ASHRAE Research Project 1561-RP
Procedure to Adjust Observed Climatic Data for Regional or Mesoscale Climatic Variations

Final Report

Prepared for:
ASHRAE

By:

Xin Qiu, PhD
Novus Environmental

Fuquan Yang, PhD
Novus Environmental

Hamish Corbett-Hains, PEng
Novus Environmental

Michael Roth, PhD, PEng
Klimaat Consulting & Innovation

August, 2015

Executive Summary

Advanced weather forecasting models such as the Weather Research and Forecasting model (WRF) are capable of integrating meteorological data from a large number of observational sources (surface and upper air stations, radar, satellite and modeling data) in order to provide high-resolution climate data for locations where observed data is not available. This report presents the results of a WRF modeling campaign designed to cover ten significant climate regions across North America. Model results were compared against mesoscale monitoring data in order to assess the model's performance for a single year's hourly weather. Furthermore, a long-term climate model evaluation was performed by running WRF over 4 regions in North America for 8 years. Overall, the model performed well against observed temperature and humidity, reasonably well against observed wind, and relatively poorly against observed solar and precipitation. Guided by this evaluation, a complete mesoscale numerical modeling procedure was developed for coastal, mountain valleys, mountain plateaus, and major city centers, to provide Site Specific Climate Data, i.e., an accessible, freely-available software solution for developing localized climate data.

Table of Contents

Table of Figures

Table of Tables

1 Introduction

Current architectural and engineering practice involves careful consideration of local meteorology as a key factor in many design projects. Parameters such as wind speed, temperature, humidity, precipitation, and incoming radiation can all influence a building's design and on-going performance. ASHRAE members will typically use climate data derived from multi-year (minimum 8 years, typical 30 years) of measured data, provided by the ASHRAE society for locations worldwide[1]. However, this data is often taken from meteorological stations which are remote from a study site by 10s to 100s of kilometers (e.g. airports) or perhaps in completely different terrain (urban vs. rural, mountain vs. valley) which may not represent local climate at the design site. Factors such as urban heat island effects, sea-land breezes, and varying terrain can all significantly alter meteorological conditions between a weather station (often at an airport outside of an urban center) and an actual project site. Alternatively, there may be many stations located in close proximity to a design site and the user must select the most appropriate data source. For instance, the three major airports serving New York City (JFK, LaGuardia and Newark) are all located within a 15 km (9.3 mi) radius of the city center (Manhattan) and have a 0.8 °C (1.5 °F) difference in annual minimum temperature, a 2.4 °C (4.3 °F) difference in annual maximum temperature, a difference of almost 5,000 cooling degree hours (over 74 °F) and a difference in 5% extreme annual wind speed of 1 m/s (2.3 mph). While this may not seem like much of a difference, considering the number of HVAC systems in the city being designed to meet these conditions, the accuracy of climate data is important. Furthermore, the weather stations on Manhattan Island have questionable reliability. The challenging question commonly is which airport data should be treated as representative to a site on Manhattan Island

There are a number of techniques available to derive the meteorology of a given a site in the absence of a suitable observational station. These techniques often consist of some form of interpolation or "downscaling" of either the nearest stations or from a gridded large-scale climate dataset (e.g. NARR[2], MERRA[3], or CFSR[4]). This interpolation can be as simple as spatial linear interpolation of neighboring observational stations, to forming linear regression models based on predictors such as elevation or proximity to a coast (e.g. PRISM). At the far end of complexity, this "interpolation" can be guided by the solution of physical equations describing the conservation of mass, momentum, energy, and moisture, coupled with representations of atmospheric radiation and cloud formation. These "mesoscale" models represent the meteorology on grids approaching the microscale, i.e. 1 km in horizontal extent.

Although these mesoscale models can be used to refine historical climate data at a very high spatial resolution by applying physical principles to interpolate between areas where meteorological conditions are known, inherent with mesoscale model's complexity and power is

[1] (ASHRAE Technical Committee 4.2, 2009)
[2] NARR: North American Regional Reanalysis (http://www.esrl.noaa.gov/psd/data/narr/)
[3] MERRA: The Modern Era Retrospective-Analysis for Research and Applications (http://gmao.gsfc.nasa.gov/research/merra/)
[4] CFSR: The Climate Forecast System Reanalysis (http://cfs.ncep.noaa.gov/cfsr/)

a barrier to entry for first time users; it can be difficult to execute the model without fully understanding the physics schemes and mechanisms which are being used. Additionally, significant computational abilities are required to execute these models. In an effort to provide engineers and designers with more accurate meteorological data, ASHRAE Technical Committee 4.2 (Climatic Information) has sponsored research project 1561-RP to develop a protocol for modeling site-specific meteorological data with next-generation meteorological models.

ASHRAE RP-1561 was created with two overarching goals in mind:

- To develop a methodology for ASHRAE members to harness the power of modern mesoscale modeling techniques in order to derive meteorological conditions specific to a study area; and
- To evaluate this methodology against available observational data in a variety of geographic categories including coastal, mountain valley, mountain plateau and major cities.

Here, we first provide an overview of the modeling methodology developed and then evaluate the performance of the methodology in estimating both hourly meteorology, i.e. the "weather", and design elements such as the 99% heating dry-bulb temperature, i.e. the "climate". We finish with a discussion and evaluation of a simplified methodology, targeting more intermediate computer and skill levels.

2 Description of Mesoscale Modeling Procedure

The following is an overview of the methodology developed to generate site-specific meteorological time series. We discuss both an advanced methodology (referred to as Novus WRF), suitable for those with "advanced" computer and meteorological skills and with access to extensive computer hardware, and a more "simplified" methodology (referred to as WRF EMS), suitable for those with more intermediate skills and hardware. This section focuses on the discussion and evaluation of the advanced methodology. A separate discussion and brief evaluation of the simplified methodology is provided in Section 5.

2.1 Weather Research and Forecasting Model (WRF)

There are several numerical meteorological models commonly used within the meteorological community for high-resolution weather forecasting which are capable of producing high-quality gridded hourly climate data. The Weather Research and Forecasting model (WRF) was used in this study due to its ability to generate high-resolution and reliable climate data at any location within its modeling domain. The WRF model is a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. WRF is suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers. The WRF model is developed and maintained as part of a collaborative effort principally among U.S. government agencies, universities, national laboratories and international communities (Skamarock et al., 2008).

2.2 Domain Configuration

The modeling procedure begins with the establishment of the model domain. For a given location, a series of grids are created with the site in question typically central. Domains defined in this report are very large based on the requirements of scope of work (SOW) in RP-1561. The final recommended domain sizes for ASHRAE members are much smaller than those in this study for ease of computing. The parent or first domain is a large domain spanning approximately 1800 km × 1800 km (1120 mi × 1120 mi) with 36 km × 36 km (22.5 mi ×22.5 mi) cells. Nested in this domain is a medium resolution child domain which is 730 km × 730 km (450 mi × 450 mi). Nested in this domain is an even finer resolution child domain which is 280 km × 280 km (170 mi × 170 mi). The grid resolution in this domain is approximately 4 km (2.5 mi). Finer resolution is achievable, down to a grid approximately 1.3 km (0.8 mi).

Mesoscale modeling is three-dimensional; the methodology divides the atmosphere from the ground surface to the top of the troposphere (around 100 hPa (1.45 psi)) into 35 vertical layers. Model topography is interpolated from the 30 arc sec (about 1 km) USGS Geophysical Data Center global data which is included with the WRF preprocessing system. Vegetation type and land use data at 1 km (0.6 mi) horizontal resolution are also interpolated to the WRF grids.

2.3 Input Data

The mesoscale model uses the output from another three-dimensional model or "reanalysis dataset" as an input. The input model must have a greater area of coverage and is processed as follows:

- The larger model is interpolated to the WRF model and serves as the initial state of the atmosphere, i.e. the "initial conditions"; and,
- The larger model also provides the air masses that enter the WRF model along the lateral boundaries as the simulation evolves, i.e. the "boundary conditions".

The data used in this report consisted of modeled 32 km (22.5 mi) North American Regional Reanalysis (NARR) data produced and distributed by the National Centers for Environmental Prediction (NCEP) of NOAA as well as upper air and surface observational data from the same dataset used to produce the ASHRAE climate data found in the Handbook of Fundamentals. While NARR was used throughout this project as a WRF model input, readily available alternatives include CFSR and MERRA, two datasets with global reach.

In addition, actual observational data is directly ingested into the model. This observational data includes NCEP global surface and upper air observational weather data, including airport ground stations, radiosondes, pibals and both aircraft reports and satellite data with measurements of pressure, geopotential height, air temperature, dew point temperature, wind direction and wind speed. This observed meteorological data was used to "nudge" the WRF outputs toward real world data, ensuring it does not "drift" too far away from reality. Note that the simplified methodology, in a significant departure, does not use this observational station nudging, reducing both complexity and overall simulation time.

2.4 Model Physics

Given the model domain setup and the initial and boundary conditions, physical parameters describing the various physics modules within the WRF model were selected. These module selections are crucial for correct modeling results due to the fact that none of the physics options are universal; different physics settings could produce different model results. In this study, the physics models were selected based on geographic location, season, and our experience on modeling parameters in past WRF modeling studies. After analyzing a series of model sensitivity tests in WRF-ARW model version 3.4.1, the following physics options were recommended for the final WRF simulation:

- **Cumulus Parameterization.** This is a cloud and precipitation physics option applied to coarse WRF domains (36 km (20 mi) and 12 km (7.5 mi)) in order to parameterize cloud processes size within a grid cell as a function of larger-scale weather processes and conditions. Here, the Kain-Fritsch cumulus parameterization scheme (Kain and Fritsch, 1993) was used. This scheme uses sophisticated cloud mixing to determine entrainment/detrainment within the cloud and removes all available buoyant energy. Updraft and downdraft properties were also predicted by this module. A Cumulus Parameterization scheme was not used in the fine 4 km (2.5 mi) domain as the model is able to start explicitly resolving cloud dynamics at this resolution.
- **Planetary Boundary Layer Scheme.** Planetary boundary layer (PBL) physics models estimate low-level wind, temperature, turbulence, cloud cover, and radiation. The Yonsei University scheme (YSU) was selected for this study. This is a non-local-K (diffusivity) scheme with an explicit entrainment layer and parabolic K profile within an unstable mixed layer.
- **Explicit Moisture Scheme.** In the WRF model, ten algorithms can be chosen from to represent the cloud microphysics scheme applied to the fine grid domain (4 km (2.5 mi)). The explicit moisture scheme is used to resolve cloud formation and precipitation within a grid cell. Among the microphysics packages for clouds and precipitation tested, the single-moment WSM6 package (Hong and Lim, 2006) was selected for the WRF simulation.
- **Radiation Schemes.** The Rapid Radiative Transfer Model (RRTM) (Mlawer et al., 1997) longwave radiation scheme was used in the modeling. This longwave radiation scheme is a highly accurate and efficient method that uses a correlated-k model to represent the effects of the detailed absorption spectrum. The model accounts for water vapor, carbon dioxide and ozone. For short-wave solar radiation, the Dudhia shortwave radiation scheme (Dudhia, 1989) was used, which considers clouds, clear-sky absorption and scattering efficiently. Significantly improved results for precipitation and shortwave radiation can be generated provided the levels of atmospheric constituents such as ozone and aerosols are available as inputs. However, for the purposes of this project, robust domain-wide radiation schemes were utilized.
- **Land Surface Scheme.** The commonly used Unified Noah land-surface model (Chen and Dudhia, 2001) – a 5-layer thermal diffusion scheme – was used to simulate surface skin temperature.
- **WRF Three-Dimensional Analysis Nudging.** WRF can use both 3D and 2D (surface) reanalysis gridded data to nudge the output. Note that WRF surface analysis nudging can

take advantage of more frequent gridded surface analyses (e.g., hourly or 3-hourly) than the 3D analyses (e.g., 6-hourly or 12-hourly), and the former are applied within the model diagnosed PBL.

- **WRF Observational Nudging.** This is a data assimilation method for WRF modeling to utilize available climate/weather observation data from ground-based or upper air meteorological monitoring processes, such as, weather station, upper air balloon, aircraft and satellite data. The widely used method is Newtonian relaxation nudging which consists of adding an empirical term to the prognostic equations that nudges the solution towards the observations. In this study, conventional weather station and upper air sounding data were used in the observational nudging process.

2.5 Model Simulation

First, a multi-year run is split into a number of 3-day periods (i.e., 72-hr chunks). Each period is preceded by adding a 12-hour "spin-up" time at the beginning. For example, a single period may be setup to run from Jan 1st to Jan 3rd, 2014. The "spin-up" added on the front of this simulation is a 12 hour portion (12:00 to 24:00 in Dec 31), that represents the time needed for the simulation to adjust to the imprecise initial conditions. Thus total run period becomes 84-hr. Beyond this spin-up time, a realistic flow field has developed from which a usable portion of the simulation is generated. Splitting the simulation into fully-contained, independent periods allows minimizing model cumulative (systematic) errors along runtime, and running different time periods in parallel, if resources are available. That is, the simulation does not have to be run in sequence saving a significant amount of time.

2.6 Post-processing

The output of the WRF simulations are a massive set of files (~1TB/year/study area) consisting of hourly three-dimensional fields. As the ultimate goal of the methodology is to produce a time series of the relevant meteorological quantities at the surface, these three-dimensional files need to be mined. Armed with such a time series, derived products such as Typical Meteorological Year (TMY) files or climatic data sheets following Handbook of Fundamentals (HOF) criteria can be generated.

3 WRF Model Meteorological Evaluation

One of the main objectives of this project was to determine if the WRF model and proposed methodology sufficiently approximates meteorological data to ASHRAE's standards for distribution. Furthermore, this evaluation sought to provide reasonable estimates of the model's biases and errors. To address these objectives, it was necessary to complete both an operational evaluation (i.e., quantitative, statistical, and graphical comparisons) as well as a more phenomenological assessment (i.e. qualitative comparisons of observed features vs. their depiction in the model.

In this Section, annual simulations (2008) from multiple domains were evaluated to determine how well the proposed methodology correlates against the hourly observed meteorological data (i.e. the "weather"). In Section 5, for a series of 8-year simulations, we examine how well the

methodology correlates against the observed statistical data (i.e. the "climate").

3.1 Domain Configuration

For the purposes of the weather evaluation, the WRF model was evaluated over five large modeling domains within North America. The largest domain has 176 x 138 grid cells, each 36 km (22.5 mi) by 36 km (22.5 mi), covering most of North America. The middle domains have approximately 160 × 140 12 km (7.5 mi) grid cells. The smallest domains, roughly centered on the New York, Florida, New Orleans, Los Angeles, San Francisco, Denver, and Seattle have approximately 145 x 160 grid cells, each with 4 km (2.5 mi) by 4 km (2.5 mi) grid resolution. The purpose of 4-km (2.5-mi) domain sections is to cover the most representative land use categories: coastal locations, mountain valleys, mountain plateaus and major city centers.
Figure 1 shows only the eight most inner 4-km (2.5-mi) domains. 4-km domains which were within a shared 36 km (22.5 mi) modeling domains are indicated by a common colored fill.

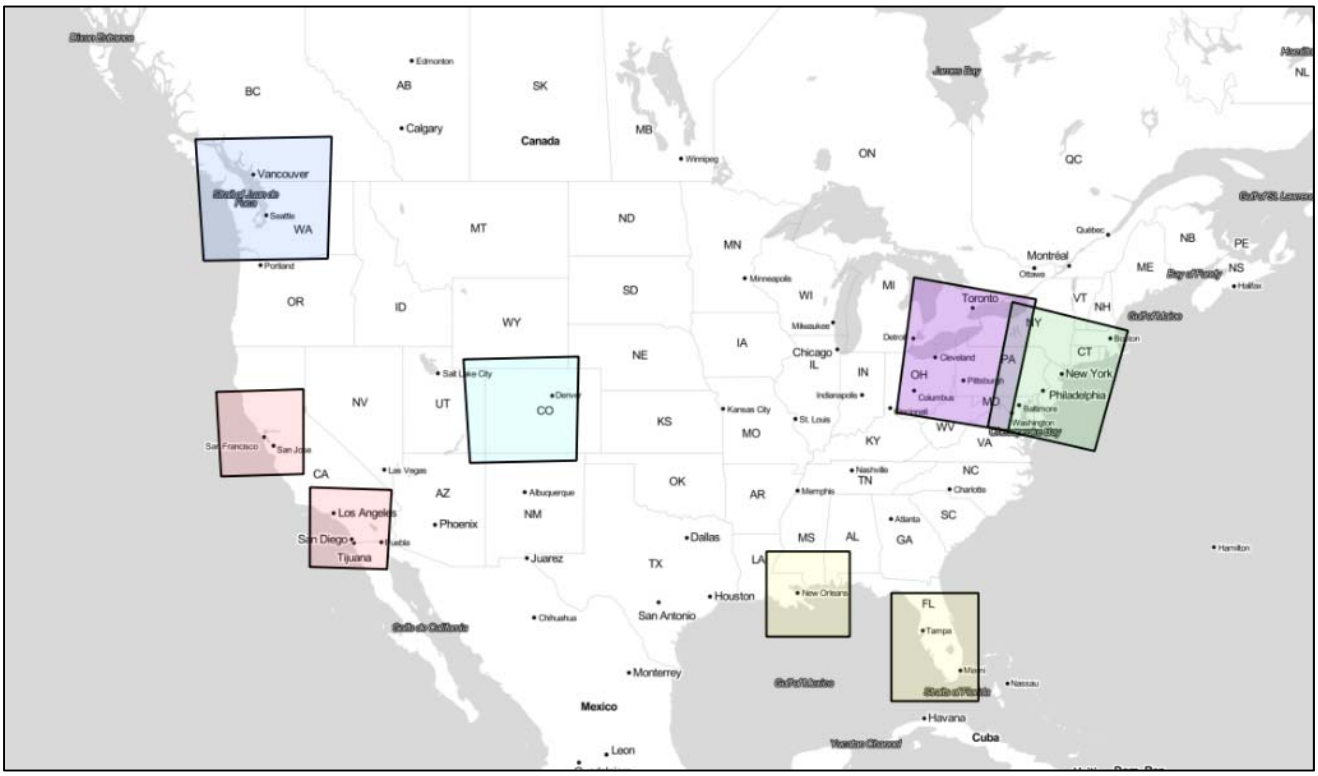

Figure 1 – WRF Model 4-km (2.5-mi) Domains; Color-coding Denotes Common 12-km (7.5 mi) Parent Domains (not illustrated)

3.2 Performance Criteria

The WRF evaluation focuses on comparisons of specific modeled hourly meteorological parameters to observed hourly data. It is recommended that comparative observation and model locations be paired as closely as possible in space and time (U.S. EPA, 2007). Typical statistical parameters of the key climate and meteorological factors include: correlation coefficient (R – a

measure of the linear correlation between two variables), mean bias (MB – the average difference between two datasets), mean normalized bias (MNB – a scaled measure of the difference between two datasets), mean gross error (MGE – the average absolute difference between two datasets), mean normalized gross error (MNGE – a scaled measure of the average absolute difference between two datasets), root mean square error (RMSE – the average absolute difference between two datasets with weighting given to larger differences), and an index of agreement (IOA – a measure of the agreement between two datasets). These metrics are commonly used for weather forecasts, local climate modeling and air quality studies.

