runs were objectively evaluated by computing performance statistics using the WRF-Model Evaluation Tools v5.0 (WRF-MET, available at http://www.dtcenter.org/met/users/) software package. WRF-MET is developed and maintained at the National Center for Atmospheric Research (NCAR) and has been designated the official verification tool for validating simulations generated by NCAR's Advanced Research WRF (referred to simply as WRF in this document). WRF-MET generates a comprehensive set of model performance statistics by comparing model fields against a variety of observational datasets. For the purposes of point verification using individual observations, WRF-MET has been designed to ingest both prepBUFR-format files from NCEP (an archive is also maintained by NCAR at http://rda.ucar.edu/datasets/ds337.0/) and Meteorological Assimilation Data Ingest System (MADIS) files from NOAA (http://MADIS.NOAA.gov/).

Results shown here are based on the PrepBUFR files, which contain conventional data (e.g., surface reports, rawinsondes, and some satellite-based observations) in BUFR format (a WMO standard) that have been prepared for ingest into NCEP's operational data assimilation. As such, they are highly quality controlled and available in a timely manner for both North America and globally. All the plots and tables generated by WRF-MET for domains d03 (4 km) and d04 (1.33 km) will be made available on AER's public ftp server. Descriptions of the contents of each of these files and how to obtain them from our ftp server is provided in Appendix B. Only a brief summary of these results are provided here.

Summary statistics of simulated vs. observed near-surface variables from the innermost domain (d4) are shown in Figure 25. Shown are the hourly domain-wide bias and rms error in 2 m temperature and 10 m wind speed, based on all surface-observing stations over land – the corresponding statistics for ship and buoy observations (not shown) have comparable magnitudes. There is considerable variation of each of these over the 10-day period, with the majority of times having a temperature bias/rmse below 1 to 2 K, wind speed bias between 0.5 and 1.5 m s^{-1}, and wind speed rmse below 2 m s^{-1}. The geographic variation of the WRF performance can be determined from Figure 26, which shows the bias and rms error (computed separately for each station, over all observation times). This plot reveals some general trends (e.g., a slight overprediction of wind speed over land, and an underprediction over water), as well as identifies outlier stations (e.g., examination of station KDRK revealed a large number of bad wind observations with a reported zero wind speed, which explains the large error values shown in Figure 26). Wind summary statistics in the form of wind rose plots for observed and simulated winds for individual stations can provide additional insights: the results for Houston (KHOU, Figure 27) show that while the general pattern is well reproduced, there are some discrepancies in the details, such as the tendency to overpredict the wind speed for southeast winds, but underpredict it for northeast winds. For an even more detailed evaluation, we also generated time series plots (of modeled and observed values) at individual stations (an example is shown in Figure 28 for Houston). This can identify additional aspects of the simulation: for example, the slightly negative bias of 2 m temperature at KHOU is largely due to nighttime temperatures that are too cold on most days, which is only partially offset by overpredicted daytime maxima on some days.

We also used WRF-MET to generate performance statistics for the entire month of September 2013 for the 4 km domain (d03). Summary statistics for all the observation stations in the Houston–Galveston area, shown in Figure 29, are presented in Tables 2 and 3. The bias (Table 2) and rms error (Table 3) are shown for each station and the last three lines show the land, water and overall bias and rmse for all the stations in the Houston–Galveston area. The overall bias for wind speed was 0.77 m s $^{-1}$, and was greater over land (0.99 m s^{-1}) than over water (0.16 m s^{-1}). For temperature the overall bias was 0.34 K and was of a different sign over water (-0.38 K) than over land (0.65 K). The overall rmse for wind speed was 1.92 m s^{-1} which was only slightly greater over land (1.95 m s $^{-1}$) compared to over water (1.85 m s^{-1}). For temperature the overall rmse was 1.68 K and was greater over land (1.85 K) than over water (1.29 K). These statistics are generally consistent with those of other meterological modeling studies of the Houston – Galveston area *[Byun et al., 2011; Chen et al., 2011b; Lee et al., 2012]*.

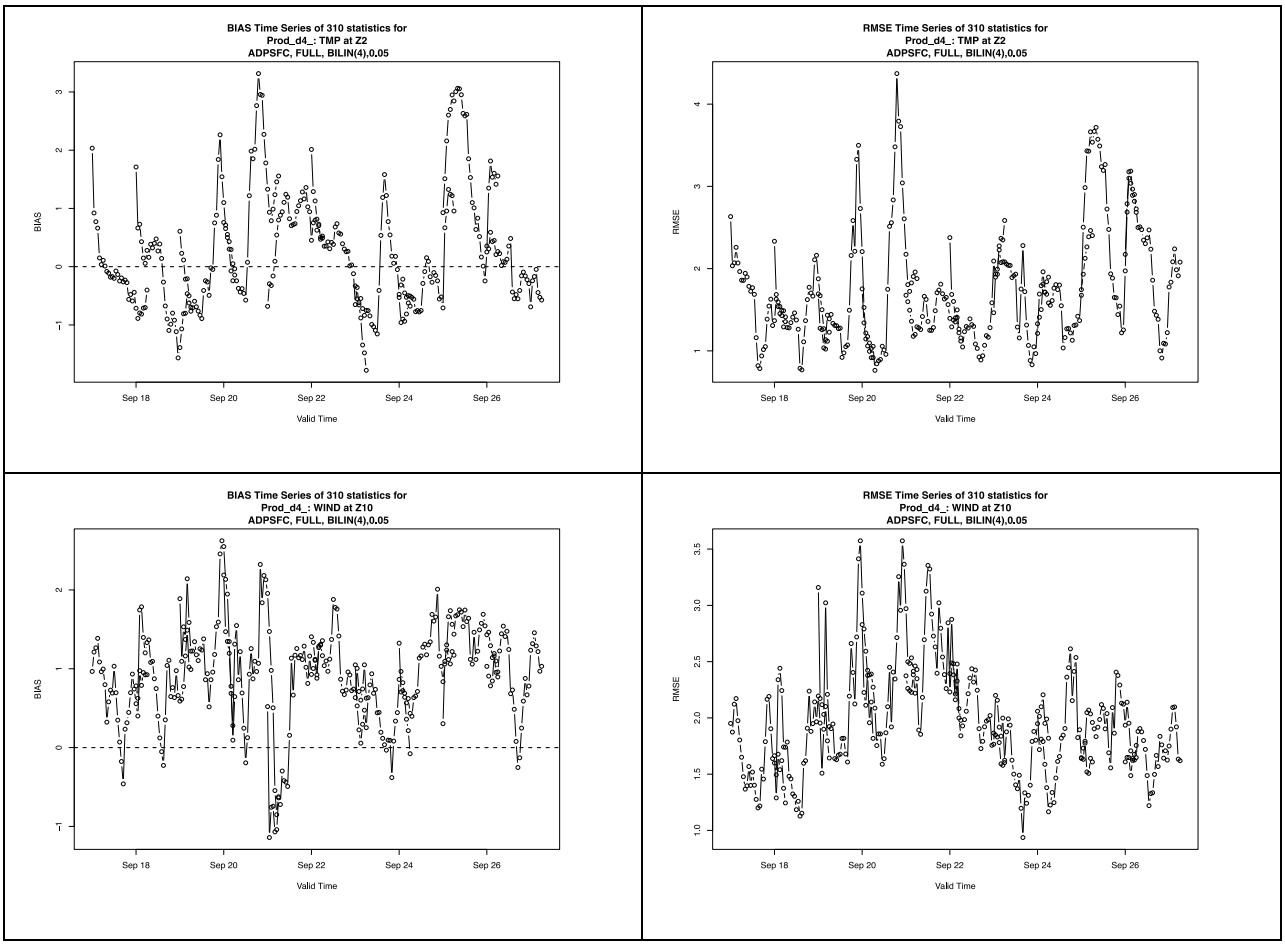

Figure 25. Domain-wide summary statistics (bias, left, and rmse, right) for 2 m temperature (K, top) and 10 m winds (m s^{-1} , bottom). Results are shown for all 30 hours of each forecast, resulting in overlapping lines for hours 0000-0600 UTC of each day.

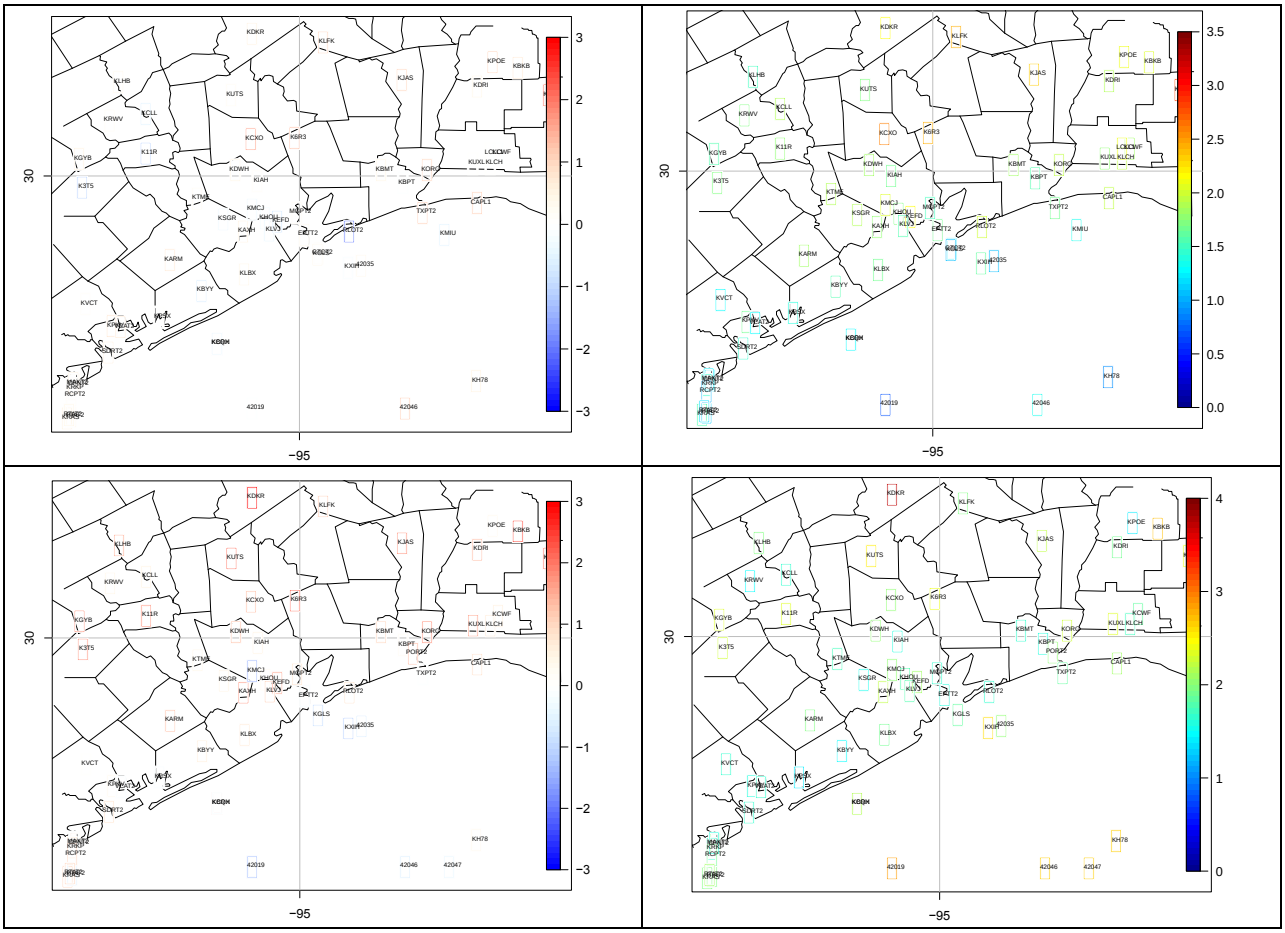

Figure 26. Time-averaged statistics (bias, left, and rmse, right) for the observing stations in domain d4, for 2 m temperature (K, top) and 10 m wind speed (m s^{-1}, bottom). Results shown here are for land and buoy observations, for forecast hours 7-30 only.

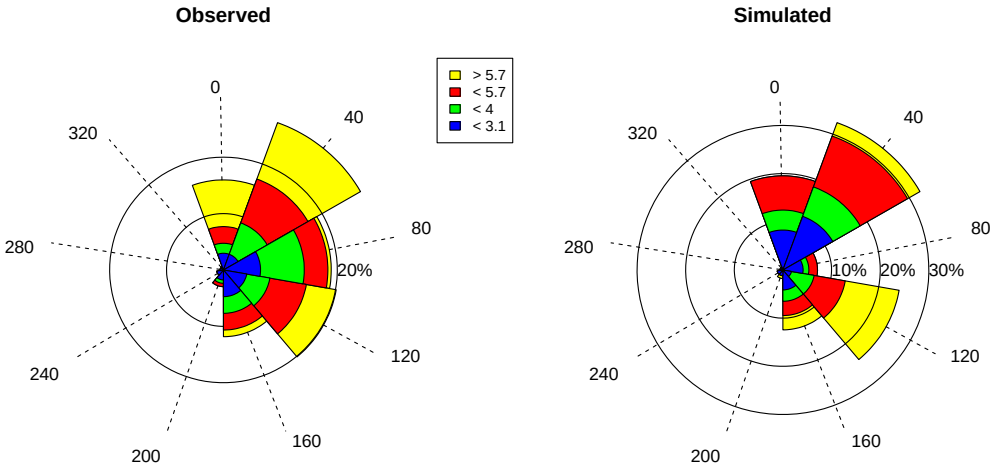

Figure 27. Wind rose plot for observed (left) and simulated (right) 10 m winds at KHOU. Results are only shown for the 169 observations (out of a total of 238) with reported wind speeds in excess of 3 knots (1.54 m s^{-1}). Direction bins were chosen to conform to the reporting practice of reporting wind direction to the nearest 10 degrees. Length of pie segments correspond to the frequency of occurrence, color coding is by wind speed bin (in m s^{-1}).

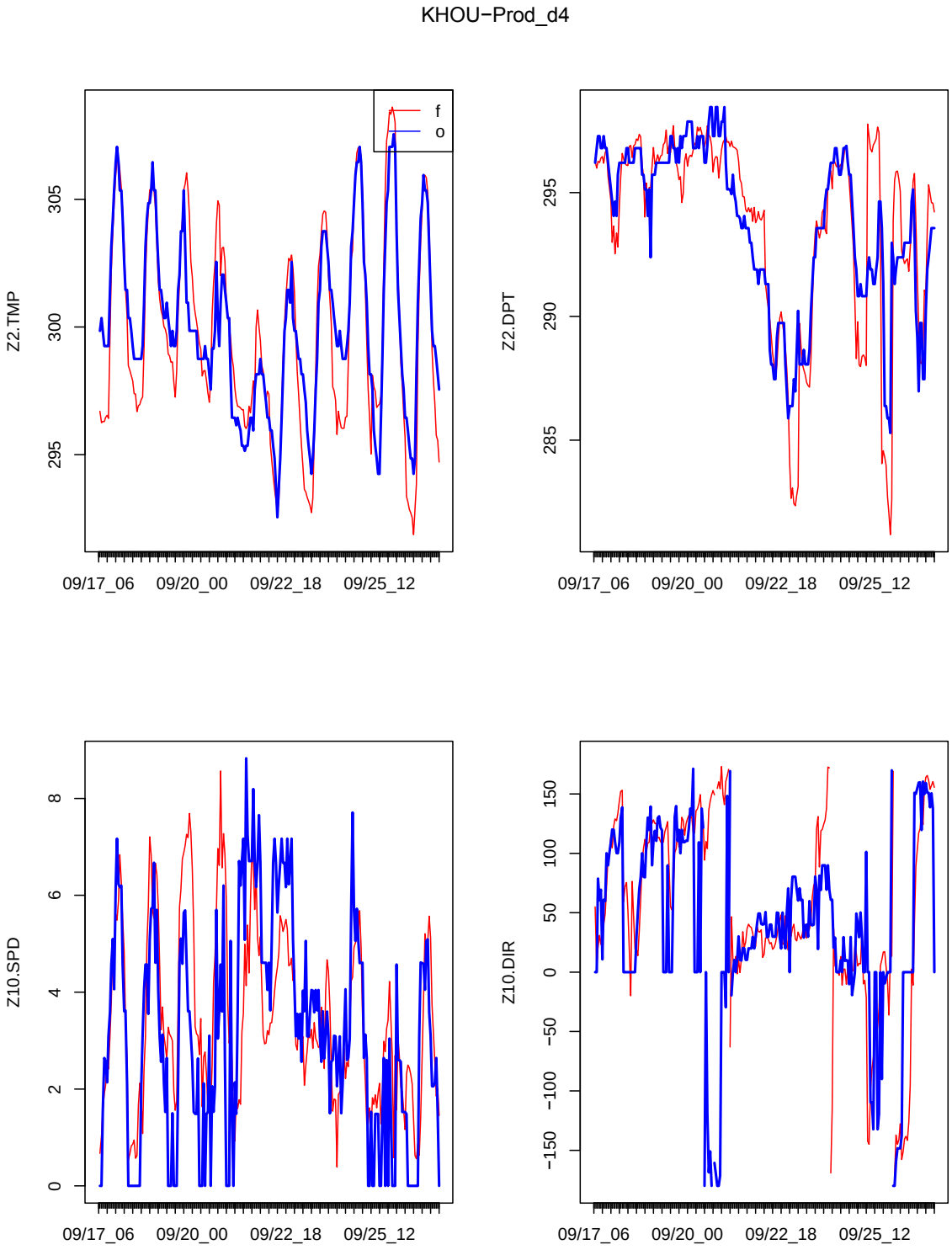

Figure 28. Example time series plot of forecast (f, red) and observed (o, blue) values for 2m temperature (top left, K), dew point (top right, K), wind speed (bottom left, m s^{-1}), and wind direction (bottom right, degrees).