Performance criteria and benchmarks for WRF meteorological model results were based on the evaluation protocol of Zhang et al. (2006) and Wu et al. (2008) which are recommended for the evaluation of mesoscale meteorological models such as the WRF model used in this study. The typical suggested benchmarks for temperature, humidity ratio, wind speed and wind direction for both regular and complex terrains are given in Table 1. They are defined as:

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}(Model - Obs)^2}$$

Mean Gross Error (MGE):

$$MGE = \frac{1}{N}\sum_{1}^{N}|Model - Obs|$$

Mean Bias (MB):

$$MB = \frac{1}{N}\sum_{1}^{N}(Model - Obs)$$

In this paper, three important meteorological parameters were evaluated including:

- Surface temperature (at 2 m (6.5 ft) elevation in Celsius (Fahrenheit)),
- Wind speed (at 10 m (33 ft) in m/s (mph)) and direction (in degrees), and
- Specific humidity (at 2 m (6.5 ft) in g/kg (grains/lb)).

3.3 Observational Mesonet Data

The Mesonet network of meteorological stations was chosen given its high spatial resolution and availability of data. The Mesonet network has stations all across the continent which are used to monitor mesoscale severe weather events such as thunderstorms, tornados and winter storms.

Importantly, as Mesonet station data was not used as an input to the WRF modeling, they can provide an unbiased evaluation of the WRF results.

The WRF model was assessed against available Mesonet station data in four different geographic categories (coastal, mountain valley, mountain plateau and major cities), as presented in in Table 2. The quality of the Mesonet data varies significantly from station to station; 148 Mesonet stations contained sufficient data (>80% by time) and were located within the 4-km (2.5 mi) modeling domains. Maps showing the station locations can be found in Appendix B. Model evaluation was performed over eight 4-km grid resolution modeling domains which were grouped into four different geographic categories mentioned above. Note that as some stations fall into two land-use categories, the total number of stations presented in the comparison graphs is higher than the 148 stations used.

Table 1 – Performance Criteria for WRF Meteorological Model Results

Terrain	Wind Speed	Wind Direction	Temperature	Humidity
Regular Terrain (Coastal, Major Cities)	RMSE: < 2 m/s (4.5 mph) MB: < ±0.5 m/s (1.1 mph)	MGE: < 30 deg MB: < ±10 deg	MGE: < 2 °C (3.6 °F) MB: < ±0.5 °C (0.5 °F)	MGE: < 2 g/kg ($2*10^{-3}$lb/lb) MB: < ±0.75 g/kg ($0.75*10^{-3}$lb/lb)
Complex Terrain (Mountain Valley, Mountain Plateau)	RMSE: < 2.5 m/s (5.6 mph) MB: < ±1.0 m/s (2.2 mph)	MGE: < 50 deg MB: < ±10 deg	MGE: < 3.5 °C (6.3 °F) MB: < ±1.0 °C (1.8 °F)	MGE: < 2 g/kg ($2*10^{-3}$lb/lb) MB: < ±1 g/kg ($1*10^{-3}$lb/lb)

Table 2 – Categories of Cites for Mesonet/WRF Comparison

City	Number of Mesonet Stations	Coastal	Mountain Valley	Mountain Plateau	Major City
New York	13	x			x
Boston	13	x			x
Toronto	14	x			x
Orlando	5				x
New Orleans	13	x			x
Harrisburg	12		x		
San Francisco	8	x			x
Palm Springs	10		x		
Los Angeles	15	x			x
Denver	3			x	
Vancouver	15	x	x		x
Colorado Springs	11			x	
San Bernardino	9		x		
Sacramento	5			x	
Seattle	2	x			x

3.4 Temperature

WRF generated temperature values for every 4 km (2.5 mi) model grid from the deepest soil layer to the ground surface to the top of atmosphere (the modeled domain cap). To compare with observational data, the 2 m (6.5 ft) temperature was extracted and used in this study. For each station, we calculated three statistics (MBE, MGE, and R). A total of 148 Mesonet locations were used to check the modeling performance.

Figure 2 contains box-and-whisker plots of WRF modeled temperature performance compared against Mesonet data across the four geographic categories. In Figure 2 through Figure 7 the boxes indicate the 25% to 75% quartile range of the relevant statistic across the stations in each category while the whiskers indicate the 10% and 90% percentiles. The greyed regions indicate the performance criteria from Table 1.

Comparison of the model results to measured (Mesonet) data is also summarized in Table 3. The model temperatures track the observations with a median correlation coefficient (R) of 0.95; the variability of the model temperatures match the observed values well within performance criteria (average MB = -0.35 $^{\circ}$C (0.63 $^{\circ}$F)) as shown in Figure 2.

The WRF model performance varies between stations within the modeling domains with MB's ranging over the 148 measurement locations from -4.1 $^{\circ}$C to 1.9 $^{\circ}$C (-7.4 $^{\circ}$F to 3.5 $^{\circ}$F) and MGE's ranging from 0.7 $^{\circ}$C to 4.2 $^{\circ}$C. Temperatures were predicted best in simple terrain. In complex terrain the model performed within the allowable range for MGE; for mountain plateaus the

model was within the allowable range for MB; the model on average was outside the allowable range for MB for mountain valleys (-1.15 °C vs -0.23 °C (-2.07 °F vs -0.41 °F)).

Figure 3 shows the cumulative distribution function and probability distribution function for the modeled and measured temperature data across the four geographic categories. This comparison shows that with the exception of the mountain valley climate zone the modeled temperature tracks measured temperature very well. In the mountain valley climate zone, the model has a slight positive bias. Note that due to measurement and unit conversions, observed temperatures may snap to the nearest °C or °F. Subsequent additional unit conversions can lead to temperatures preferentially spilling out of some histogram bins into adjacent bins, resulting in gaps in the histograms.

Table 3 – WRF Temperature Performance vs Mesonet across four Geographic Categories

Climate Zone	MB (°C | °F)	MGE (°C | °F)	R
Coastal	-0.01 | -0.02	1.46 | 2.63	0.95
Mountain Valley	-1.15 | -2.07	2.12 | 3.82	0.96
Mountain Plateau	-0.23 | -0.41	1.98 | 3.56	0.97
Major City	-0.02 | -0.04	1.48 | 2.66	0.95

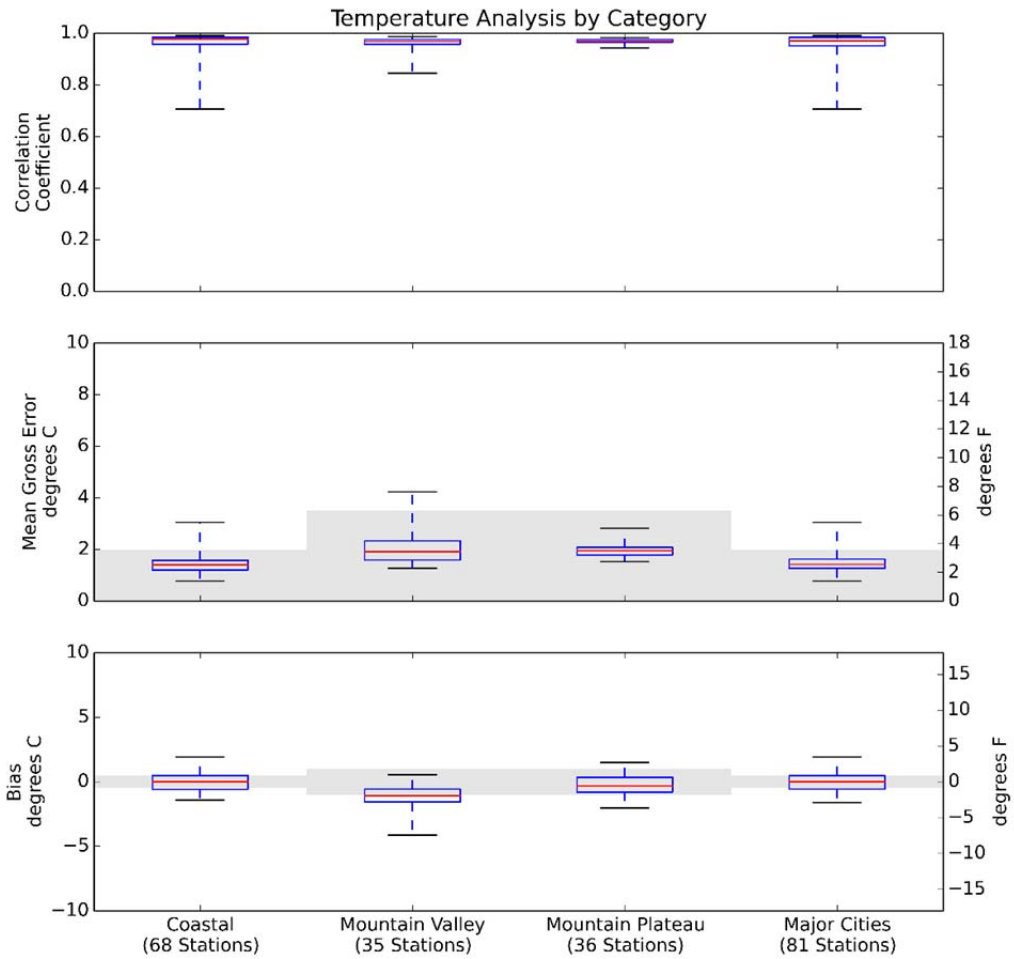

Figure 2 – Comparison of Modeled (WRF) and Measured (Mesonet) Temperatures across Four Geographic Categories. The boxes indicate the 25% to 75% quartile range of the relevant statistic across the stations in each category while the whiskers indicate the maximum and minimum values. The greyed regions indicate the performance criteria from Table 1.

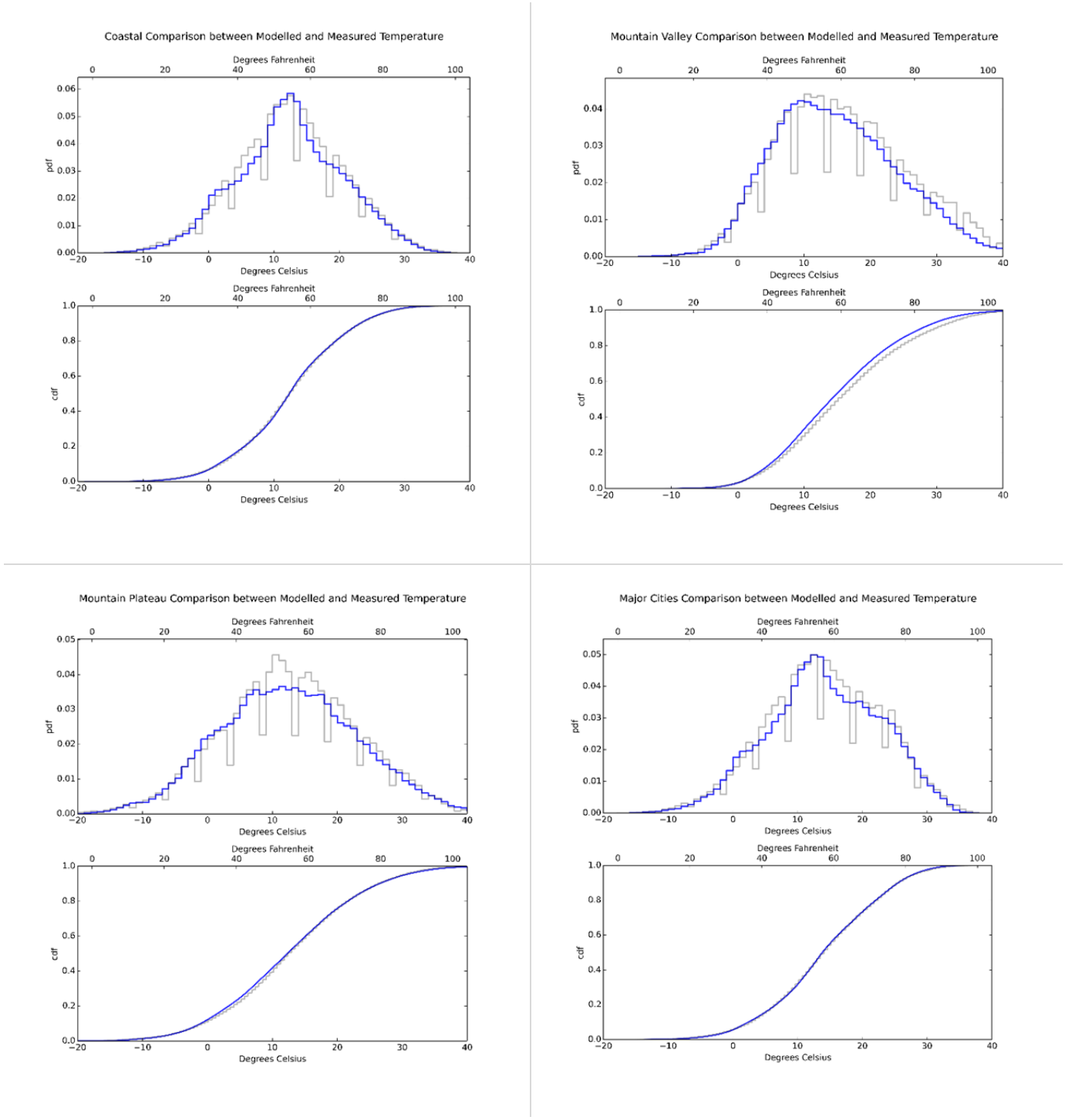

Figure 3 – Cumulative Distribution Function and Probability Distribution Function for WRF (blue)-Mesonet (grey) Temperature Comparison

3.5 Humidity

Specific humidity is defined as the mass of water vapor per unit mass of moist air in units of g/kg (grains/lb). Table 4 shows the summary of WRF predicted specific humidity performance. The mean MGE's for the WRF 2-meter (6.5 ft) specific humidity were within the 2 g/kg (14 grains/lb) threshold for every climate zone, and three of the four geographic categories had maximum MGE's below the threshold. In fact, only the *Major Cities* climate zone had any values outside the allowable range. It can be seen that WRF produced a consistently negative bias, although the median MB's for every climate zone are within the allowable range (±0.75 g/kg (± 5.25 grains/lb) for simple terrain, ±1 g/kg (±7 grains/lb) for complex terrain). Additionally, the range of correlation coefficients seen (0.83-0.93) indicates that the WRF model is matching measured data well.

Figure 4 shows humidity ratio across the four modeling zones (geographic categories). The variability of the model humidity ratio matches the observed values well within performance criteria (greyed area). Figure 5 shows the cumulative distribution function and probability distribution functions for the humidity ratio comparison. These results show that for all regions, and especially for mountain plateaus, the model has a tendency to slightly over-predict at low humidity levels (i.e. relatively dry ambient conditions) however overall modeling performance are mostly within the acceptable criteria ranges (grey area)

Table 4 – WRF Humidity Ratio Performance vs Mesonet across four Geographic categories

Climate Zone	MB (g/kg, grains/lb)	MGE (g/kg, grains/lb)	R
Coastal	-0.24 | -1.68	0.87 | 6.09	0.92
Mountain Valley	-0.48 | -3.36	0.97 | 6.79	0.87
Mountain Plateau	-0.64 | -4.48	1.08 | 7.56	0.83
Major City	-0.30 | -2.10	0.97 | 6.79	0.93

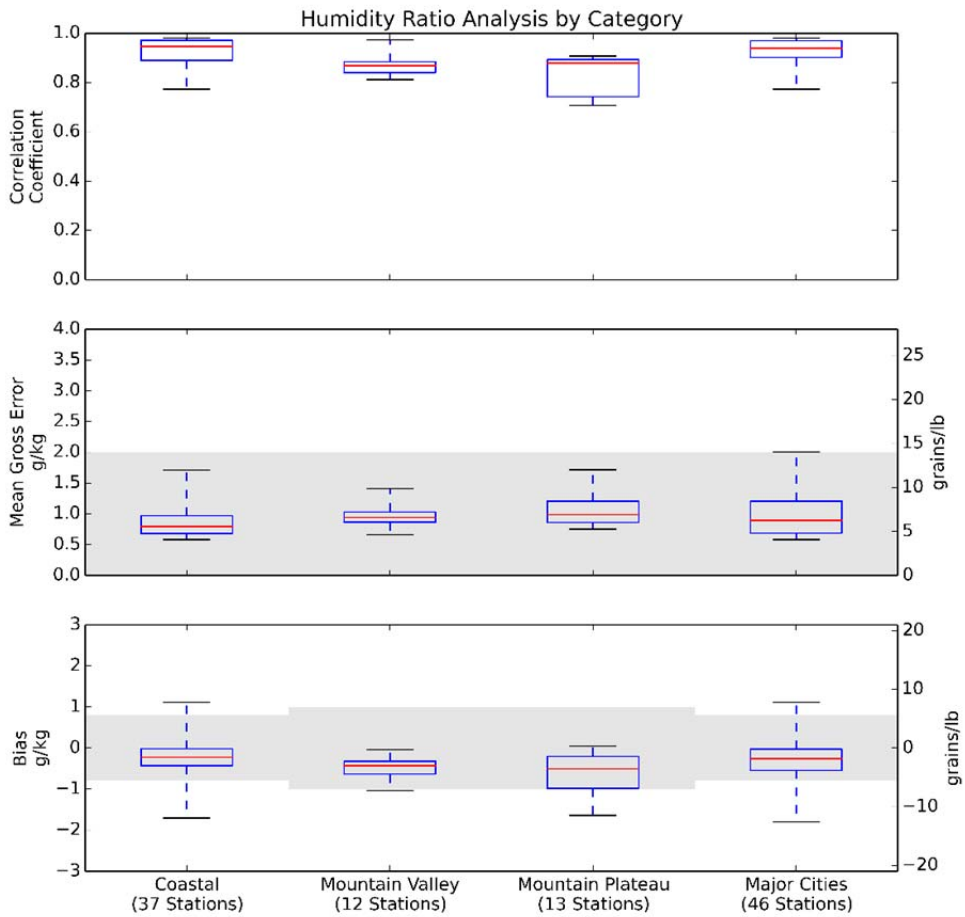

Figure 4 – Comparison of Modeled (WRF) and Measured (Mesonet) Humidity Ratio across Four Geographic Categories

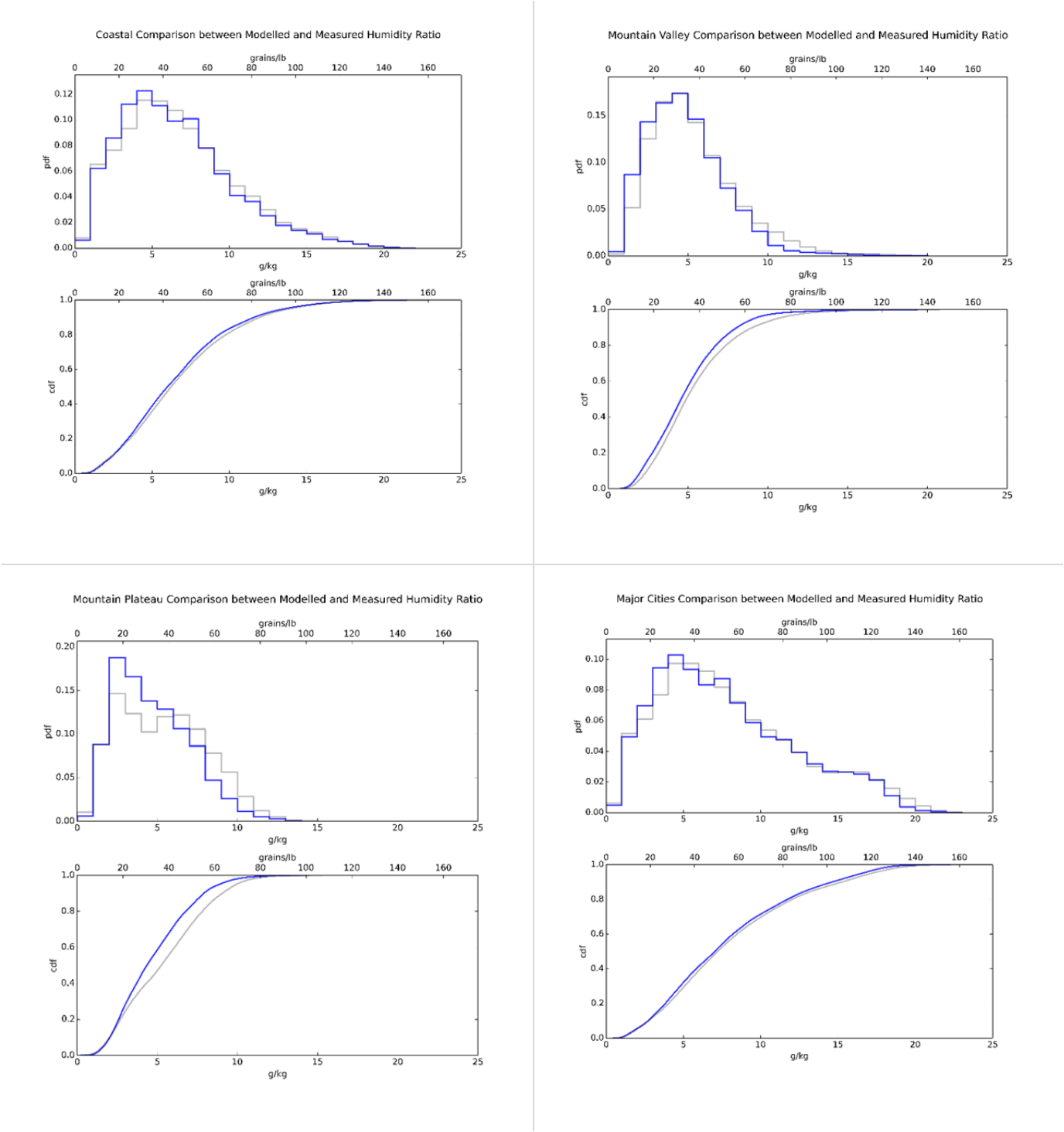

Figure 5 – Cumulative Distribution Function and Probability Distribution Function for Humidity Ratio across Four Geographic Categories

3.6 Wind

Table 5 shows the benchmarks for WRF-predicted wind speed performance. Figure 6 illustrates WRF modeled wind speed performance by the four geographic categories. Low wind speeds (below 1 m/s (2.2 mph)) were removed from this comparison as weather stations return unreliable measurements below wind 1 m/s. The MBs for the WRF 10-meter (33 ft) wind speed were slightly outside the allowable range for every climate zone (0.5 m/s (1.1 mph) for simple terrain, 1 m/s (2.2 mph) for complex). The RMSEs were all between 2.2 m/s and 3.3 m/s; every climate zone except for mountain plateaus was slightly outside the acceptable range (2 m/s (4.9 mph) for simple terrain, 2.5 m/s (5.6 mph) for complex terrain). Biases are mostly within or very close to the acceptable criteria range.

Wind direction is defined as the direction the wind blows from, in degrees (0 to 360) clockwise from North. Table 1 shows the allowable ranges for WRF predicted wind direction performance. Table 6 shows that the modeled MB was within the allowable range ($\pm 10^{o}$) for every climate zone. Figure 7 illustrates WRF modelled wind direction performance by the four geographic categories. MGE was within the allowable range for complex terrains (<50 degrees), while the Coastal and Major City geographic categories were outside the allowable range (<30 degrees). Biases are mostly within the acceptable criteria range.

It is not surprising that microscale wind variations (observational speed and direction) do not agree with the mesoscale data as well as temperature. Many of the Mesonet stations are poorly sited, being in close proximity to large structures (buildings) or vegetation (tree canopies), not located at airports or other well-situated sites. Thus, individual observational error can be large, resulting in large mean gross and root mean square errors. However, when combining all of the stations in a given category or region, the variations even out, providing a satisfactory mean bias.

When assessing wind speed and direction, a more meaningful way to determine the model's accuracy is to use data taken from a more representative dataset such as the NOAA network of weather stations. These stations collect data without mesoscale interference as they are typically situated in areas like airports where they are separated from obstructions which may cause localized interference.

Table 5 – Summary of WRF Wind Speed Performance

Climate Zone	MB (m/s | mph)	RMSE (m/s | mph)	R.
Coastal	0.98 | 2.2	2.23 | 5.0	0.41
Mountain Valley	2.11 | 4.7	3.31 | 7.4	0.38
Mountain Plateau	1.36 | 3.0	2.42 | 5.4	0.37
Major City	1.09 | 2.4	2.28 | 5.1	0.40

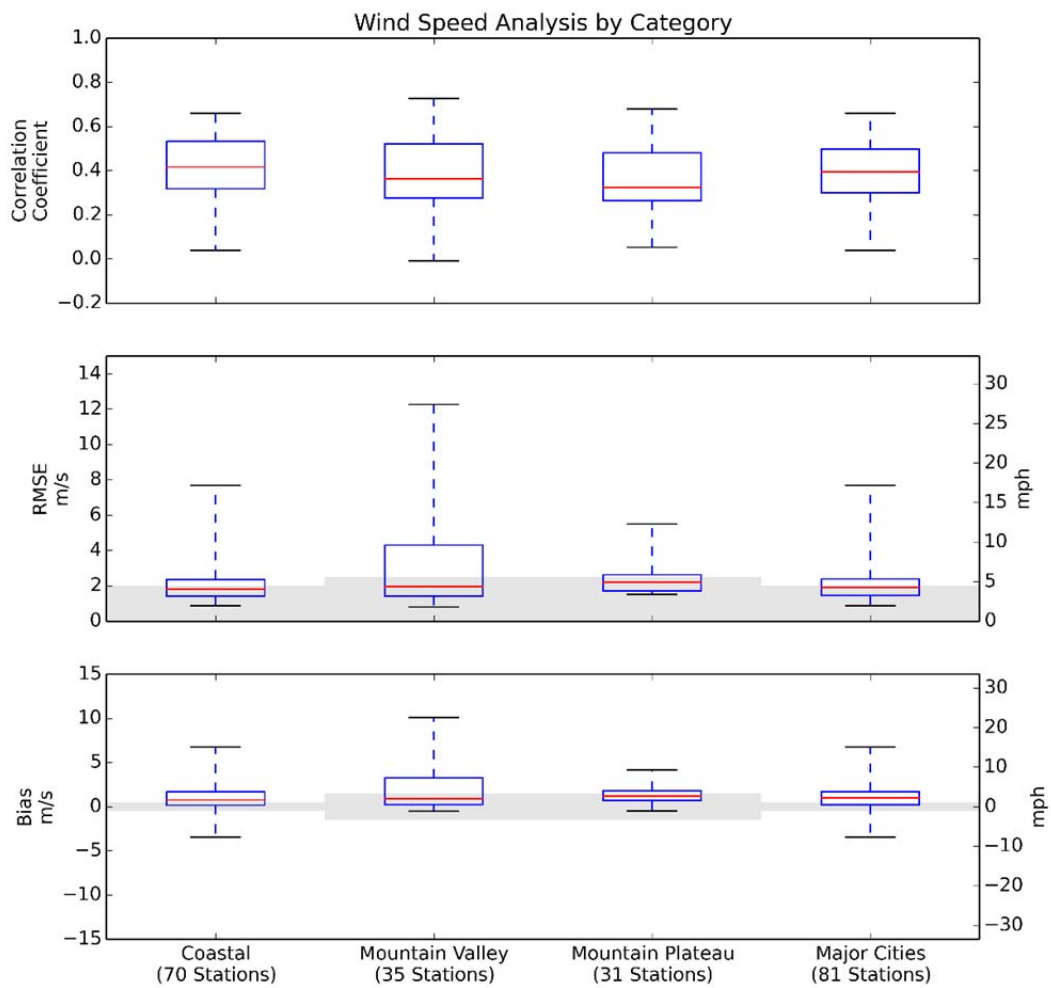

Figure 6 – Comparison of Modeled (WRF) and Measured (Mesonet) Wind Speeds across Four Geographic categories

Table 6 – Summary of WRF Wind Direction Performance

Climate Zone	MB (degrees)	MGE (degrees)	R.
Coastal	3.8	39.9	0.40
Mountain Valley	4.6	45.2	0.38
Mountain Plateau	2.9	49.8	0.29
Major City	2.3	40.2	0.39

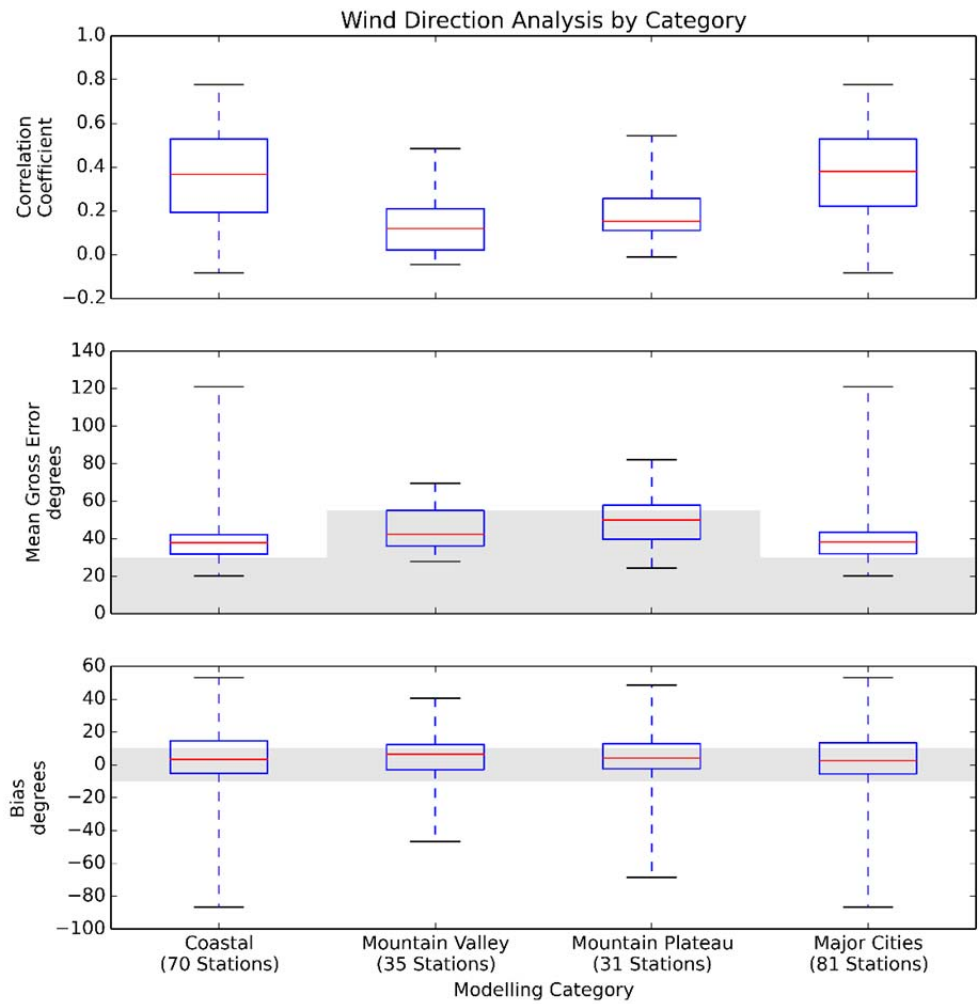

Figure 7 – Comparison of Modeled (WRF) and Measured (Mesonet) Wind Speeds across Four Geographic categories

4 Climate Evaluation

In order to evaluate the ability of the methodology to represent long term climatic trends and statistics, modeling was conducted for an 8-year period, 2005-2012. 8 years was chosen as it represents the shortest allowable measurement period for entry in the HOF data. The aim of the modeling is to generate information sufficient to estimate the parameters in the HOF tables, with the exception of clear-sky radiation coefficients which require a separate modeling procedure.

Whereas the previous "weather" evaluation concentrated on how well the methodology matched the observations on an hour by hour basis, the focus here is how accurately the proposed methodology can represent averages (e.g. average daily temperature in July), percentiles (e.g. 1% annual wind speed), and integrated values (e.g. average incident daily solar radiation received at the ground in January).

We start with a similar modeling methodology as for the meteorological evaluation, though here we limit the domains to four 4-km regions surrounding the cities of New York, Denver, San Francisco, and Orlando, as show in Figure 8. All four regions represent challenging modeling environments for differing reasons:

- New York is close to the moderating influence of the Atlantic but sees a large seasonal contrast and snowfall;
- Denver is within the influence of the Rocky Mountains and subject to rapid changes in weather;
- San Francisco is adjacent to westerly winds blowing over cold Pacific waters which can produce its characteristic fogs; and,
- Orlando has many of the hallmarks of a tropical climate and, due to the adjacent Gulf of Mexico to the west and Atlantic to the east, can see moisture convergence, intense rainfall, and high humidity, in addition to the risk of hurricane activity.

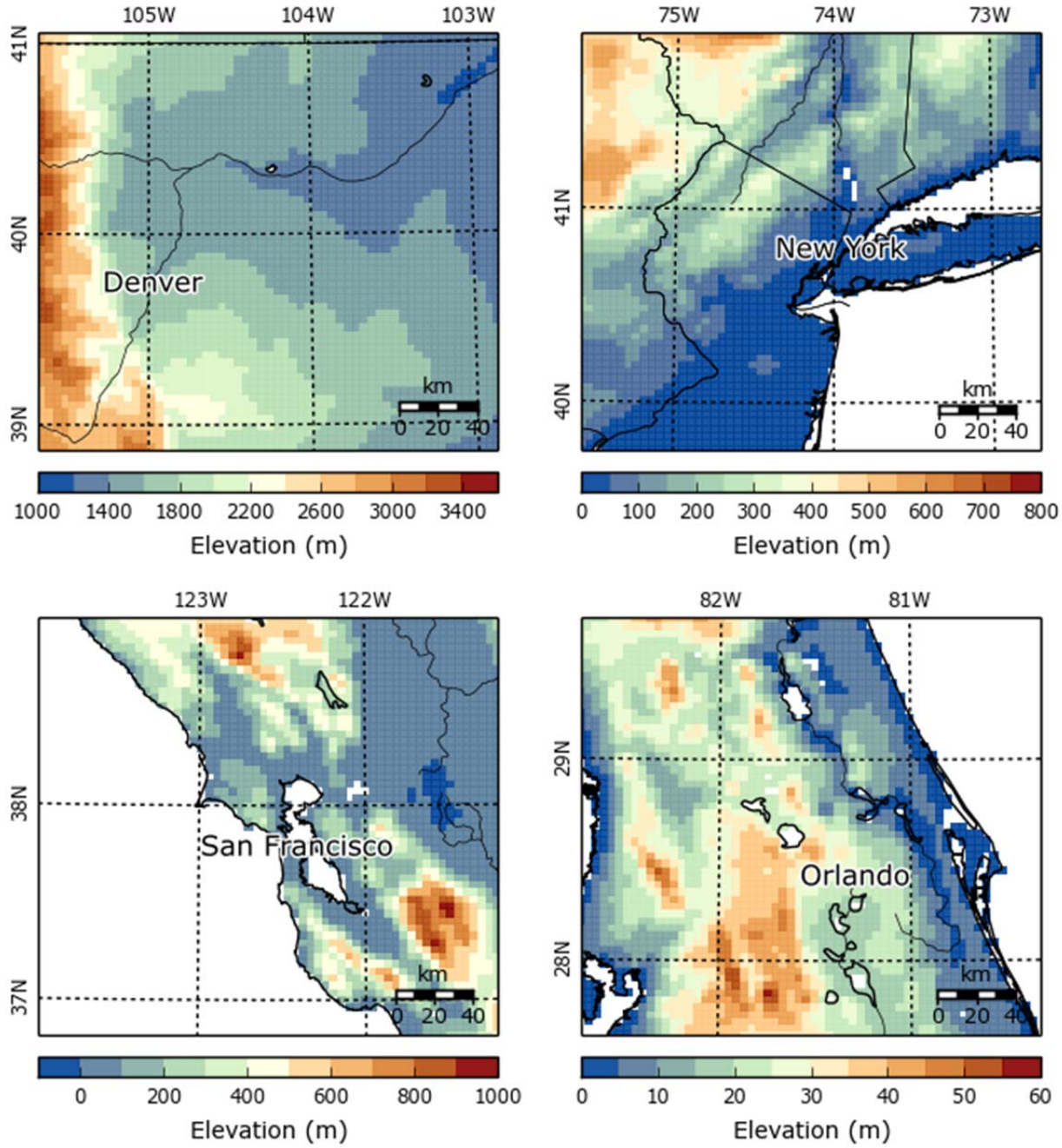

Figure 8 – Elevation differences for the four regions modeled

4.1 Datasets

A number of both gridded datasets and observational meteorological stations are used for comparative purposes:

- **NARR (2005-2012)**: Though this dataset was used as both initial and boundary conditions for temperature, humidity, and wind, the precipitation and solar radiation

derived in WRF is completely independent of NARR. In addition, the NARR dataset was used to evaluate the consistency of the modeling effort;

- **CFSR (2003-2010):** The Climate Forecast System Reanalysis, version 1, provides historical data over the period from 1979-2010. The more recent version 2 provides historical data from 2011 to the present and at a higher resolution but was not utilized;
- **PRISM (2005-2012):** PRISM is a downscaled, gridded dataset for average temperature and precipitation statistics only derived by interpolating observational stations and correcting for elevation, proximity to coastlines, and other parameters. The resolution is approximately 4 km (2.5 mi) over the continental United States;
- **NREL (1998-2009):** NREL provides gridded solar data at approximately a 10 km (6.2 mi) resolution based on the Perez model;
- **HOF (1986-2010):** The climatic data found in the HOF 2013 provides a long term comparison;
- **NSRDB (1991-2010**): The NSRDB database is derived from largely the same locations and data as the HOF yet also contains some quantities not currently found in the HOF including solar radiation and average wind speed; and,
- **MADIS (2005-2012)**: The MADIS Mesonet provide a further, independent source of all variables. While there are on the order of 100s of stations operating in each region, only approximately 10% of these stations had a sufficiently complete record over the 8 years required. However, the few stations that were sufficient were of excellent quality.