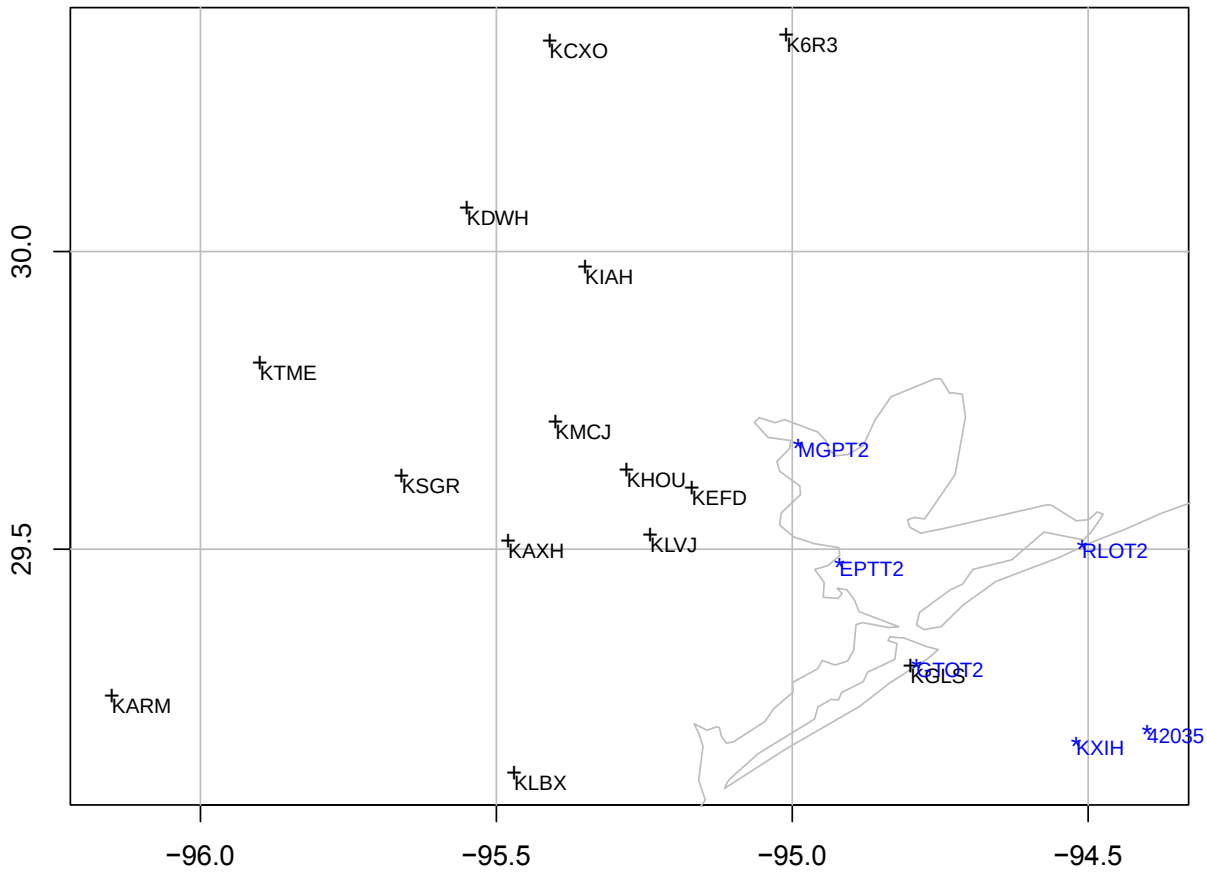

Figure 29. Map of the stations used to generated the WRF-MET performance statistics presented in Tables 2 and 3 for the the Houston-Galveston area using the 4 km output of the WRF production runs for September 2013.

Table 2. WRF-MET bias statistics for the Houston-Galveston area for the September 2013 4 km WRF production runs.

Station	Temp (K)	Dewpoint (K)	Relative Humidity (%)	Wind Speed (m/s)	Wind Direction (degrees)
42035	0.12	2.54	10.39	-0.12	7.11
EPTT2	0.03	NA	NA	-0.36	-7.90
GTOT2	-0.18	NA	NA	NA	NA
K6R3	1.20	-0.09	-6.43	1.94	28.67
KARM	0.63	0.79	-0.26	1.14	8.42
KAXH	0.83	0.35	-3.09	1.64	20.10
KCXO	1.39	0.75	-4.38	1.38	21.89
KDWH	1.37	-0.67	-9.63	1.40	6.30
KEFD	**-0.20**	**-0.73**	**-2.96**	**1.59**	**15.19**
KGLS	**0.13**	**-0.21**	**-2.89**	**0.28**	**9.93**
KHOU	0.59	0.02	-2.41	0.83	9.79
KIAH	**1.30**	**-0.69**	**-8.91**	**0.59**	**1.18**
KLBX	0.38	0.36	-1.17	1.04	5.31
KLVJ	0.23	0.38	0.36	1.35	6.74
KMCJ	0.61	0.19	-1.87	-0.38	-0.16
KSGR	0.53	-0.20	-3.38	0.70	0.06
KTME	0.16	0.49	-0.41	0.33	17.90
KXIH	-0.45	2.60	12.87	-0.41	-10.61
MGPT2	0.00	NA	NA	0.76	-7.65
RLOT2	-1.83	NA	NA	0.92	17.70
land	**0.65**	**0.05**	**-3.39**	**0.99**	**10.81**
water	**-0.38**	**2.57**	**11.63**	**0.16**	**-0.27**
all	**0.34**	**0.37**	**-1.51**	**0.77**	**7.89**

Table 3. WRF-MET rmse statistics for the Houston-Galveston area for the September 2013 4 km WRF production runs.

Station	Temp (K)	Dewpoint (K)	Relative Humidity (%)	Wind Speed (m/s)	Wind Direction (degrees)
42035	0.75	2.74	11.87	1.78	39.83
EPTT2	1.23	NA	NA	1.67	42.25
GTOT2	1.05	NA	NA	NA	NA
K6R3	2.31	2	12.75	2.38	98.72
KARM	1.76	1.72	8.60	1.82	67.26
KAXH	1.70	1.68	9.06	2.19	83.12
KCXO	2.55	2.01	13.13	2.10	87.59
KDWH	2.49	1.95	14.60	2.07	77.34
KEFD	**1.71**	**2.23**	**11.33**	**2.14**	**74.29**
KGLS	**0.95**	**1.10**	**7.49**	**1.69**	**44.48**
KHOU	1.68	1.64	9.29	1.95	66.90
KIAH	**2.27**	**1.89**	**13.44**	**1.81**	**68.93**
KLBX	1.59	1.44	8.07	1.87	70.64
KLVJ	1.47	1.71	8.58	2.02	75.70
KMCJ	2.01	1.63	10.58	1.76	52.10
KSGR	1.73	1.71	9.58	1.72	66.01
KTME	1.65	1.85	10.07	1.73	75.06
KXIH	1.44	2.81	14.83	2.16	41.35
MGPT2	1.18	NA	NA	1.76	42.57
RLOT2	2.11	NA	NA	1.87	48.15
land	**1.85**	**1.75**	**10.47**	**1.95**	**72.01**
water	**1.29**	**2.77**	**13.35**	**1.85**	**42.83**
all	**1.68**	**1.88**	**10.83**	**1.92**	**64.33**

3.3.2 Evaluation of September 25 -26th High O_3 Episodes with 1.33 km Data

The WRF runs were also evaluated on the high-resolution 1.33 km domain to determine how well the winds and temperatures were simulated at individual sites that were affected by the sea breeze and Galveston Bay breeze and/or high O_3 levels during the high O_3 episodes of September 25th and 26th. In addition, the simulation of solar radiation, which plays a major role in O_3 production, was evaluated at these sites.

The highest O_3 levels on September 25th were recorded on the western shore of Galveston Bay at the Texas City, Seabrook, and La Porte monitoring stations (see Figure 1). As discussed in Section 2, the WRF configuration selected for the production runs (CFG3) generally did a reasonable job of simulating the winds along the Texas coastline (including Galveston Bay) as well as the development of the sea and bay breezes on the afternoon of September 25th. On Galveston Island, WRF simulated the sea breeze beginning just after 1900 UTC, a little more that 2 hours after it was observed (see Figure 15). After the sea breeze onset the WRF winds

were initially a little more southerly than observed and about 2 m s^{-1} too fast but became more consistent with the observations toward the evening. The simulated temperatures at Galveston generally tracked the observed values except during the 2 hour period between the observed and simulated start of the sea breeze. At Seabrook and Texas City, WRF captures the Bay breeze wind shift from northerly to easterly at 1700 UTC as well as the shift to southerly at 2200 UTC as the Gulf of Mexico breeze became the dominant forcing (Figures 30 and 31). However, at La Porte the wind shift from northerly to southeasterly observed at 2100 UTC was delayed in the simulation until ~0030 UTC September 26th and was too weak with wind speeds remaining mostly less than 2.5 m s^{-1} (see Figure 16). In general simulated temperatures tracked the observed at Seabrook and Texas City but were about 2 K too warm at their peak at Texas City. At La Porte, the temperature errors were consistent with the delayed onset of the bay breeze.