It is important to realize that the MADIS data comparisons with WRF output are most meaningful since they are the most direct in the same period of time with WRF hourly output data.

4.2 Qualitative and Quantitative Comparison

To evaluate the modeling effort against all statistics currently found in the HOF would be a monumental undertaking. Therefore the comparison focused on select statistics for each of the major variables of temperature, humidity, wind, precipitation, and solar radiation.

For simplicity, the closest grid point to each observational station was used without any sophisticated interpolation horizontally or any corrections in elevation based on e.g. lapse rate.

Appendix A contains a total of 85 figures that include a region-by-region, statistic-by-station comparison including scatter plots of modeled versus observational data and contour plots of all the datasets.

Here, in the main body of the report, we include only the summary statistics and observations. The statistics provided are similar to what was presented in the section on weather evaluation. That is, for each observational data station or grid point time series, we calculate the summary statistic (e.g. average January dry-bulb temperature) and pair that with the summary statistic calculated for the closest available modeled point from WRF. Then, for all available pairs in the region, we calculate the mean bias error, the root mean square error, and the correlation coefficient. To provide an estimate of the robustness of these calculations, we also provide error

bars of the bootstrapped estimate of the 90% confidence intervals for these quantities.

4.3 Temperature

In terms of temperature, we provide an estimate of the representation of the average annual dry-bulb temperature (see Figure 9), the average January dry-bulb temperature (see Figure 10), and the average July dry-bulb temperature (Figure 11). In all regions except for Denver, the errors are typically within 1 °C (1.8 °F) and well-correlated. There is a larger uncertainty to the Denver results due to a fewer number of available stations with sufficient data. Extreme values of temperature were also computed representing the 1% cooling and 99% heating dry-bulb temperatures (see Figure 12 and Figure 13). Again, errors are typically within 1 °C (1.8 °F) and well-correlated.

4.4 Humidity

Extreme values for humidity were generated for the 1% dehumidification humidity and 99% humidification ratios (see Figure 14 and Figure 15). In both cases humidity values are poorly to semi-correlated with biases within approximately 0.5 g/kg (3.5 grain/lb). Again, Denver is particularly poorly represented.

4.5 Wind

Average annual, January, and July wind speeds were calculated (see Figure 16, Figure 17, and Figure 18, respectively). While the NSRDB stations are semi-correlated to the WRF projections, the MADIS results are quite poor, agreeing with the previous "weather" evaluation results. The annual 1% wind speed (see Figure 19) exhibits similar errors as the average wind conditions.

4.6 Precipitation

Precipitation was calculated for annual, January, and July average conditions (see Figure 20, Figure 21, and Figure 22, respectively). Correlation is consistently poor with biases of approximately ±50%. Precipitation is of approximately the right magnitude but with poor spatial accuracy.

4.7 Solar Radiation

Finally, solar radiation was calculated for annual, January, and July average daily global horizontal radiation received at the surface (see Figure 23, Figure 24, and Figure 25, respectively). Solar radiation is typically poorly to semi-correlated and consistently overestimated by approximately 5-15%.

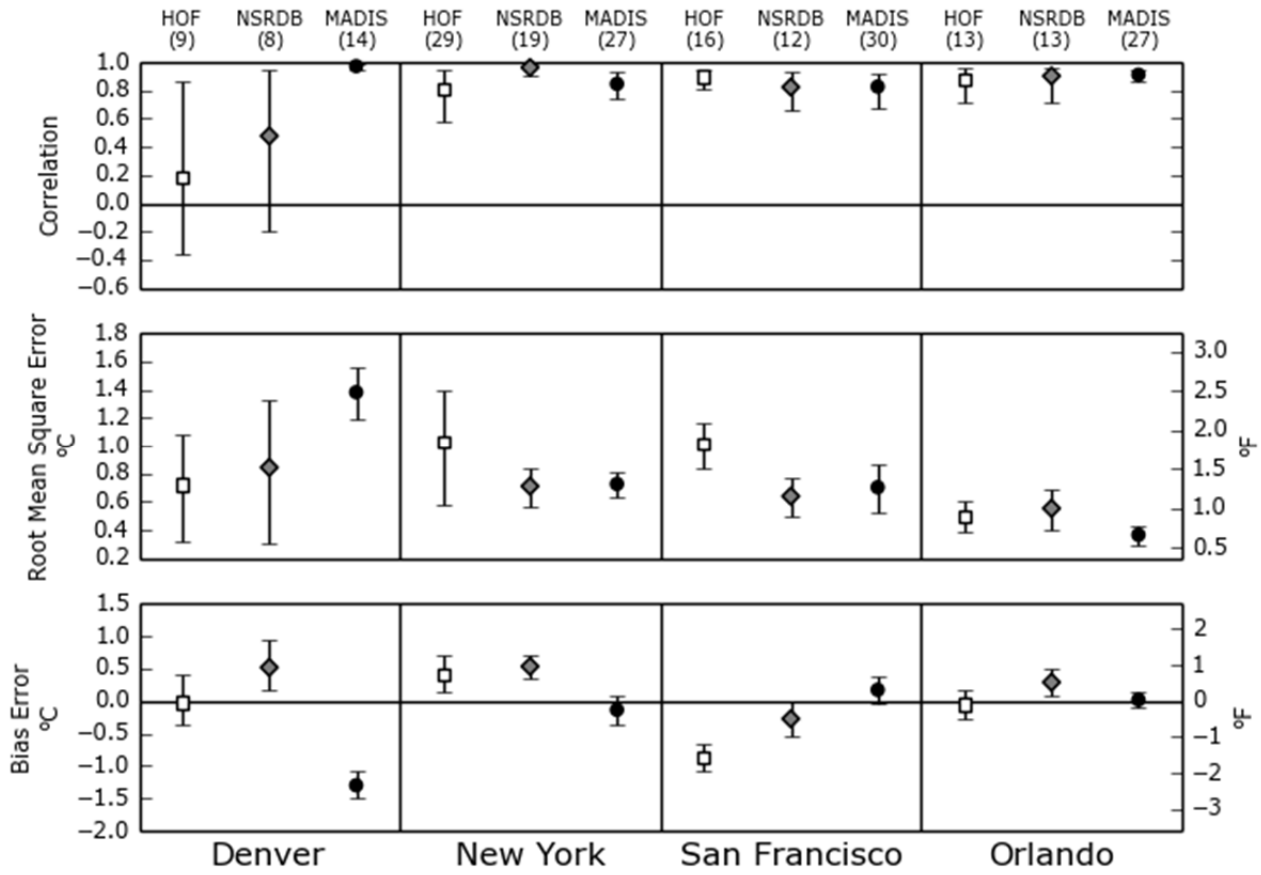

Figure 9 – Annual Average Dry-bulb Temperature. Numbers in parenthesis indicate the number of stations utilized in the statistic

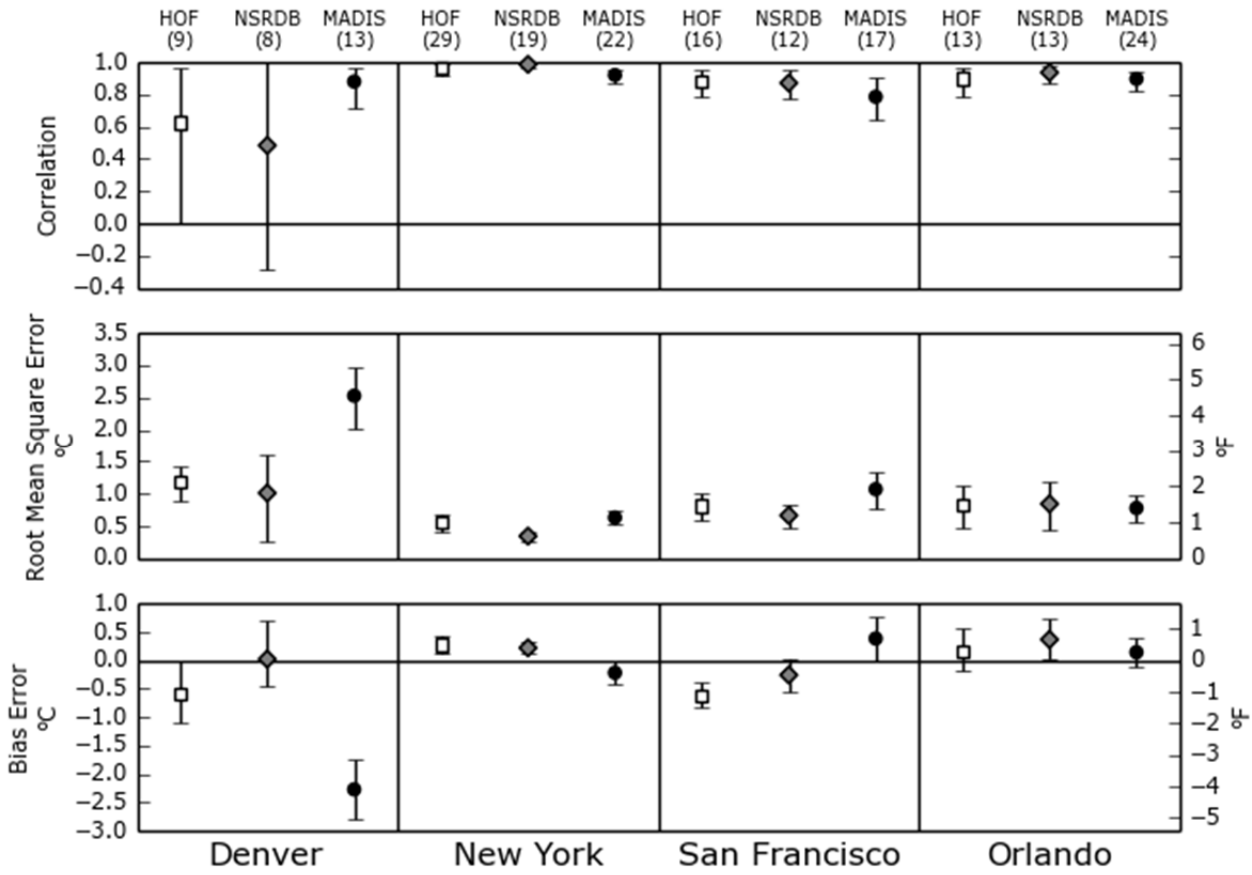

Figure 10 – Average January Dry-bulb Temperature

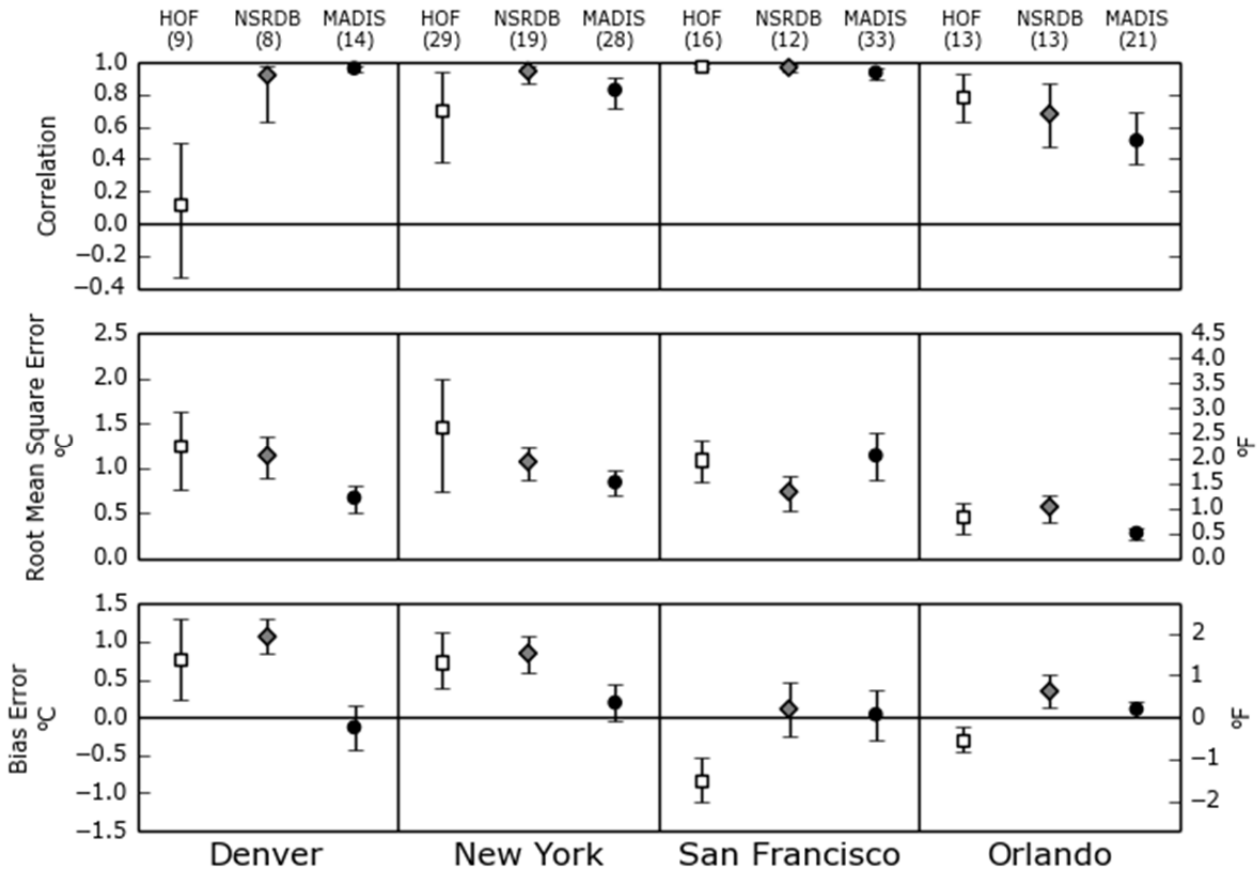

Figure 11 – Average July Dry-bulb Temperature

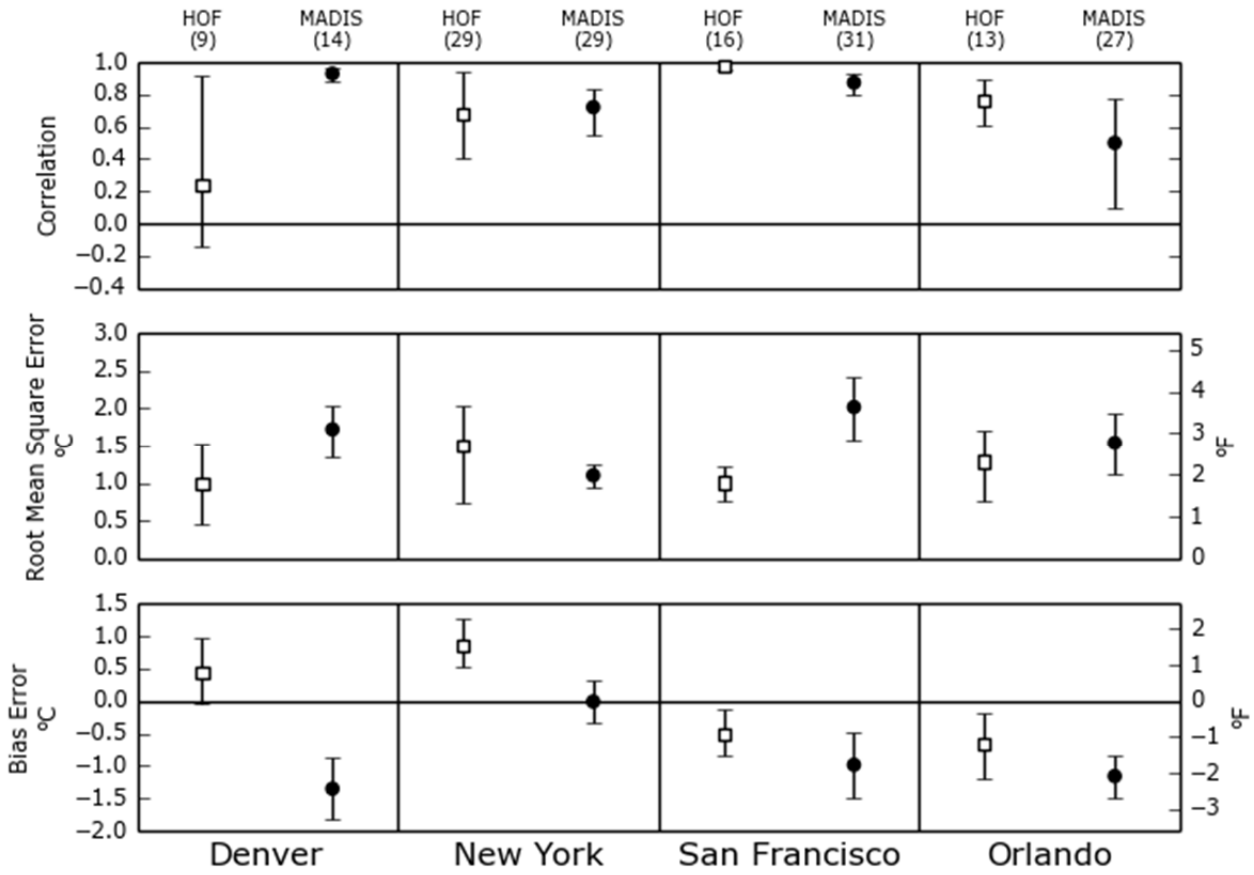

Figure 12 – 1% Cooling Dry-bulb Temperature

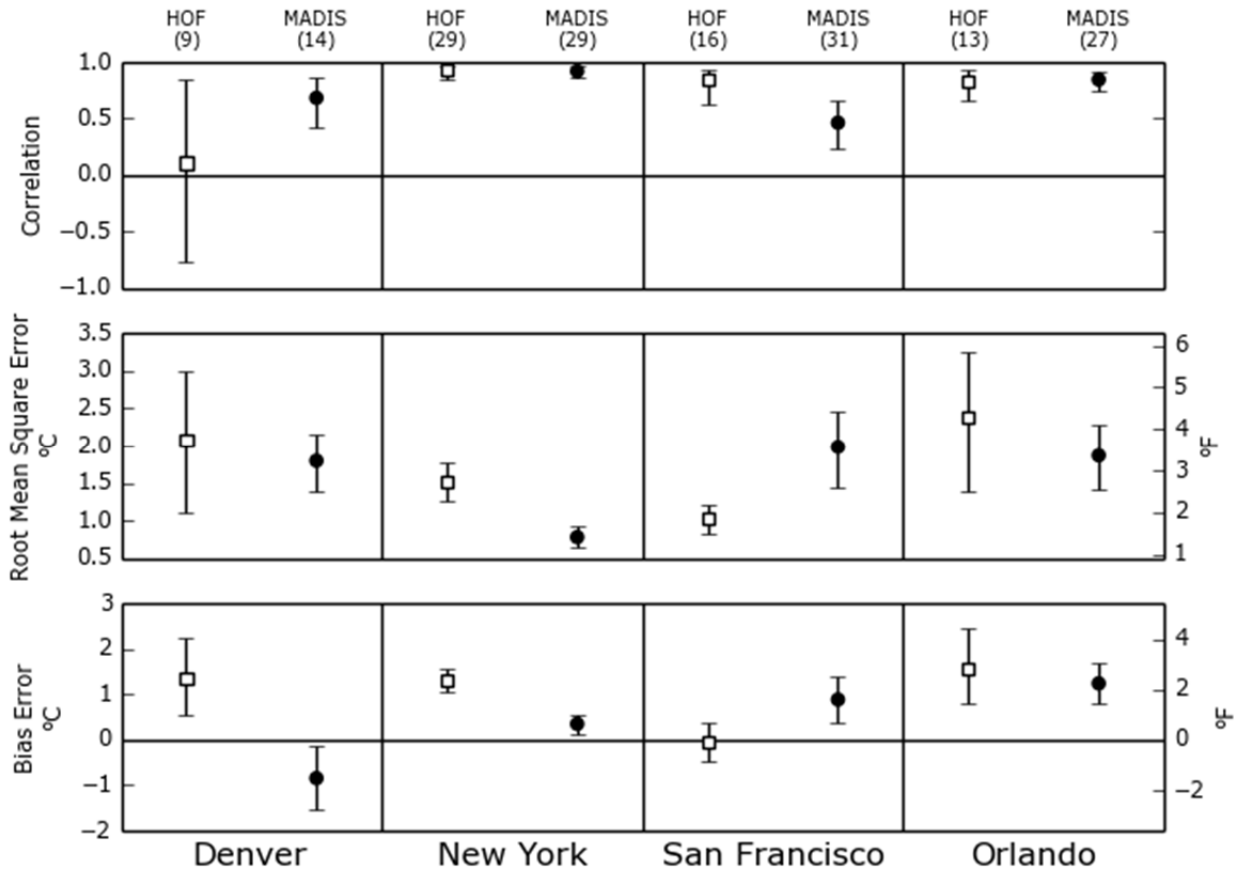

Figure 13 – 99% Heating Dry-bulb Temperature

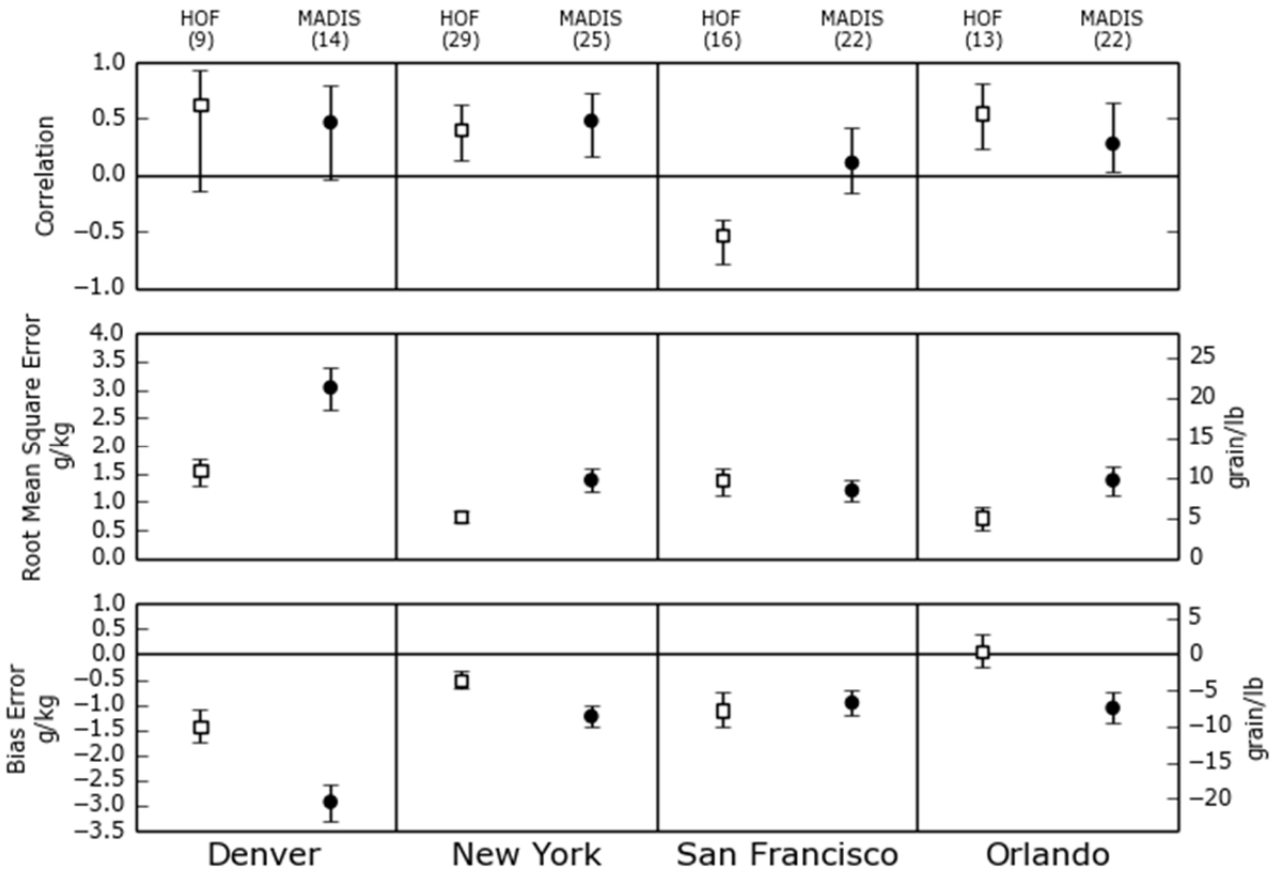

Figure 14 – 1% Dehumidification Humidity Ratio

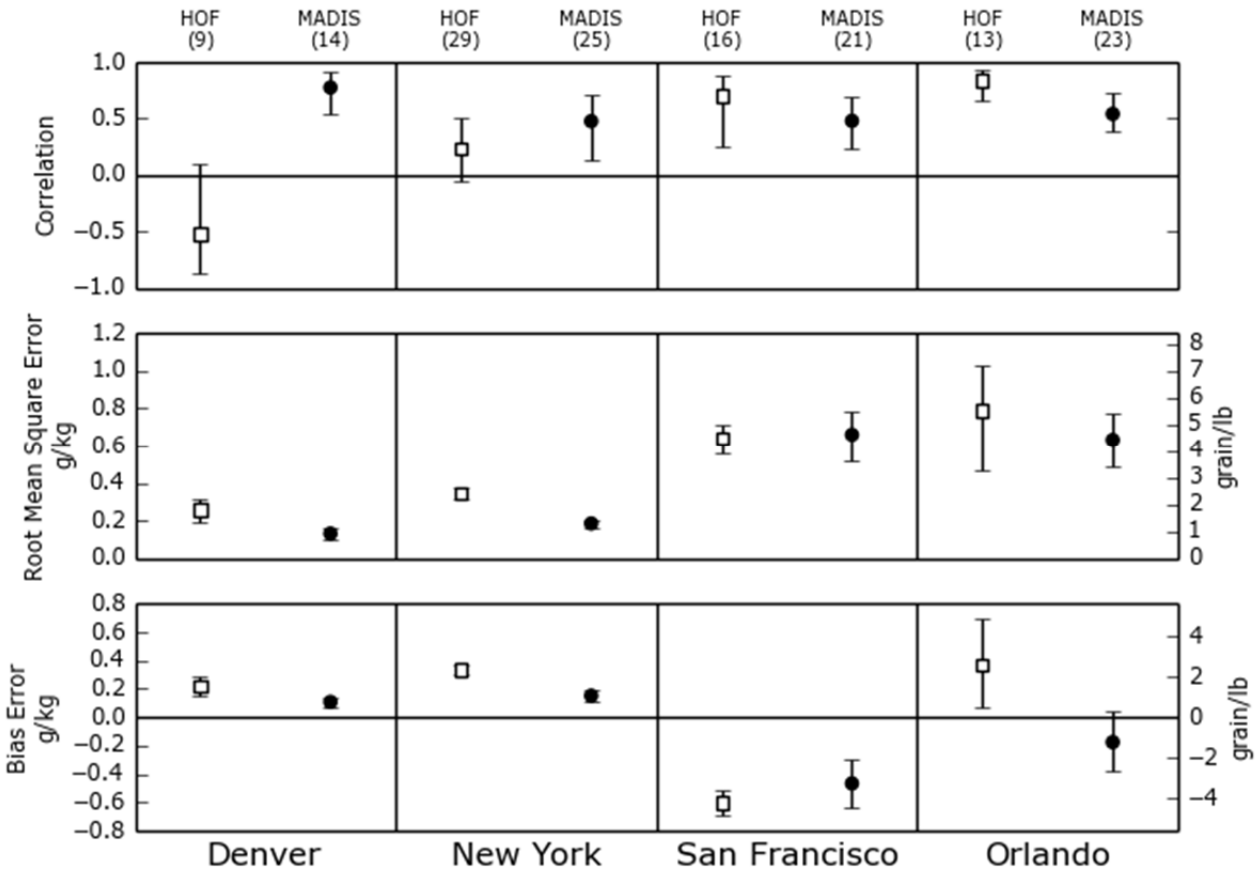

Figure 15 – 99% Humidification Humidity Ratio

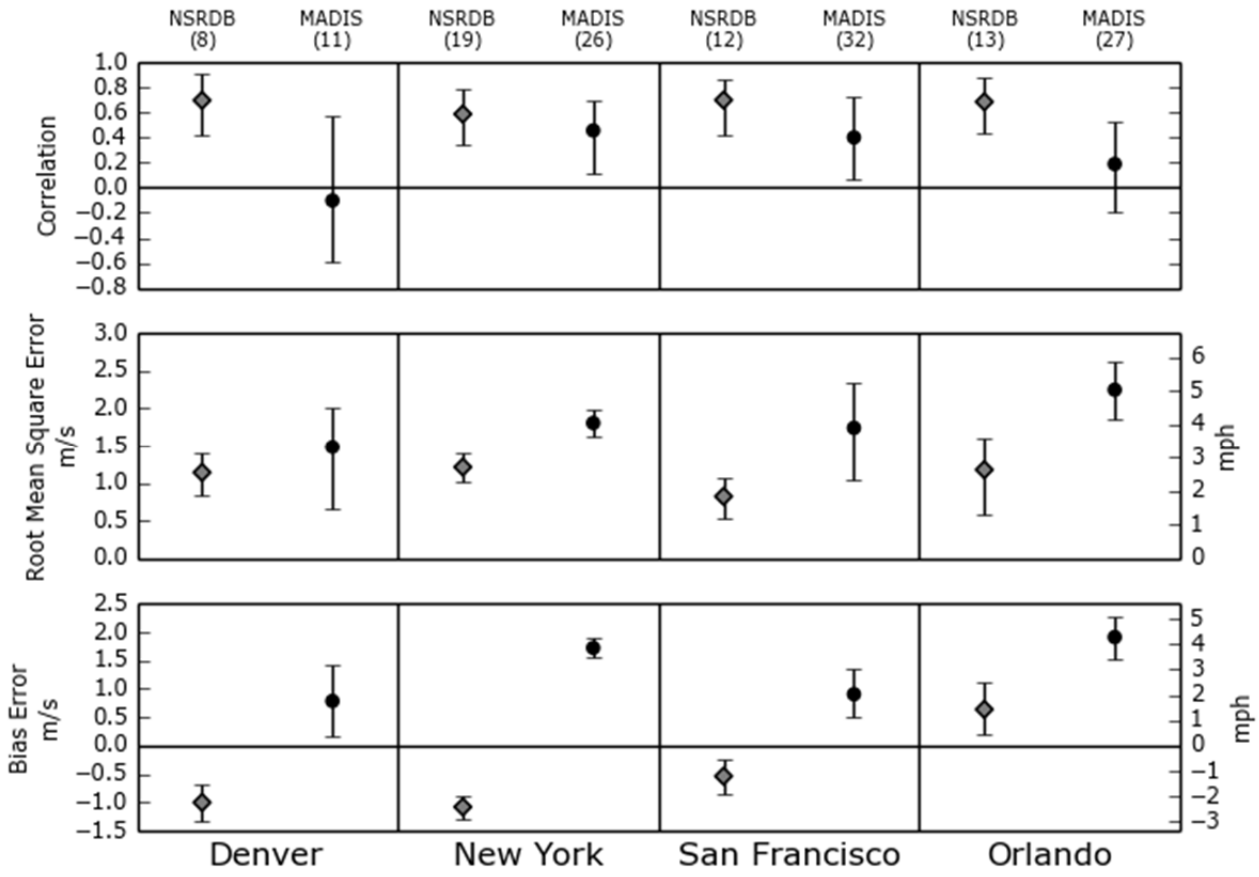

Figure 16 – Annual Average Wind Speed

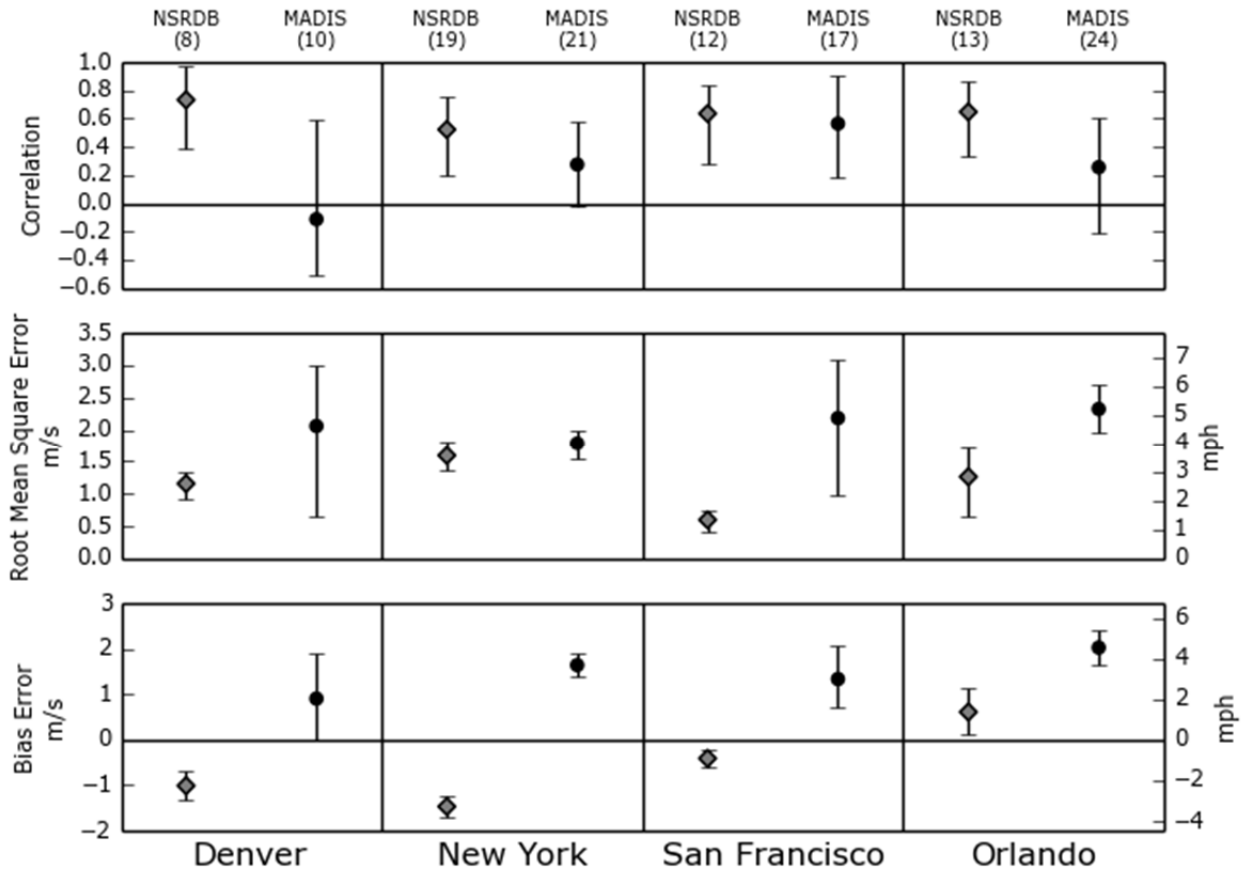

Figure 17 – January Average Wind Speed

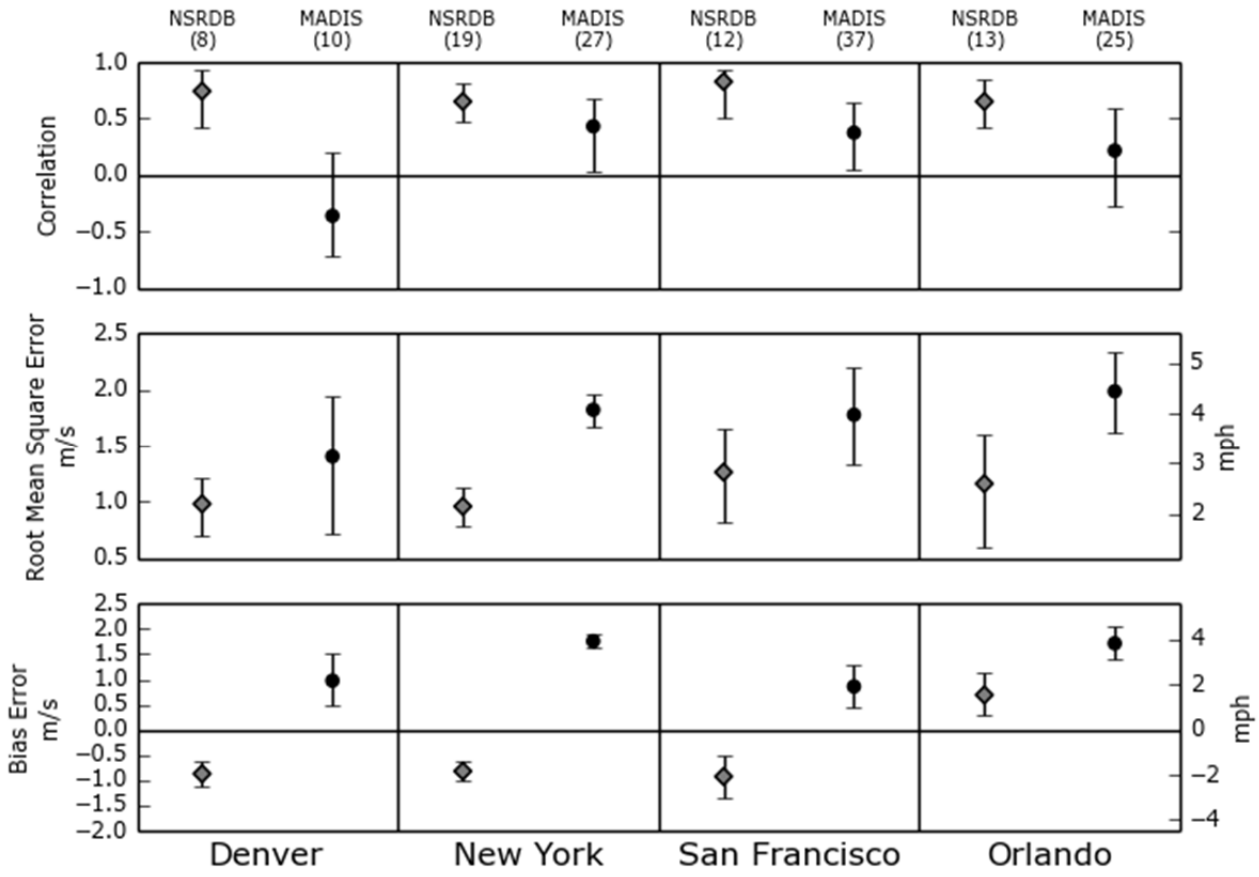

Figure 18 – July Average Wind Speed

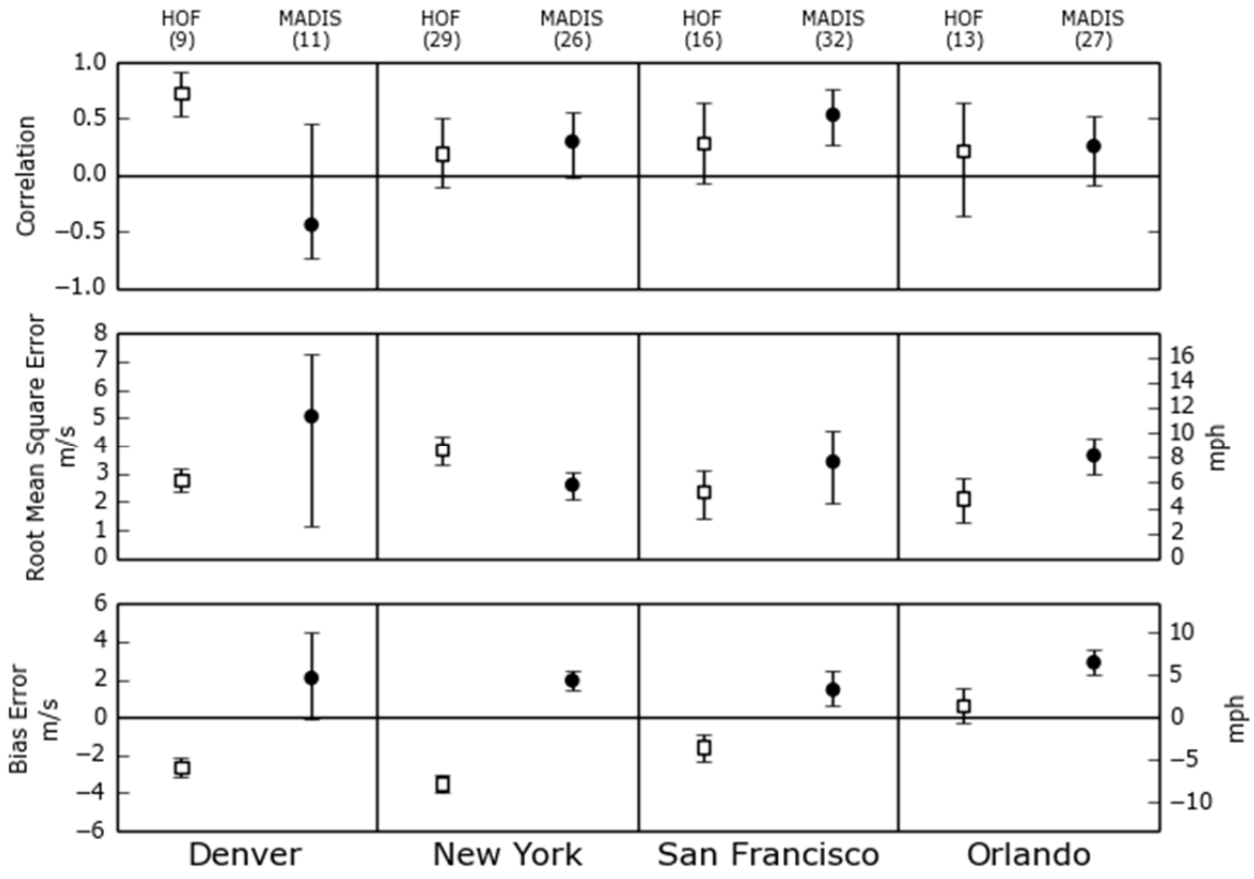

Figure 19 – Annual 1% Wind Speed

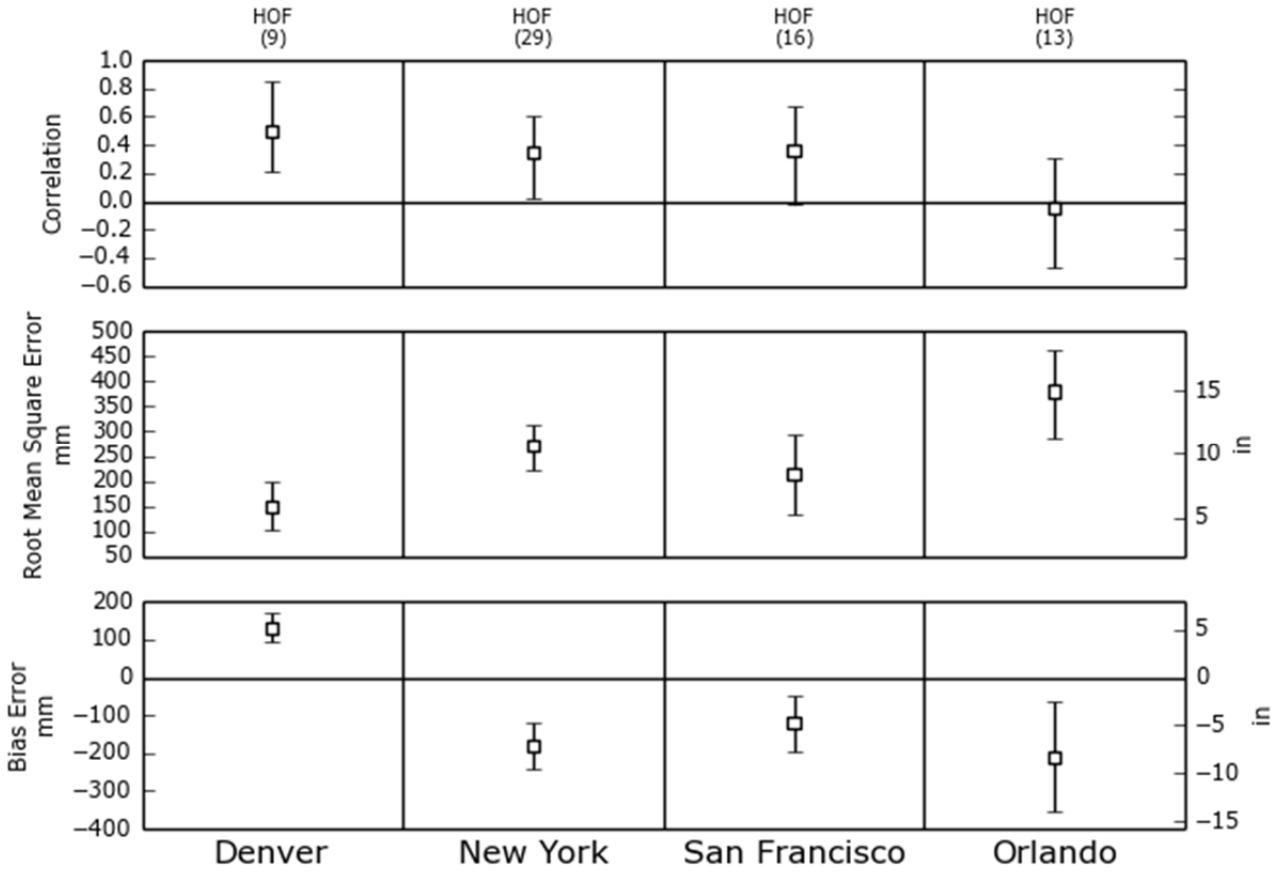

Figure 20 – Average Annual Precipitation

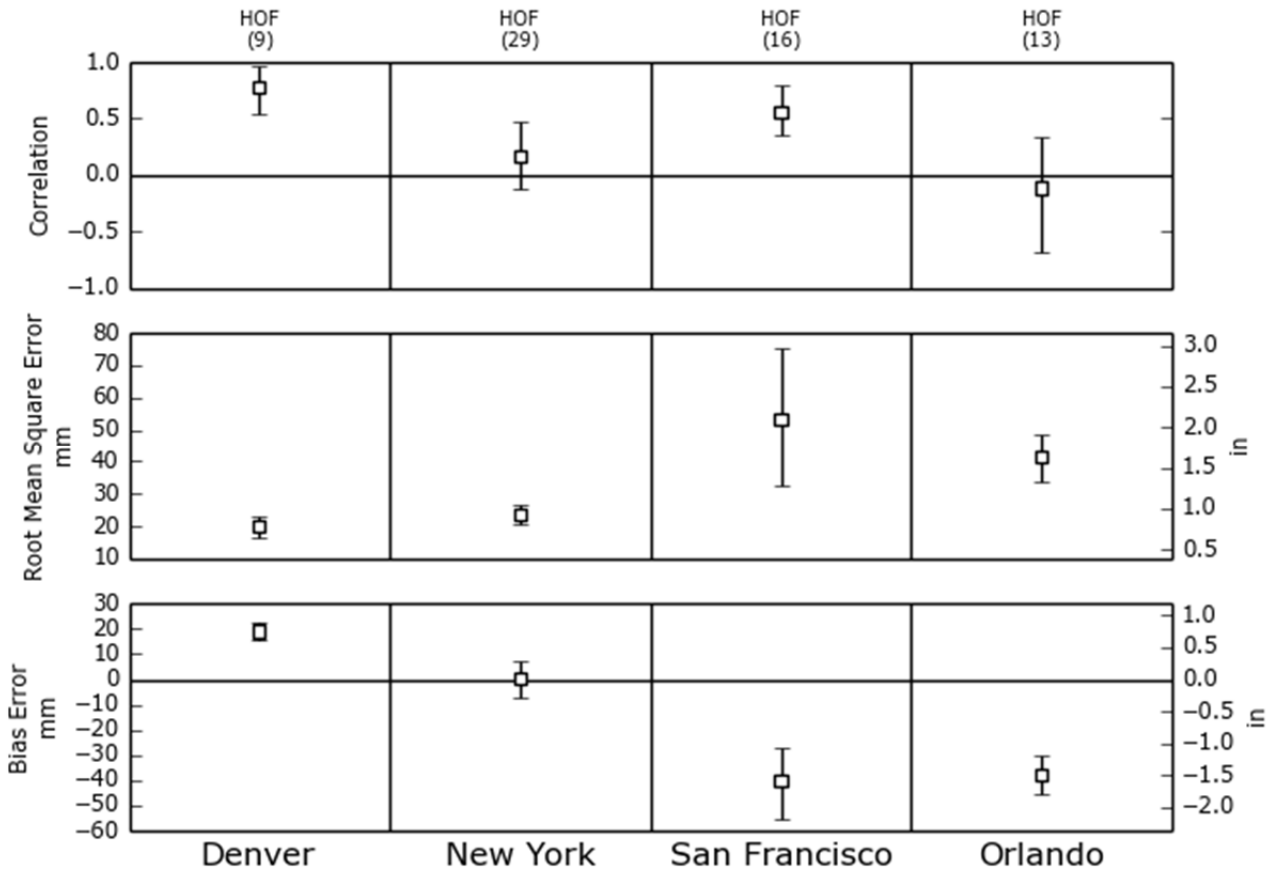

Figure 21 – Average January Precipitation

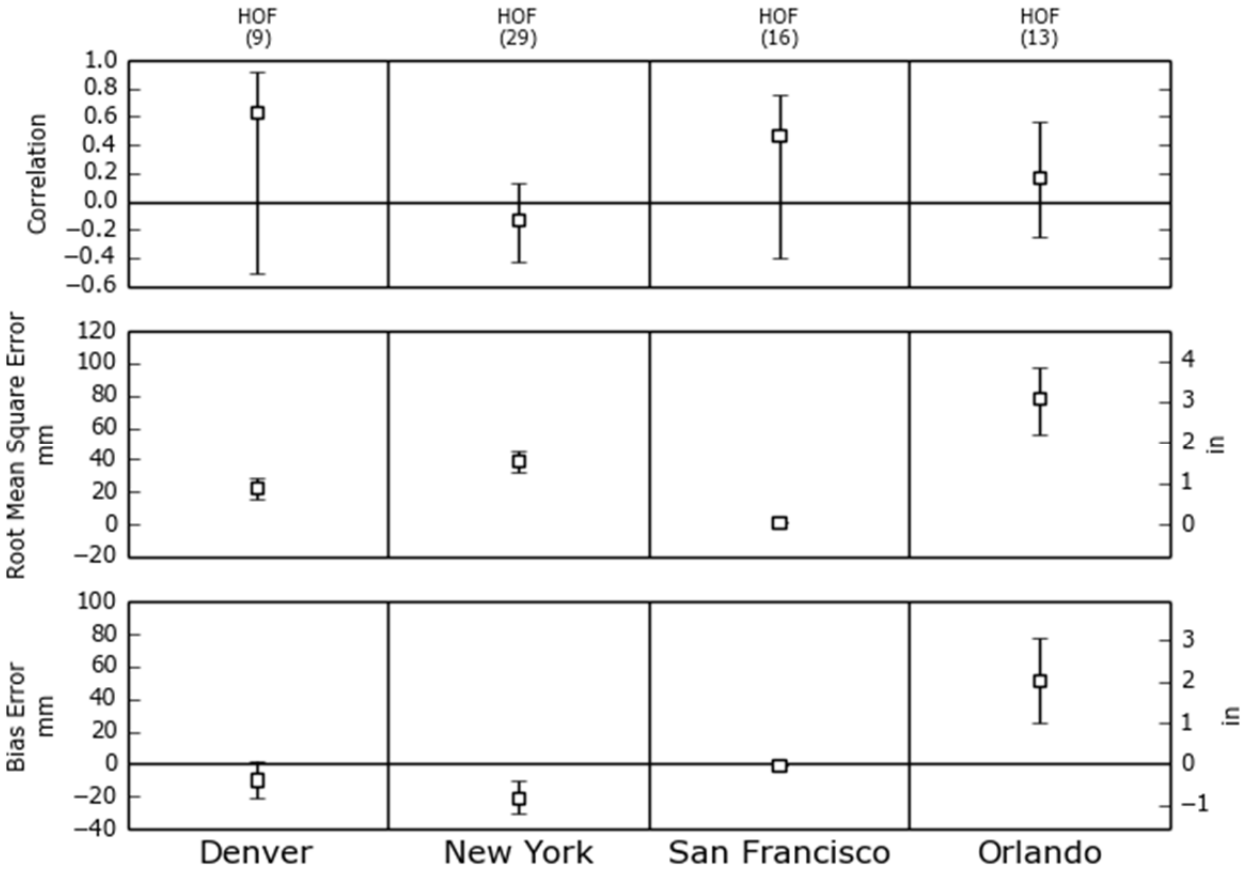

Figure 22 – Average July Precipitation

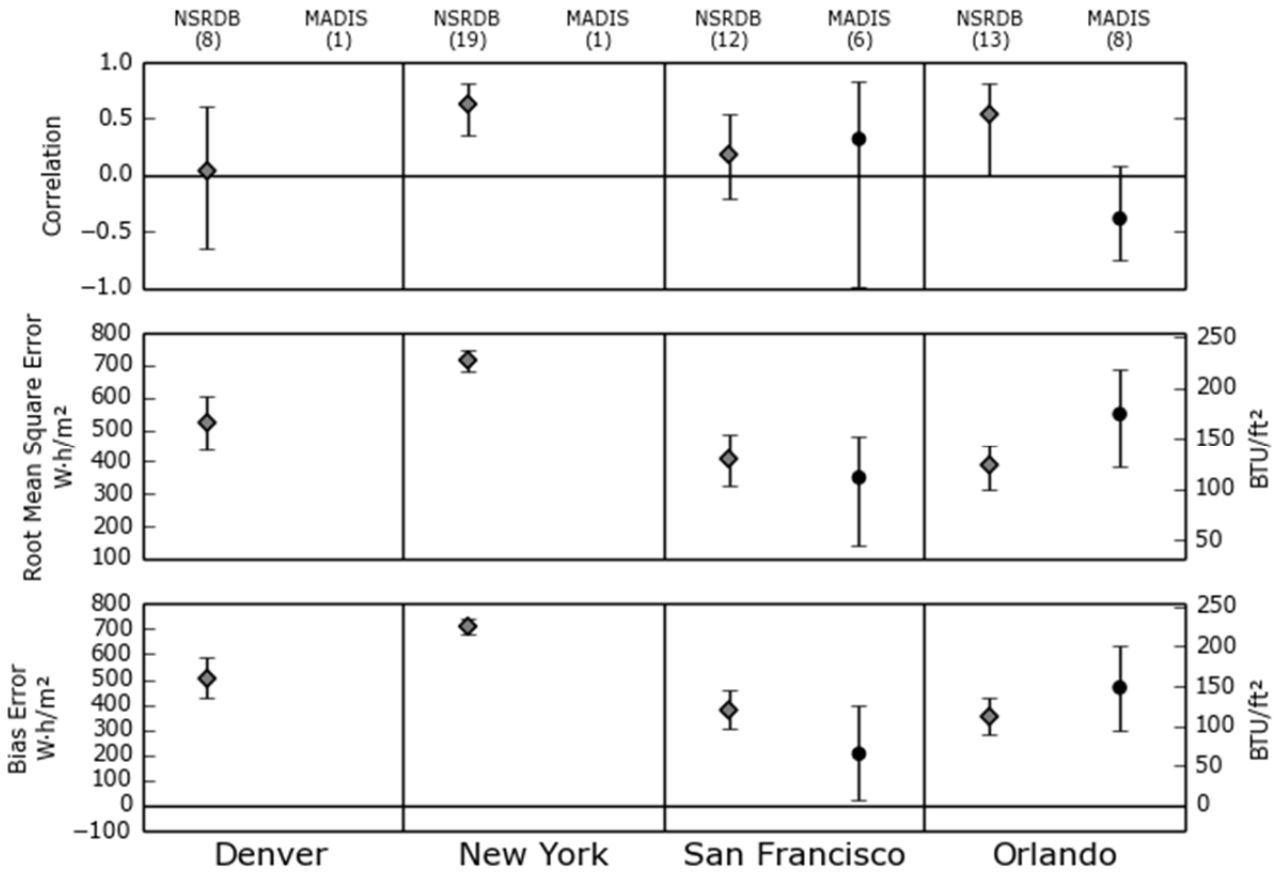

Figure 23 – Average Annual Daily Solar Radiation

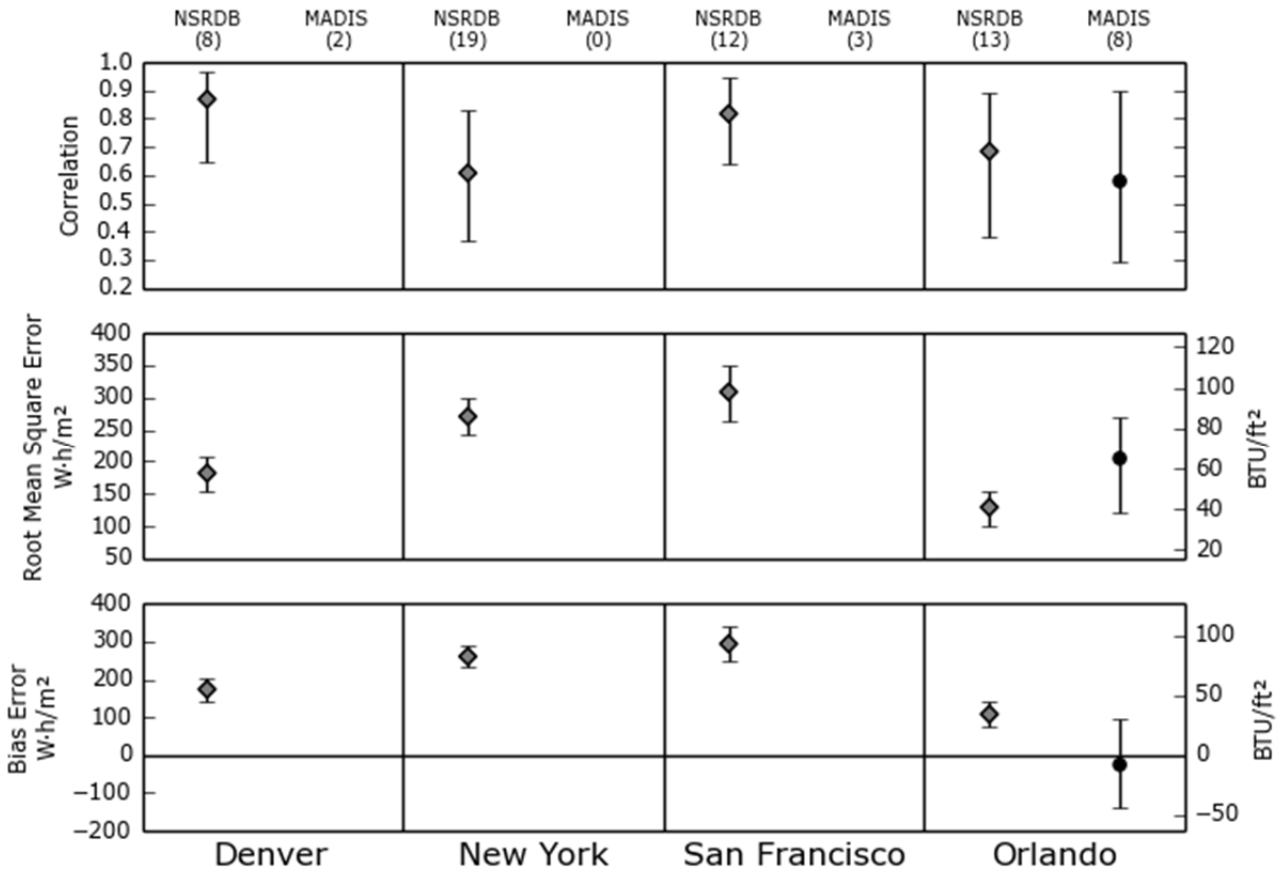

Figure 24 – Average January Daily Solar Radiation

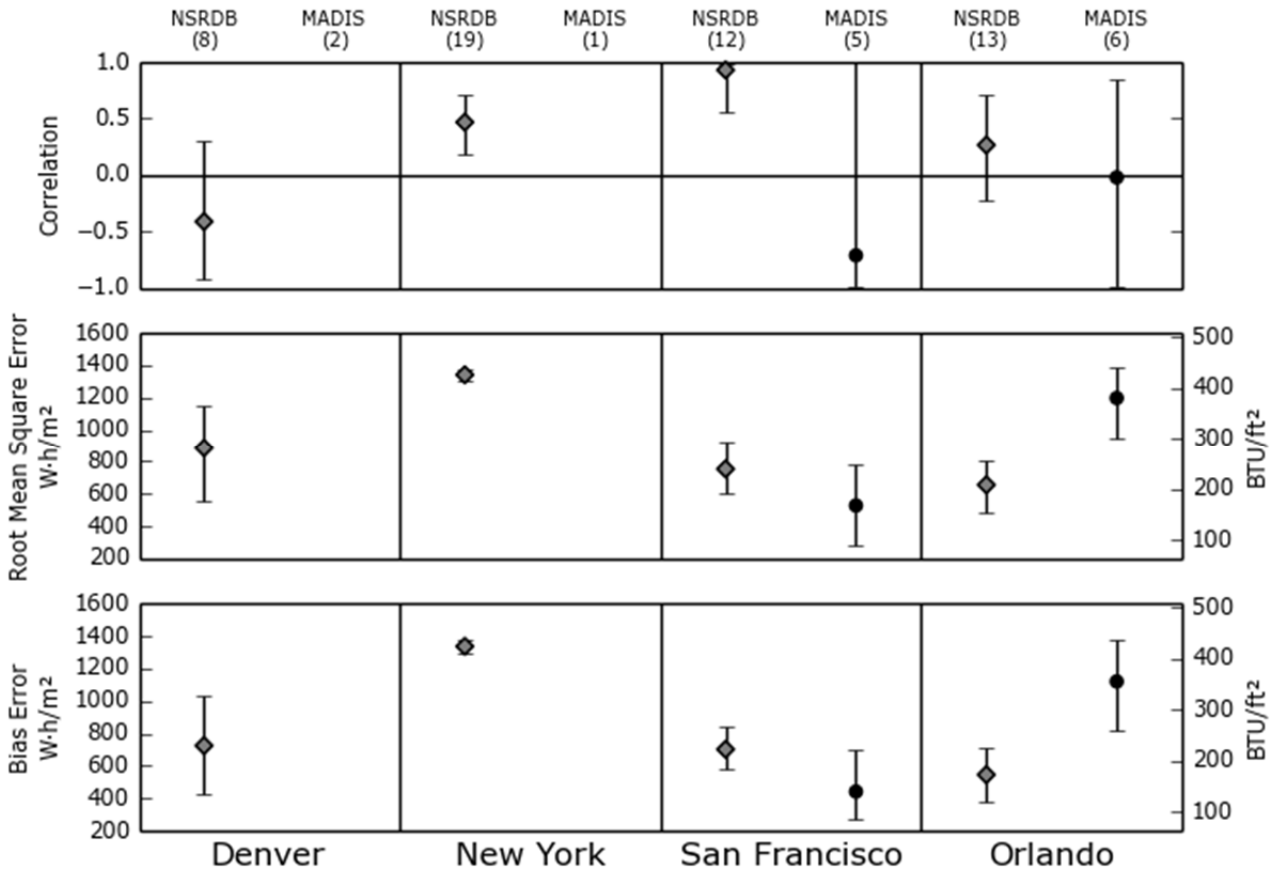

Figure 25 – Average July Daily Solar Radiation

5 Simplified (WRF EMS) versus Standard (Novus WRF) Methodology