Another important factor influencing O_3 mixing ratios is the amount of solar radiation penetrating to lower atmospheric levels and this is critically dependent on cloud cover. The GOES infrared satellite loops for September 25th (not shown) suggested mostly clear skies throughout the day over much of eatern Texas. METAR observations from around the Houston area also indicated clear skies and visibilities at or near 10 statute miles for much of the day with only "FEW" clouds (1/8 – 2/8 cloud cover) reported toward the evening hours. The WRF solar radiation time series showed no significant impacts from simulated cloud cover and generally were within 10 % of the observations at the three stations in the vicinity of Galveston Bay (Figure 32). The one notable exception was on Galveston Island where the solar radiation was dramatically lower between 1700 and 1800 UTC. This reduced solar radiation occurred at the same time on September 26th but the METAR observations from Galveston (station identification KGLS) indicated clear skies and good visibilities during those periods on both days. Therefore we suspect that either the instrument was malfunctioning or that some short-lived localized obscurent was interfering with the instrument measurments each day at those times.

On September 26th, the highest O_3 mixing ratios were observed northwest of Houston (see Figure 1). On this day an anticyclone that had been located over central Texas on September 25th moved offshore into the Gulf of Mexico producing a general south-southwest flow over much of coastal Texas (Figure 33). By late morning the winds had shifted into the southeast at Galveston Island and Seabrook, and this wind shift was captured by the WRF simulations (Figure 34 and 35). Similar to the case of September 25th, the simulated wind shift at Galveston Island lagged the observed by ~3 hours, but both the observed and modeled wind speed were light (2 to 3 m s^{-1}) during this period and only increased after ~1700 UTC when both observed and simulated winds were of the same direction. At Conroe Airport, where high O_3 levels were observed on this day, the winds remained light until 1500 UTC when they increased to ~4 m s^{-1} out of the south and then shifted to the southeast (Figure 36). The simulated winds remained light later until ~1800 UTC but then generally matched the observed in speed and direction during the remainder of the afternoon. The temperature simulation at Conroe Airport tracked the observed within 1 K throughout the entire day.

Similar to September 25th GOES satellite loops and METAR observation for September 26th indicated mostly clear skies over the Houston-Galveston area. WRF also simulated clear skies and the solar radiation matched the observed within 10 % at Seabrook Friendship Park on the western shore of Galveston Bay and at Conroe Airport northwest or Houston (Figure 37).

Overall the WRF simulations captured the wind features associated with the developing sea and Galveston bay breezes on high O_3 episode days. The most noteable issue with the simulations were that the initation of the sea breeze and associated wind shifts were delayed by 2 to 3 hours at some stations. However, these runs are an improvement over the previous WRF runs reported by *Loughner et al. [2014]* in which the timing and strength of the simulated sea breeze featured greater errors.

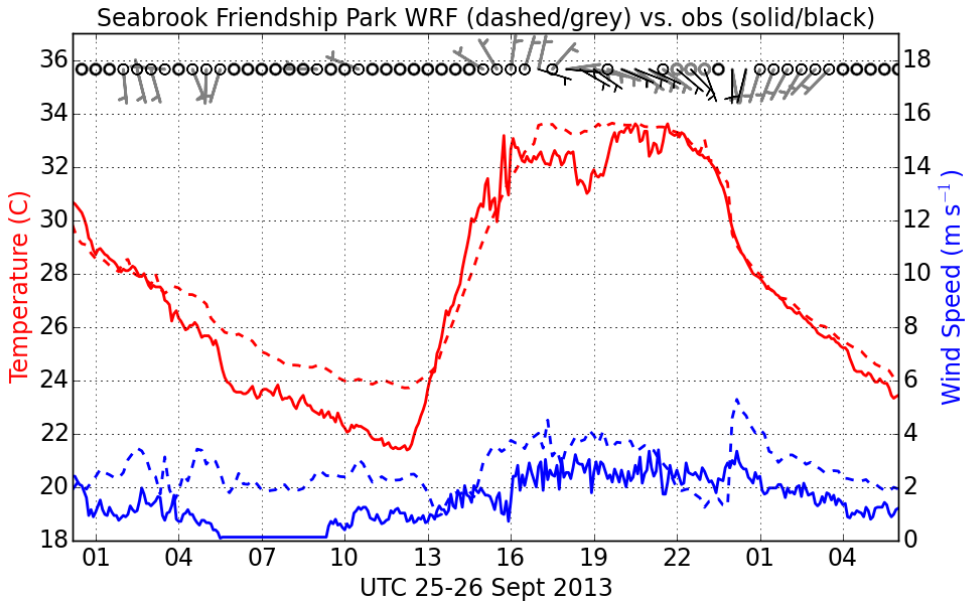

Figure 30. Observed (solid/black) and WRF Production run (CFG3) 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Seabrook Friendship Park starting 0000 UTC September 25, 2013.

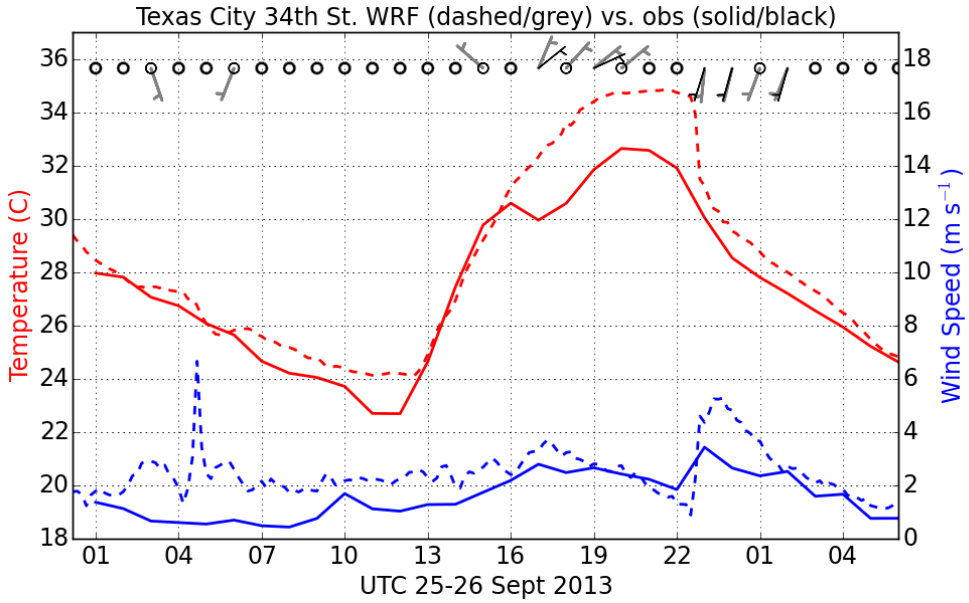

Figure 31. Observed (solid/black) and WRF Production run (CFG3) 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Texas City 34th Street starting 0000 UTC September 25, 2013.

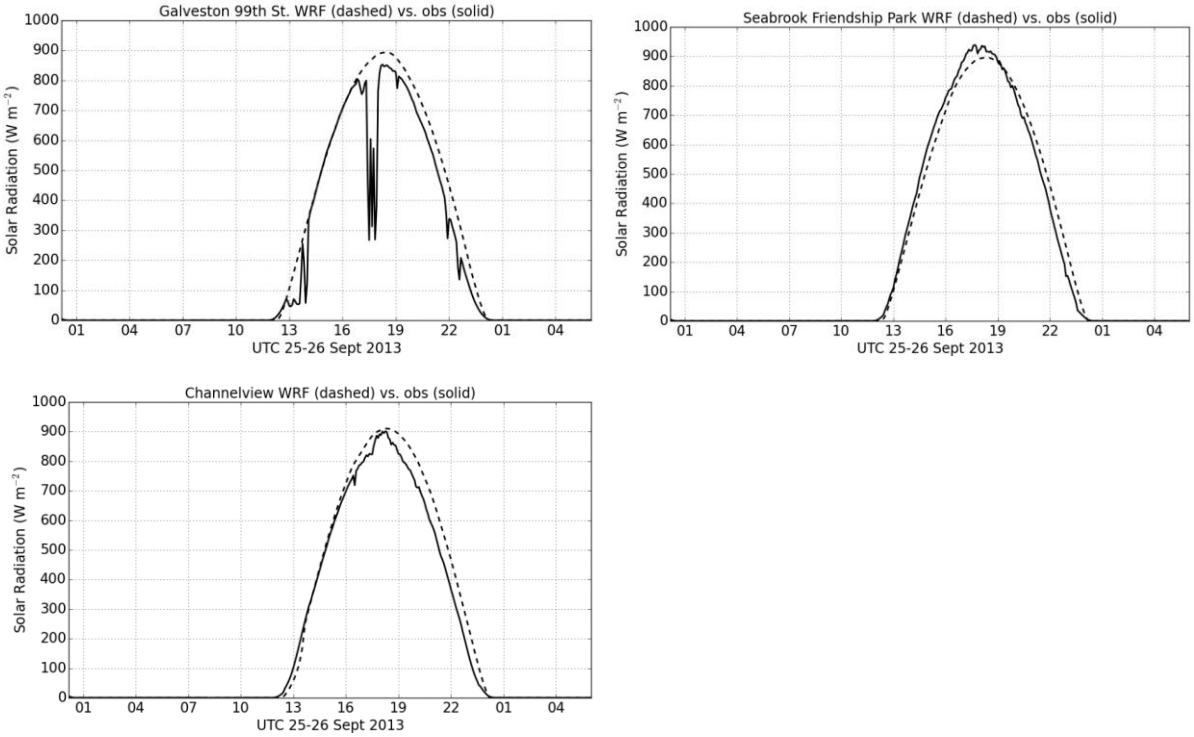

Figure 32. Observed (solid) and WRF Production run (CFG3) 1.33 km (dashed) solar radiation (w m^{-2}) for Galveston (top left), Seabrook (top right) and Channelview (bottom right) staring 0000 UTC September 25, 2013.

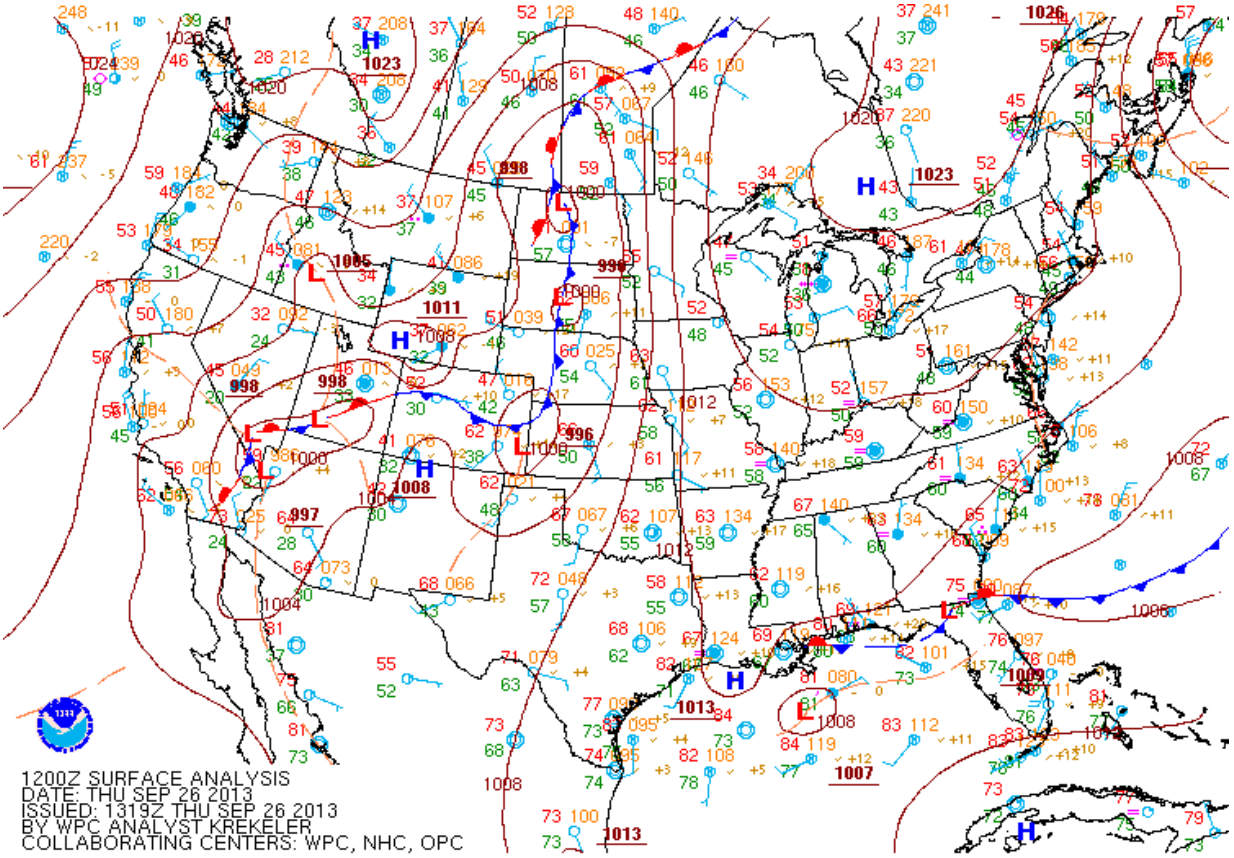

Figure 33. NOAA Weather Prediction Center surface analysis showing sea level pressure (mb, solid lines) analysed fronts and observations valid 1200 UTC September 26, 2013. Plot available for download from www.wpc.ncep.noaa.gov.

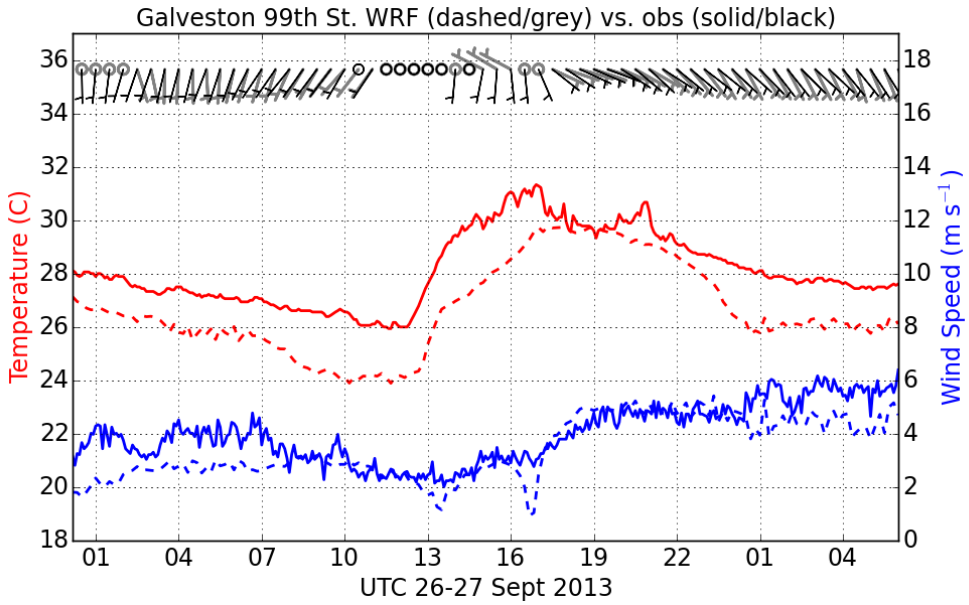

Figure 34. Observed (solid/black) and WRF Production run (CFG3) 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Galveston 99th Street starting 0000 UTC September 26, 2013.

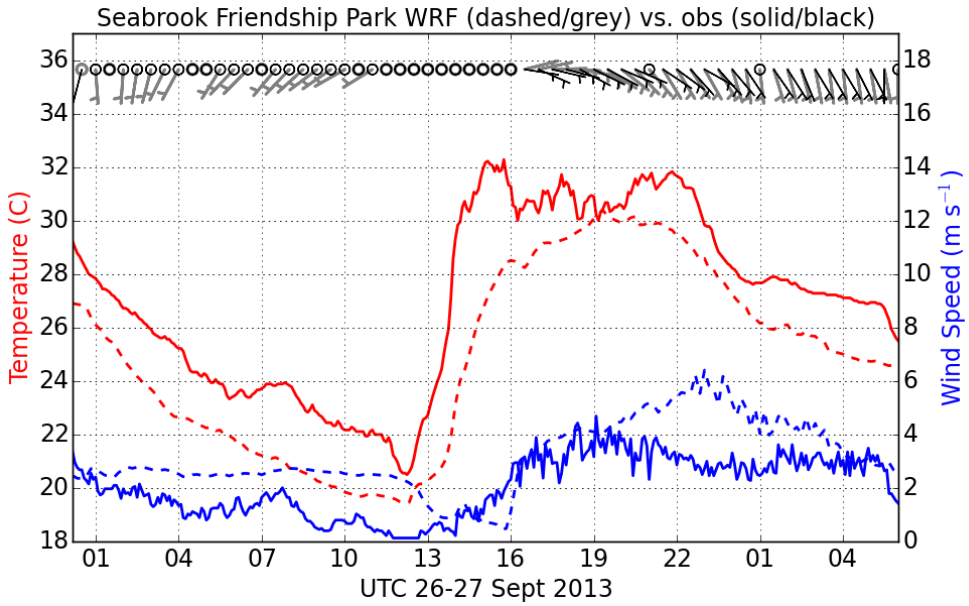

Figure 35. Observed (solid/black) and WRF Production run (CFG3) 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Seabrook Friendship Park starting 0000 UTC September 26, 2013.