The preceding sections evaluated the Novus WRF methodology against observational data. However, this methodology is extremely difficult to setup without advanced meteorological knowledge and significant computational resources. This method can be prohibitively difficult for new users and therefore we are recommending the use of a pre-compiled, simplified modeling package called WRF EMS. We thus have produced a methodology to implement WRF EMS which is detailed online[5]. To evaluate the difference between the methodologies, modeling was performed for New York City for 2005.

5.1 The Simplified Method: WRF-EMS Based WRF Modeling System

The NOAA/NWS Science and Training Resource Center (STRC) Environmental Modeling System (EMS) is a complete, full-physics, state-of-the-science numerical weather prediction

[5] Simplified WRF EMS methodology: http://klimaat.ca/emspy

(NWP) package that incorporates dynamical cores from both the National Center for Atmospheric Research (NCAR) Advanced Research WRF (ARW) and the National Center for Environmental Predictions' (NCEP) non-hydrostatic mesoscale model (NMM) releases into a single end-to-end forecasting system. All the capability of the NCEP and NCAR WRF models are retained within the EMS; however, the installation, configuration, and execution of each core has been greatly simplified to encourage their use throughout the operational, private, and university forecasting and research communities.

The project team for RP-1561 has developed a set of user friendly scripts to support potential ASHRAE members perform modeling with the WRF EMS system.

There are some differences between the WRF EMS Methodology and the Novus WRF Methodology. Table 7 summarizes the key differences. Note that although observational nudging is not available in the current version of WRF EMS, most of the differences will not significantly affect model functions as both models use the same dynamic and physics cores.

Table 7 – Differences between WRF Methods

Novus WRF	WRF EMS
Grid, Spectral, and Observational Nudging	Grid & Spectral Nudging Only
User-controlled Physics Options	Subset of full Physics Options available
User supplies input data	Model retrieves input data from online repository as needed
User prepares all scripts for all stages of modeling	EMS setup template scripts for user to use
User has full control of model run settings	User has more limited control of model run settings
Can be run on large scale cluster or supercomputer	Targets local multi-core PC or workstations
Large community and forum based support	Smaller community and forum based support

In order to evaluate the WRF EMS package, a comparison was made between the Novus WRF model, WRF EMS, and observation data. The following sections present the results of this comparison.

5.2 Novus WRF vs. WRF EMS

The two WRF modeling approaches were compared using the same methods used to validate the WRF model for ASHRAE. Given the limits of computational time, the evaluation focused on the three main airports around New York City (JFK, LaGuardia, and Newark). Figure 26, Figure 27, and Figure 28 show the results of this comparison for temperature, wind direction, and wind speed, respectively.

It can be seen that for temperature, both models perform similarly, with Novus WRF performing slightly better. For wind direction and wind speed, it can be seen that the WRF EMS performs to the same standard as the Novus WRF model.