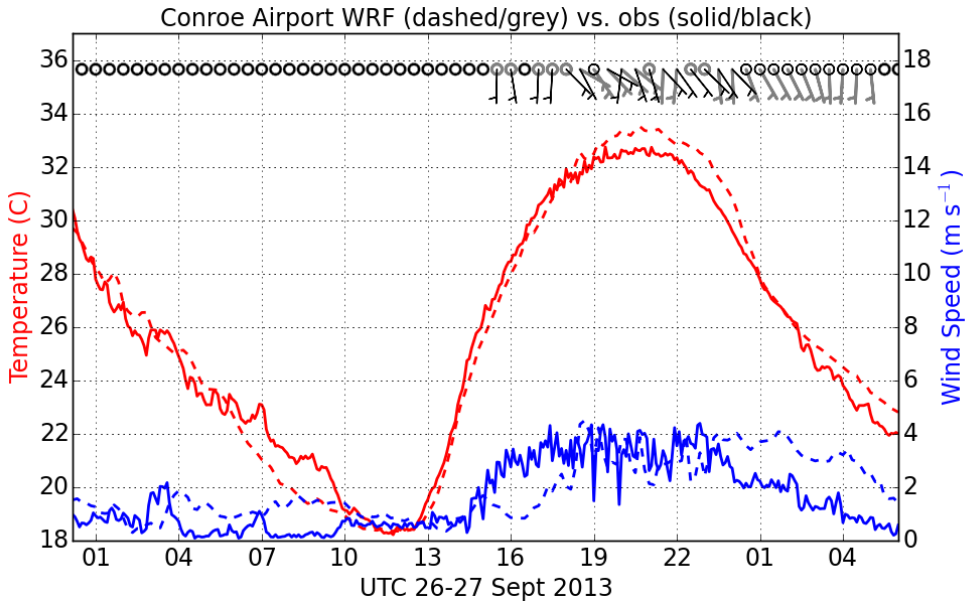

Figure 36. Observed (solid/black) and WRF Production run (CFG3) 1.33 km (dashed/grey) time series of 10 m winds and 2 m temperature for the observation station at Conroe Airport starting 0000 UTC September 26, 2013.

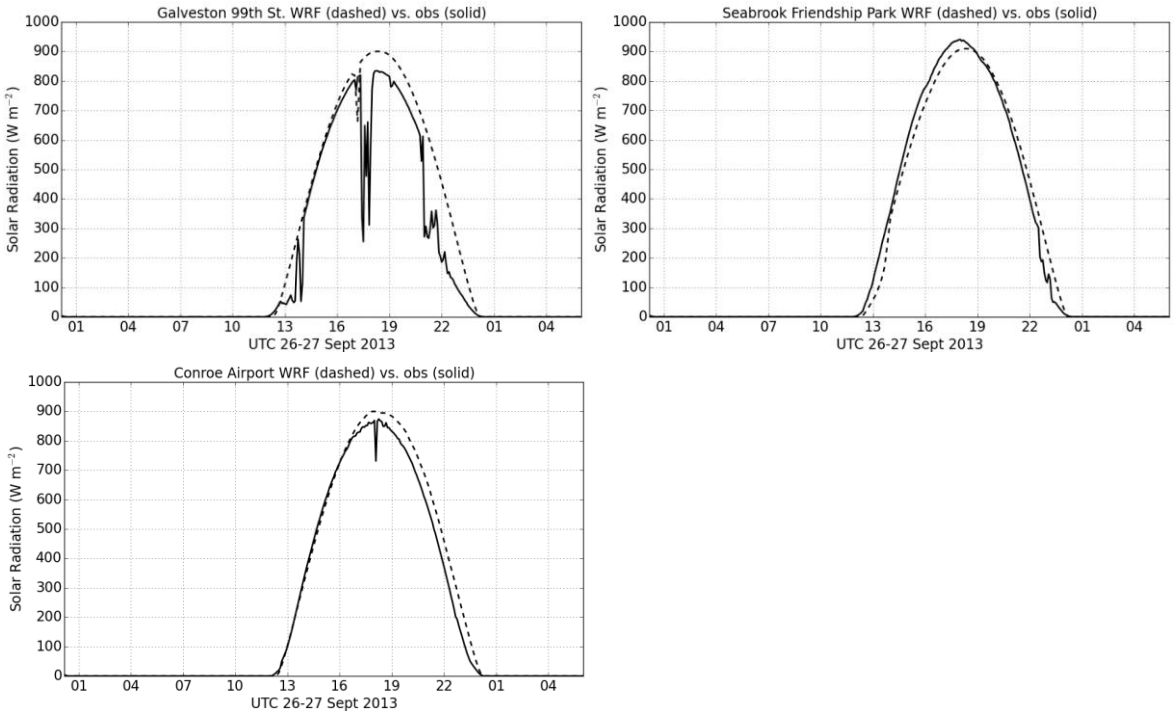

Figure 37, Observed (solid) and WRF Production run (CFG3) 1.33 km (dashed) solar radiation (w m^{-2}) for Galveston (top left), Seabrook (top right) and Conroe (bottom right) for September 26, 2013.

4. Conclusions and Recommendations for Future Study

The research performed for this study suggests that WRF run at high spatial resolution (1.33 km) can simulate fine-scale meteorological features, such as bay and sea breezes, that may contribute to high O_3 episodes in the Houston area provided that it is configured with the appropriate physical parameterizations. One of the key findings was that the choice of PBL and LSM schemes is critical, as simulations performed with a configuration that included the MYJ PBL and Noah LSM with an activated UCM did a noticeably better job than one with the ACM2 PBL and PXLSM at simulating the Galveston Bay and Gulf of Mexico sea breezes during the September 25th, 2013 high O_3 episode. Another finding was that although SST observations from buoys indicated a diurnal temperature cycle with an amplitude of about 2 K near the Texas coast during the September 25th and 26th high O_3 days, imposing this cycle on the high resolution SST analysis input to WRF produced only minor improvements to the simulated winds at some observations stations near the coastline. Finally, the UCM is an important part of the configuration, as turning it off in the Noah LSM but keeping all other WRF options the same increased the warm temperature bias in Houston and the offshore wind bias near the Galveston Bay coastline.

Based on the findings of our WRF tests for the high O_3 days of September 25th and 26th, we used the best WRF configuration to simulate the meteorology at 1.33 km for the 10 days preceding and including the two high O_3 episode days of DISCOVER-AQ. In addition, we simulated the entire month of September 2013 during which the DISCOVER-AQ campaign took place at 4 km resolution. All the WRF outputs were validated against observational data and were processed to provide inputs for the CMAQ and CAMx chemical transport models. Both the raw WRF data and processed data were provided to TCEQ for further analysis and retrospective air quality simulations.

The WRF output was validated with a quantitative statistical error analysis generatated by the WRF-MET software package and with qualitative comparisons of the 1.33 km WRF output with obsrvations at locations impacted by the sea and bay breezes or high O_3 mixing ratios during the high O_3 episodes of September 25th and 26th. All of the WRF-MET output plots and tables generated for the quantitative evaluation are being provided to TCEQ.

Summary statistics of simulated vs. observed near-surface variables from the innermost 1.33 km domain indicate that the hourly domain-wide bias and root mean square error (rmse) in 2 m temperature and 10 m wind had considerable variation over the 10-day period, with the majority of times having a temperature bias/rmse below 1 to 2 K, wind speed bias between 0.5 and 1.5 m s^{-1}, and wind speed rmse below 2 m s^{-1}. A set of month-long WRF performance statistics generated for a sub-section of the 4 km WRF domain covering the Houston-Galveston area indicated a wind speed bias of 0.77 m s^{-1} and rmse of 1.92 m s^{-1} with both being greater over land than water. For temperature the bias and rmse were 0.34 K and 1.68 K, respectively, and were also greater over land than over water. Overall these statistics are generally consistent with those from other meteorological modeling studies in the Houston – Galveston area.

Comparisions of time series of WRF winds and temperature from the 1.33 km domain with observations at key locations indicated that WRF was generally able to capture the development of sea and bay breezes during the high O_3 episodes but that at some locations (i.e., Galveston Island on September 25th and 26th and La Porte Airport during the 25th) the wind shift and temperature changes associated with the sea and bay breezes were delayed in the model by about 2-3 hours. However, observed and modeled winds were at times light (< 2.5 m s^{-1}) during these

periods and so this error may not significantly impact air quality simulations using these data. At other locations on the western shore of Galveston Bay and at Conroe Airport northwest of Houston the WRF winds and temperatures tracked the observations well. Skies were mostly clear during September 25th and 26th and WRF solar radiation outputs were generally within 10 % of observed values.

The WRF evaluations performed for this work order primarily used surface temperature and wind observations, as these were most readily comparable to the results of the previous WRF simulation presented by *Loughner et al. [2014]*. We also looked at mixing heights derived from Doppler profiles to make an initial assessment of the impact of vertical dilution of pollutants near the locations of the O_3 exceedences. However, there are many other observations made as part of the DISCOVER-AQ campaign including ozonsesonde profiles, mixing heights derived from airborne HSRL measurements and the Doppler wind vertical profiles that could be used to perform a more in depth analysis of the fine-scale meteorological features influencing Houston air quality and how well WRF is able to simulate them. Thus future studies should include this evaluation.

There are also additional WRF configuration options that could be considered for future study. For example in this work we used a single-layer UCM to account for the urban landscape of Houston and this requires that the average building height within a grid cell be below the top of the first model layer. This requirement imposes a potentially serious limit on the WRF model's vertical resolution. Future work could employ the multi-layer version of the UCM which does not have this restriction. In addition, a newer version of NARR with updated soil moisture fields has also recently become available and it might be valuable to see if this has an impact on the WRF simulations.

Finally, this work focused only on Houston for one specific month. Future work could be extended to other cities in Texas such as Dallas or San Antonio or other time periods of interest to TCEQ, such as the month of June 2012 being used for developing the State Implementation Plans.

5. References

Byun, D., H-C. Kim, and F. Nagin, (2011), Improvement of Meteorological Modeling by Accurate Prediction of Soil Moisture in the Weather Research and Forecasting (WRF) Model, *Project Summary prepared for TCEQ by NOAA ARL*, 46 pp.

Chen. F. and J. Dudhia, (2001), Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Wea. Rev.*, *129*, 569–585.

Chen, F., and Coauthors, (2011a), The integrated WRF/urban modeling system: Development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.*, *31*, 273–288.

Chen, F., S. Miao, M. Tewari, J-W. Bao and H. Kusaka, (2011b), A numerical study of interactions between surface forcing and sea breeze circulations and their effects on stagnation in the greater Houston area, J. Geophys. Res., 116, D12105, doi:10.1029/2010JD015533.

Chen, S.-H., and W.-Y. Sun, (2002), A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan*, *80*, 99–118.

Chin, T. M., J. Vazquez, and E. Armstrong, (2013), Algorithm Theoretical Basis Document: A multi-scale, high-resolution analysis of global sea surface temperature, Version 1.3, NASA JPL, Pasadena CA, 13 pp. (available at ftp://mariana.jpl.nasa.gov/mur_sst/tmchin/docs/ATBD/).

Grell, G. A., and D. Dévényi, (2002), A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, *29*, 1693, doi:10.1029/2002GL015311.

Grell, G. A., and S. R. Freitas, (2014), A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, *14*, 5233–5250.

Hong, S.-Y., Y. Noh, and J. Dudhia, (2006), A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, *134*, 2318–2341.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, *113*, D13103, doi:10.1029/2008JD009944.

Janjić, Z. I., (1994), The step-mountain eta coordinate model: Further developments of the convection, viscous layer, and turbulence closure schemes. *Mon. Wea. Rev.*, *122*, 927–945.

Janjić, Z. I., (2002), Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp.

Jiménez, P. A., and J. Dudhia, (2013), On the Ability of the WRF Model to Reproduce the Surface Wind Direction over Complex Terrain. *J. Appl. Meteor. Climatol.*, *52*, 1610–1617. doi: http://dx.doi.org/10.1175/JAMC-D-12-0266.1

Lee, P., H-C. Kim, and F. Nagin, (2012), Investigation of nocturnal surface wind bias by the Weather Research and Forecasting (WRF)/ Advanced Research WRF (ARW) meteorological model for the Second Texas Air Quality Study (TexAQS-II) in 2006, *Project Summary prepared for TCEQ by NOAA ARL*, 26 pp.

Lin, J. C., C. Gerbig, S. C. Wofsy, et al., (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *J. Geophys. Res.*, *108*, NO. D16, 4493, doi:10.1029/2002JD003161.

Loughner, C. P., M., Follette-Cook, K. E., Pickering, M., Estes, (2014), Bay breeze enhanced air pollution event in Houston, Texas during the DISCOVER-AQ field campaign**,** presentation at *13th Annual CMAS Conference*, October 27-29, 2014, Chapel Hill NC, available at https://www.cmascenter.org/conference/2014/agenda.cfm.

Mellor, G. L., and T. Yamada, (1982), Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys., 20,* 851–875.

Mitchell, K., and Coauthors, (2005), The Community Noah Land- SurfaceModel (LSM) user's guide, public release version 2.7.1. National Centers for Environmental Prediction/Environmental Modeling Center Tech. Doc., 26 pp. [Available online at ftp://ftp.emc.ncep.noaa.gov/mmb/gcp/ldas/noahlsm/ver_2.7.1/ Noah_LSM_USERGUIDE_2.7.1.doc.]

Nehrkorn, T., J. Henderson, M. Leidner, M. Mountain, J. Eluszkiewicz, K. McKain, and S. C. Wofsy, (2013), WRF simulations of the urban circulation in the Salt Lake City area for CO_2 modeling. *J. Appl. Meteor. Clim., 52,* 323-340.

Nehrkorn, T., M. Mountain, J. Hegarty, M. J. Iacono, Y. Barreira, P. Decola, and S. C. Wofsy, (2015), Sensitivity of high-resolution simulations over the Eastern Seaboard to surface process parameterizations. *16th Conf. on Mesoscale Processes*, Boston, MA, 3-6 Aug. 2015. American Meteorological Society, Boston, MA. Available online at https://ams.confex.com/ams/16Meso/webprogram/Paper274789.html.

Pleim, J. E., (2007), A combined local and non-local closure model for the atmospheric boundary layer. Part 1: Model description and testing. *J. Appl. Meteor. and Clim., 46,* 1383–1395.

Pleim, J. E., and A. Xiu, (1995), Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *J. Appl. Meteor., 34*, 16–32.

Pleim, J. E. and A. Xiu (2003), Development of a Land Surface Model. Part II: Data Assimilation. *J. Appl. Meteor., 42*, 1811–1822.

Skamarock, W. C., and J. B. Klemp, (2008), A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys., 227 (7),* 3465–3485.

Wang, W., C. Bruyere, M. Dudhia, and coauthors (2012), ARW version 3 modeling system user's guide. Tech. rep., MMM Division, NCAR, Boulder, CO, available from http://www.mmm.ucar.edu/wrf/users.

Xiu, A., and J. E. Pleim, (2001), Development of a land surface model part I: Application in a mesoscale meteorology model. *J. Appl. Meteor., 40,* 192–209

Appendix A. WRF Namelist File for Selected Configuration

```
 &time_control
 run_days = 1,
 run_hours = 6,
 run_minutes = 0,
 run_seconds = 0,
 start_year = 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013,
 start_month = 09, 09, 09, 09, 09, 09, 09, 09, 09,
 start_day = 25, 25, 25, 25, 25, 25, 25, 25, 25,
 start_hour = 00, 00, 00, 00, 00, 00, 00, 00, 00,
 start_minute = 00, 00, 00, 00, 00, 00, 00, 00, 00,
 start_second = 00, 00, 00, 00, 00, 00, 00, 00, 00,
 end_year = 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013,
 end_month = 09, 09, 09, 09, 09, 09, 09, 09, 09,
 end_day = 26, 26, 26, 26, 26, 26, 26, 26, 26,
 end_hour = 06, 06, 06, 06, 06, 06, 06, 06, 06,
 end_minute = 00, 00, 00, 00, 00, 00, 00, 00, 00,
 end_second = 00, 00, 00, 00, 00, 00, 00, 00, 00,
 interval_seconds = 10800,
 input_from_file = .true.,.true.,.true., .true.,.true.,.true.,
.true.,.true.,.true.,
 history_interval = 60, 60, 30, 10, 15, 15, 15, 15, 15,
 frames_per_outfile = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 restart = .false.,
 restart_interval = 2880,
 io_form_history = 2
 io_form_restart = 2
 io_form_input = 2
 io_form_boundary = 2
 io_form_auxinput4 = 2,
 auxinput4_inname = "wrflowinp_d<domain>",
 auxinput4_interval = 180,180,180,180,
 debug_level = 0
 /

 &domains
 time_step = 120,
 time_step_fract_num = 0,
 time_step_fract_den = 1,
 max_dom = 4,
 s_we = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 e_we = 163, 175, 217, 328, 251, 145, 251, 251, 251,
 s_sn = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 e_sn = 129, 139, 289, 301, 251, 149, 251, 251, 251,
 s_vert = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 e_vert = 44, 44, 44, 44, 44, 44, 44, 44, 44,
 eta_levels = 1.0000, 0.9960, 0.9900, 0.9800, 0.9700, 0.9600, 0.9500,
0.9400,
              0.9300, 0.9200, 0.9100  0.8950, 0.8800, 0.8650, 0.8500,
0.8250,
              0.8000, 0.7750, 0.7500, 0.7200, 0.6900, 0.6600, 0.6300,
0.6000,
```

```
              0.5700, 0.5400, 0.5100, 0.4750, 0.4400, 0.4050, 0.3700,
0.3300,
              0.2900, 0.2500, 0.2100, 0.1750, 0.1450, 0.1150, 0.0900,
0.0650,
              0.0450, 0.0250, 0.0100, 0.0000,
 num_metgrid_levels = 30,
 num_metgrid_soil_levels = 4,
 p_top_requested = 10000,
 dx = 36000, 12000, 4000, 1333.3333, 2000, 2000, 2000, 2000, 2000,
 dy = 36000, 12000, 4000, 1333.3333, 2000, 2000, 2000, 2000, 2000,
 grid_id = 1, 2, 3, 4, 5, 6, 7, 8, 9,
 parent_id = 0, 1, 2, 3, 4, 5, 6, 7, 8,
 i_parent_start = 1, 49, 67, 97, 5, 6, 7, 8, 9,
 j_parent_start = 1, 15, 16, 66, 5, 6, 7, 8, 9,
 parent_grid_ratio = 1,  3,  3,  3,  3,  3,  3,  3,  3,
 parent_time_step_ratio = 1,  3,  3,  3,  3,  3,  3,  3,  3,
 max_ts_locs=200,
 feedback = 1,
 smooth_option = 0
 /

 &physics
 mp_physics = 2, 2, 2, 2, 2, 2, 2, 2, 2,
 ra_lw_physics = 4, 4, 4, 4, 4, 4, 4, 4, 4,
 ra_sw_physics = 4, 4, 4, 4, 4, 4, 4, 4, 4,
 radt = 10, 10, 10, 10, 10, 10, 10, 10, 10,
 sf_sfclay_physics = 2, 2, 2, 2, 2, 2, 2, 2, 2,
 sf_surface_physics = 2, 2, 2, 2, 2, 2, 2, 2, 2,
 bl_pbl_physics = 2, 2, 2, 2, 2, 2, 2, 2, 2,
 bldt = 0, 0, 0, 0, 0, 0, 0, 0, 0,
 cu_physics = 3, 3, 3, 3, 3, 3, 3, 3, 3,
 cudt = 5, 5, 5, 5, 5, 5, 5, 5, 5,
 isfflx = 1,
 ifsnow = 1,
 icloud = 1,
 surface_input_source = 1,
 num_soil_layers = 4,
 sf_urban_physics = 0, 0, 0, 1, 1, 1, 1, 1, 1,
 num_land_cat = 33,
 maxiens = 1,
 maxens = 3,
 maxens2 = 3,
 maxens3 = 16,
 ensdim = 144,
 sst_update = 1,
 /

 &fdda
 grid_fdda = 1,0,0,0,0,
 gfdda_inname = "wrffdda_d<domain>",
 gfdda_interval_m = 180,
 gfdda_end_h = 30,30,30,30,30,
 io_form_gfdda = 2,
 fgdt = 0,0,0,0,0,
```

```
 if_no_pbl_nudging_uv = 1,0,0,0,0,0,0,0,0,
 if_no_pbl_nudging_t = 1,0,0,0,0,0,0,0,0,
 if_no_pbl_nudging_q = 1,0,0,0,0,0,0,0,0,
 if_zfac_uv = 0,0,0,0,0,0,0,0,0,
  k_zfac_uv = 10,10,10,10,10,10,10,10,10,
 if_zfac_t  = 0,0,0,0,0,0,0,0,0,
  k_zfac_t  = 10,10,10,10,10,10,10,10,10,
 if_zfac_q  = 0,0,0,0,0,0,0,0,0,
  k_zfac_q  = 10,10,10,10,10,10,10,10,10,
 guv = 0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,
 gt = 0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,
 gq = 0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,0.0003,
 if_ramping = 0,
 dtramp_min = 60.0,
 /

 &dynamics
 dyn_opt                             = 2,
 rk_ord                              = 3,
 w_damping                           = 0,
 diff_opt                            = 1,
 km_opt                              = 4,
 diff_6th_opt                        = 0,
 diff_6th_factor                     = 0.12,
 base_temp                           = 290.
 damp_opt                            = 0,
 zdamp                               = 5000.,  5000.,  5000.,  5000.,
5000.,  5000.,  5000.,  5000.,  5000.,
 dampcoef                            = 0.01,   0.01,   0.01,   0.01,
0.01,   0.01,   0.01,   0.01,   0.01,
 khdif                               = 0,      0,      0,      0,      0,
0,      0,      0,      0,
 kvdif                               = 0,      0,      0,      0,      0,
0,      0,      0,      0,
 smdiv                               = 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1,
0.1, 0.1,
 emdiv                               = 0.01, 0.01, 0.01, 0.01, 0.01, 0.01,
0.01, 0.01, 0.01,
 epssm                               = 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1,
0.1, 0.1,
 non_hydrostatic                     = .true., .true., .true., .true.,
.true., .true., .true., .true., .true.,
 moist_adv_opt                       = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 scalar_adv_opt                      = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 time_step_sound                     = 4, 4, 4, 4, 4, 4, 4, 4, 4,
 h_mom_adv_order                     = 5, 5, 5, 5, 5, 5, 5, 5, 5,
 v_mom_adv_order                     = 3, 3, 3, 3, 3, 3, 3, 3, 3,
 h_sca_adv_order                     = 5, 5, 5, 5, 5, 5, 5, 5, 5,
 v_sca_adv_order                     = 3, 3, 3, 3, 3, 3, 3, 3, 3,
 do_avgflx_em                        = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 do_avgflx_cugd                      = 1, 1, 1, 1, 1, 1, 1, 1, 1,
 /

 &bdy_control
```

```
 spec_bdy_width                      = 5,
 spec_zone                           = 1,
 relax_zone                          = 4,
 specified                           = .true., .false.,.false.,
.false.,.false., .false.,.false.,.false.,.false.,
 nested                              = .false., .true., .true., .true.,
.true., .true., .true., .true., .true.,
 /

 &grib2
 /

 &namelist_quilt
 nio_tasks_per_group = 0,
 nio_groups = 1,
 /
```

Appendix B. WRF-MET Derived Plots and Tables

The files may be obtained from the AER ftp server at ftp://ftp.aer.com/pub/jhegarty/TCEQ/. To download the files ftp to ftp.aer.com and login in anonymously using a username of anonymous and your email, then cd to pub/jhegarty/TCEQ, or use your web browser to go to the ftp link directly. The files are currently stored in compressed tar files (tar.gz) that will need to be expanded after they are downloaded. On a Linux system you can use the following command to expand the files.

tar –zxvf filename.tar.gz

After the tar.gz files have been expanded there will be many files of different types available to view.

Below is a description of the files that are available.

B.1 Notation

In the following,

- <n> refers to the domain number (3 for the 4km domain, 4 for the 1.33km domain)
- <stat> is either bias or rmse
- <levvar> denotes the level and variable:
  - Z10.DIR, Z10.SPD – 10 m wind speed (m/s) and direction (degrees)
  - Z2.DPT, Z2.RH, Z2.TMP – 2 m dewpoint (K), relative humidity (%), and temperature (K)
- <yyyymmdd>_<hh> - valid date (year/month/day) and time (hour), in UTC

B.2 Files for error statistics computed over the entire month, for each station

These statistics are computed for lead times from 7 to 30 hours.

Prod_<stat>_bubbleplot_d<n>_sfc_<levvar>__201309.pdf: bubble plots (on a map background) of error <stat> (rmse,bias) for domain <n> (2,3,4) for <levvar> (Z10.DIR, Z10.SPD, Z2.DPT, Z2.RH, Z2.TMP), similar to those in Figure 26.

The data used to produce these plots is available in text files, for example:
Prod_bias_bubbleplot_d3_sfc__201309.txt, *Prod_rmse_bubbleplot_d3_sfc__201309.txt*

Prod_d4_<stats>_diurnal<sid>.pdf: mean diurnal cycle of error <stat> for Z10.SPD and Z2.TMP for station <sid>, similar to the plots in Figure 25.

Prod_d4_windroses_<sid>.pdf: wind roses of observed and forecast winds, similar to Figure 27. Only computed for observed winds > 3 knots, using bins that align with 10 degree reporting intervals for ASOS. Sample sizes printed in graphic are for: number of matched pairs used in plot (above 3 knots), and in parenthesis: (all pairs with non-missing winds, all pairs).

Prod_d4_fcstobs_<sid>.pdf: multipanel plots of forecast and observed time series (Z2.TMP, Z2.DPT, Z10.SPD, Z10.DIR) for station <sid>, similar to Figure 28.

B.3 Single time period maps of observed/forecast values

Prod_d3_<var>_<yyyymmdd>_<hh>.pdf: for <var>=dewpts (Z2.DPT) and temps (Z2.TMP) these are plotted as color-coded concentric circles (obs: inner circle, fcst: outer annulus); for <var>=winds, plotted as wind barbs (speedx10, using standard convention otherwise), fcst:red, obs: black. Obs by themselves are plotted for <var>=obs_winds.

B.4 Summary plots of domain-wide statistics for surface (<lev>=sfc) andupper air (<lev>=upa) observations

plot_cnt_d3<lev>_201309.pdf: plots of domain-average stats as a function of lead time/level/valid time.

plot_mpr_d3<lev>_201309.pdf: plots of matched-pair stats as a function of lead time/level/valid time.