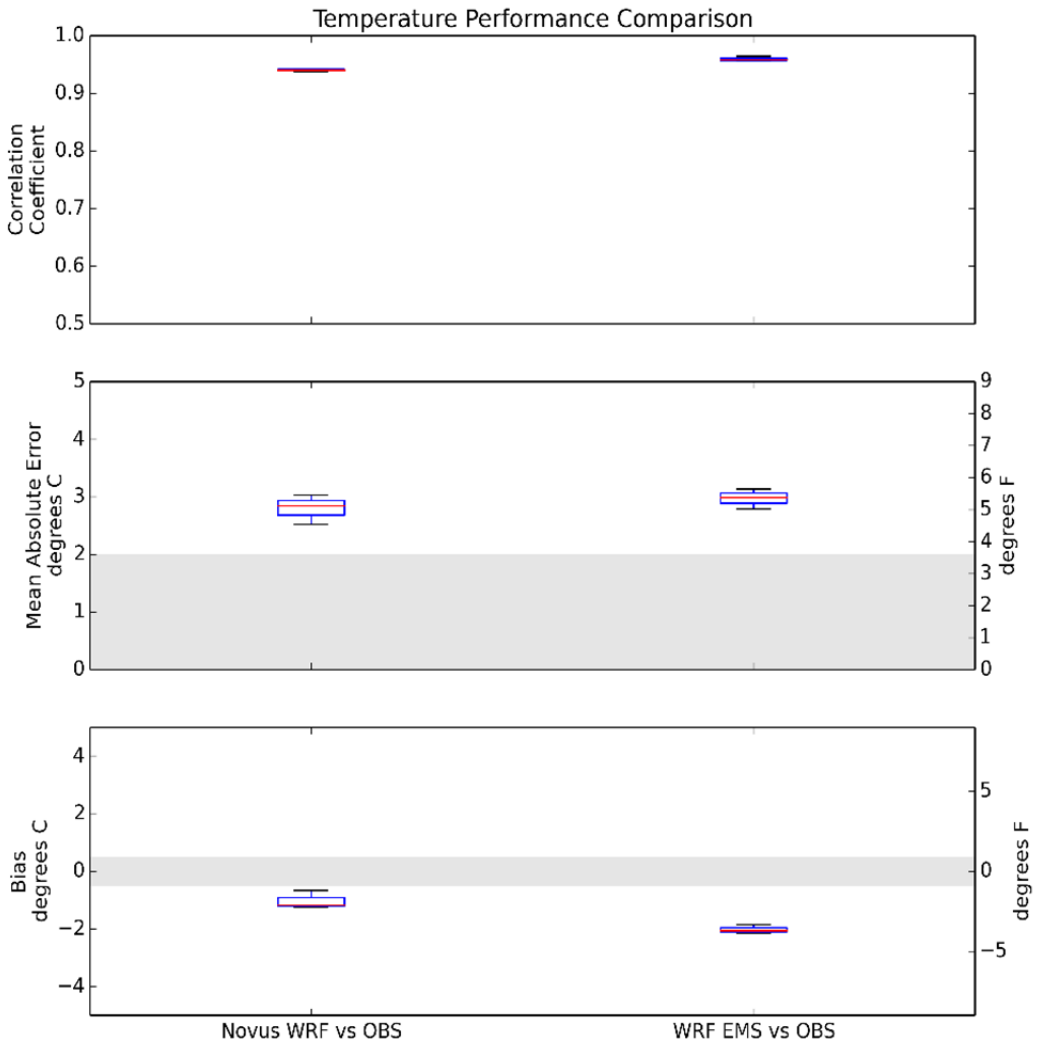

Figure 26 – Novus WRF and WRF EMS vs Observed Temperature

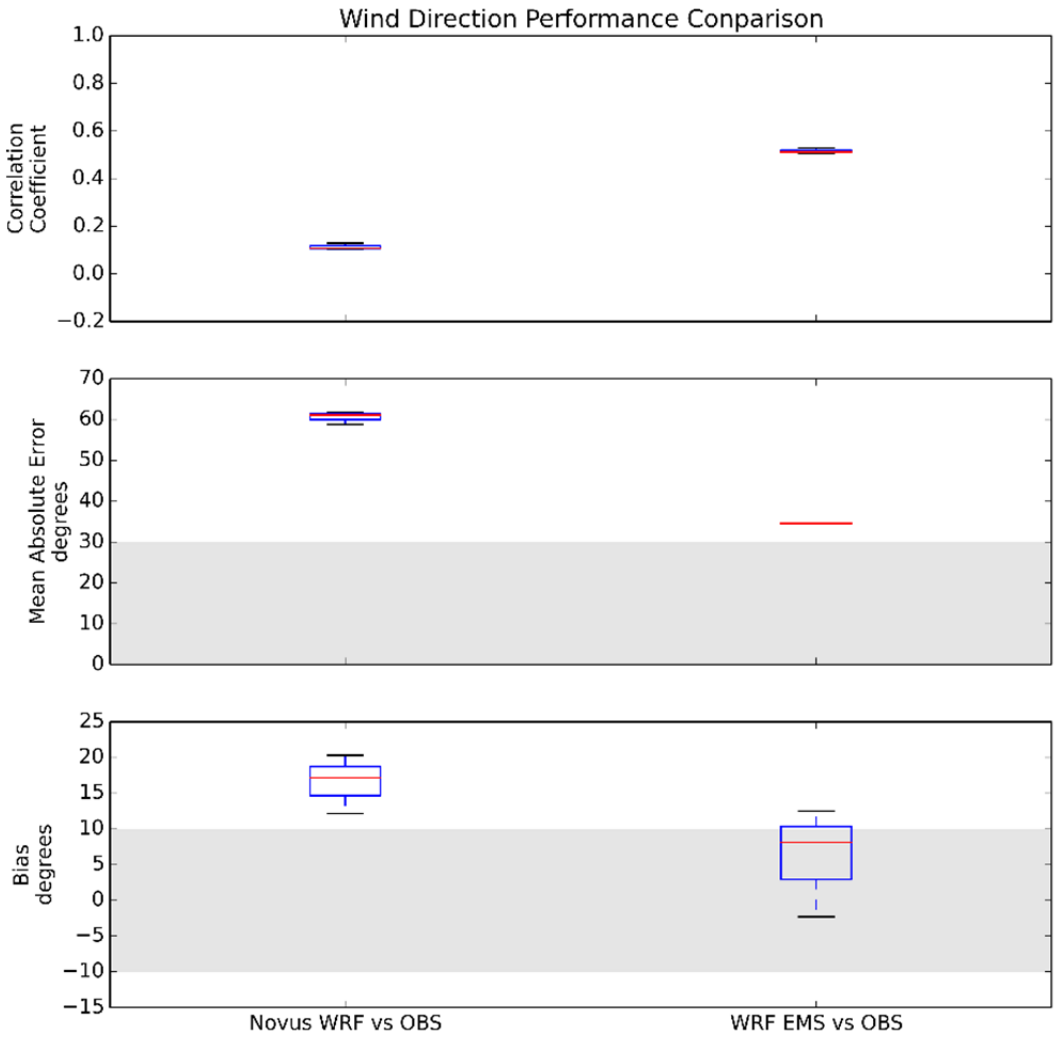

Figure 27 – Novus WRF and WRF EMS vs Observed Wind Direction

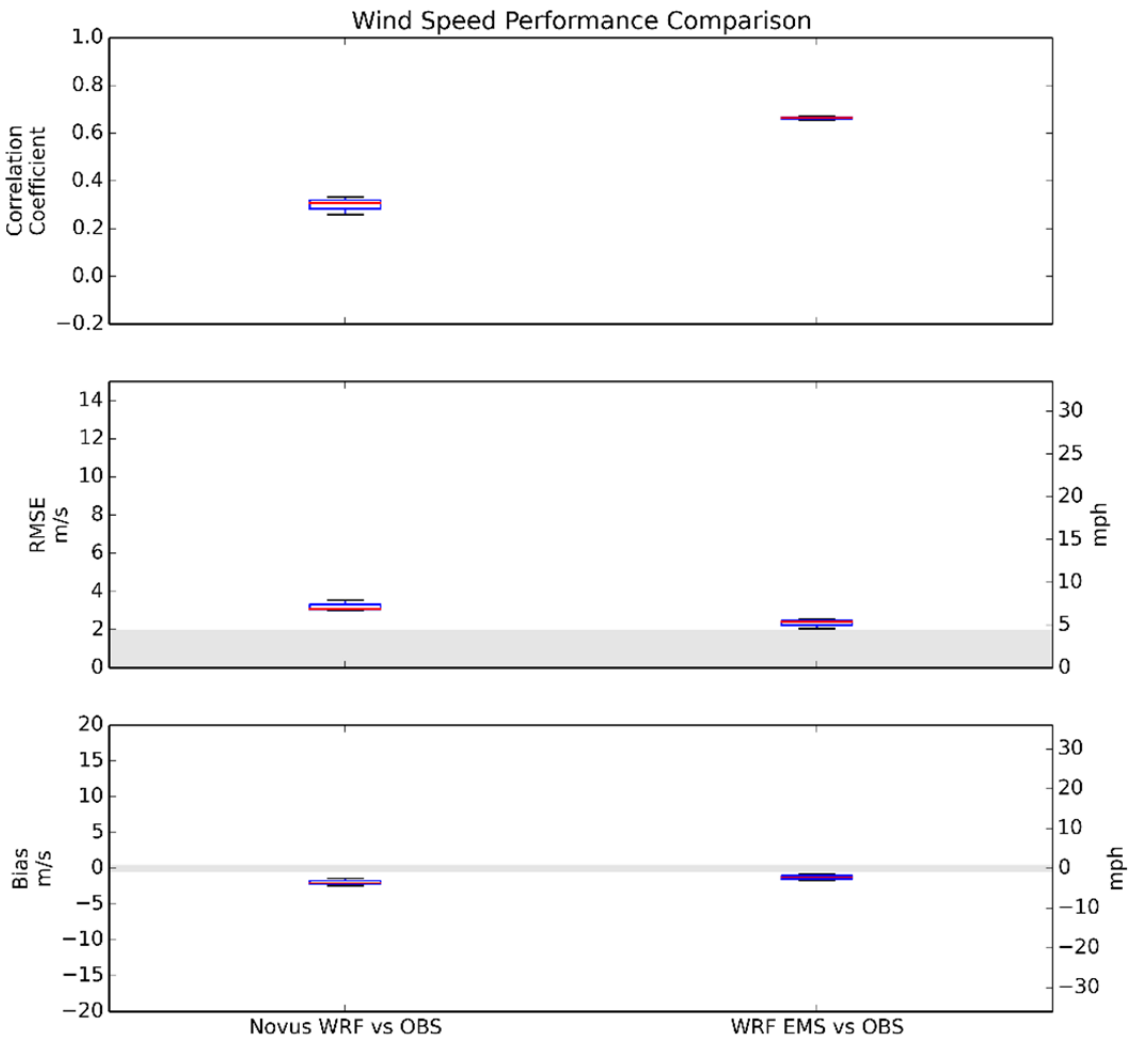

Figure 28 – Novus WRF and WRF EMS vs Observed Wind Speed

Figure 29, Figure 30, and Figure 31 show time-series of the average daily temperature at each of the three airports for both models as well as observation data.

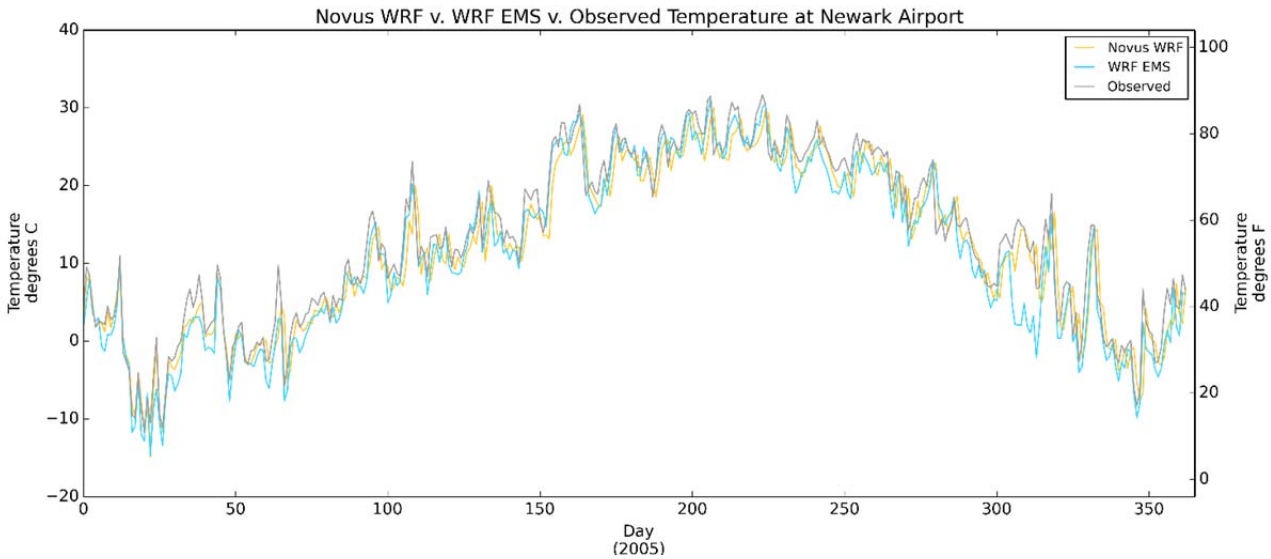

Figure 29 – Modeled (Novus WRF and WRF EMS) Temperatures against Measured Temperature at Newark Airport

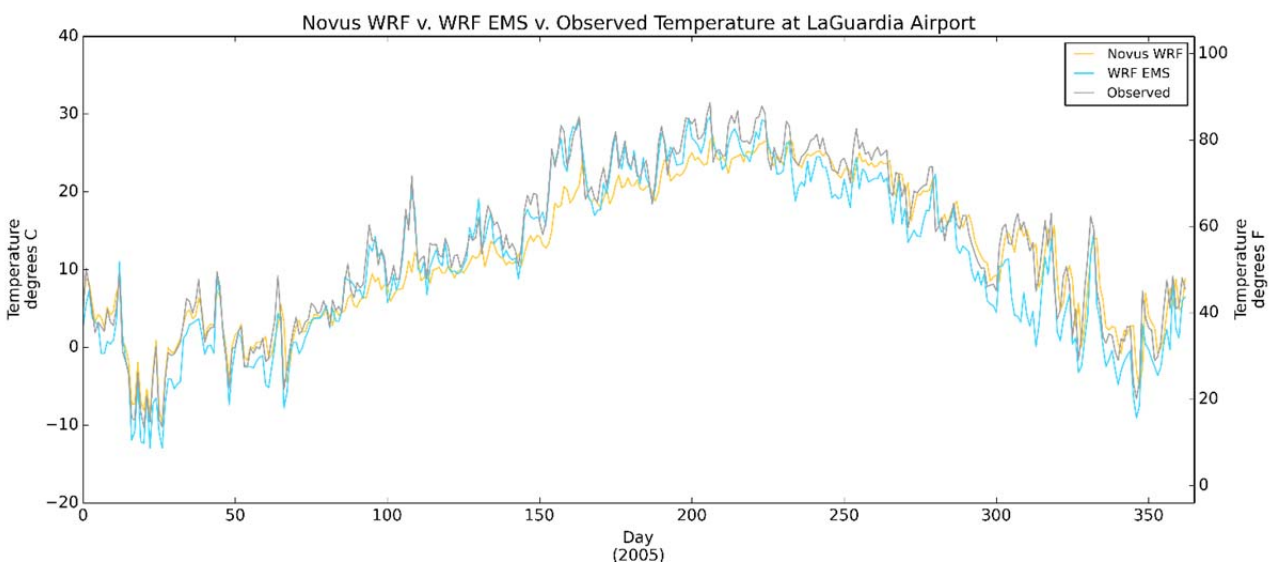

Figure 30 – Modeled (Novus WRF and WRF EMS) Temperatures against Measured Temperature at LaGuardia Airport

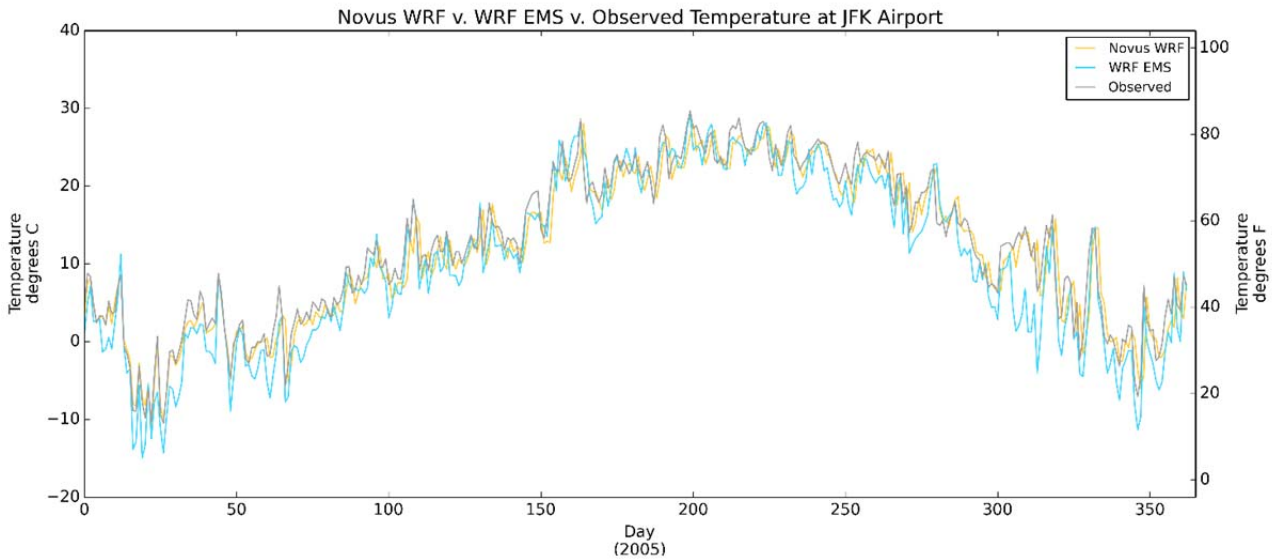

Figure 31 – Modeled (Novus WRF and WRF EMS) Temperatures against Measured Temperature at JFK Airport

Figure 29 and Figure 30 show that both models slightly under predict the observed values at Newark and LaGuardia. Figure 32, Figure 33, and Figure 34 further compare temperatures at the three monitoring locations.

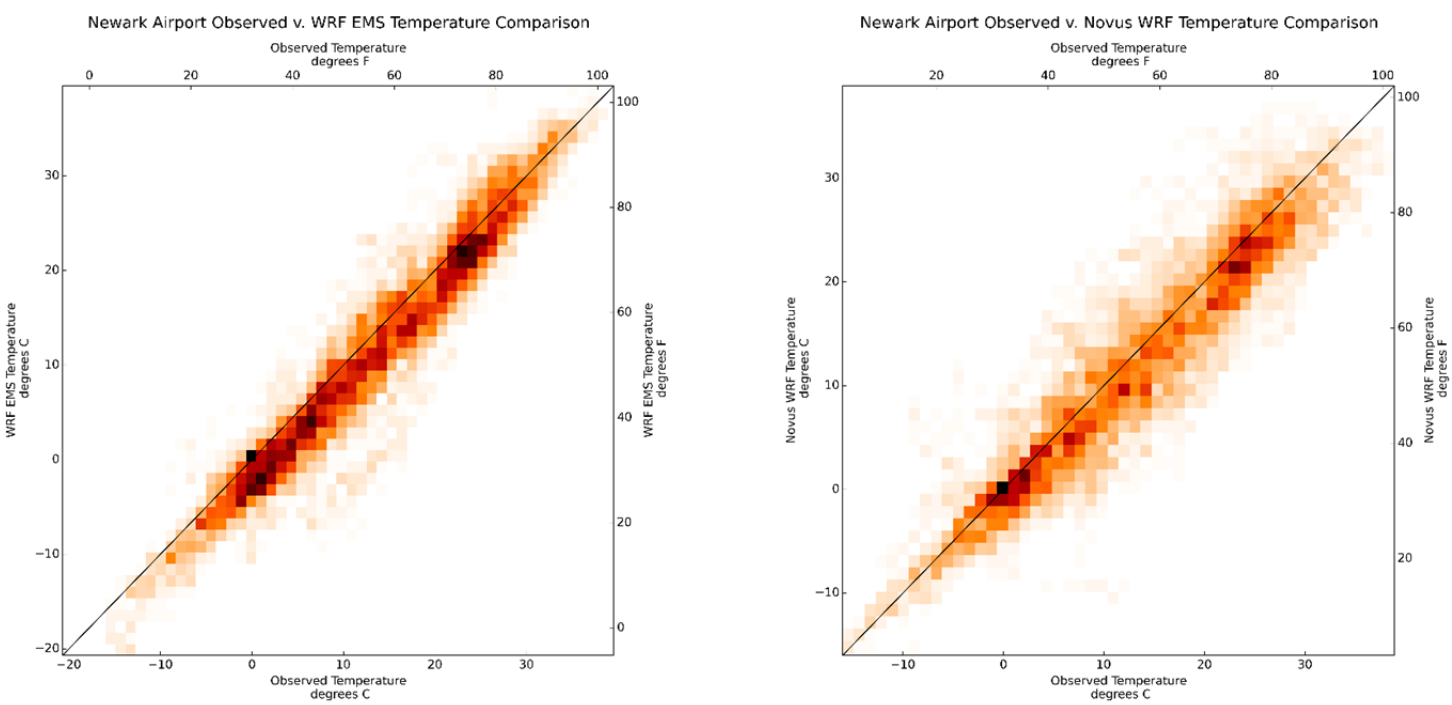

Figure 32 – Novus WRF & WRF EMS vs Newark Airport Observed Temperature

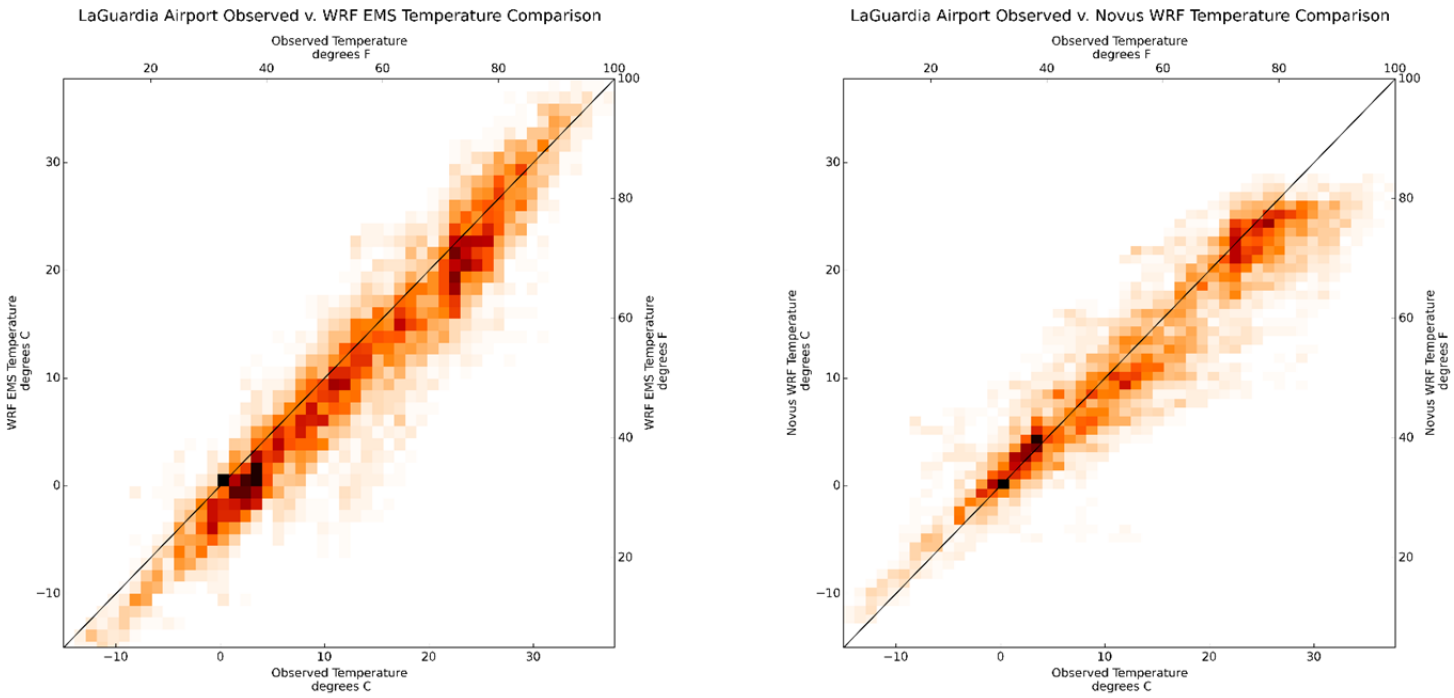

Figure 33 – Novus WRF & WRF EMS vs LaGuardia Airport Observed Temperature

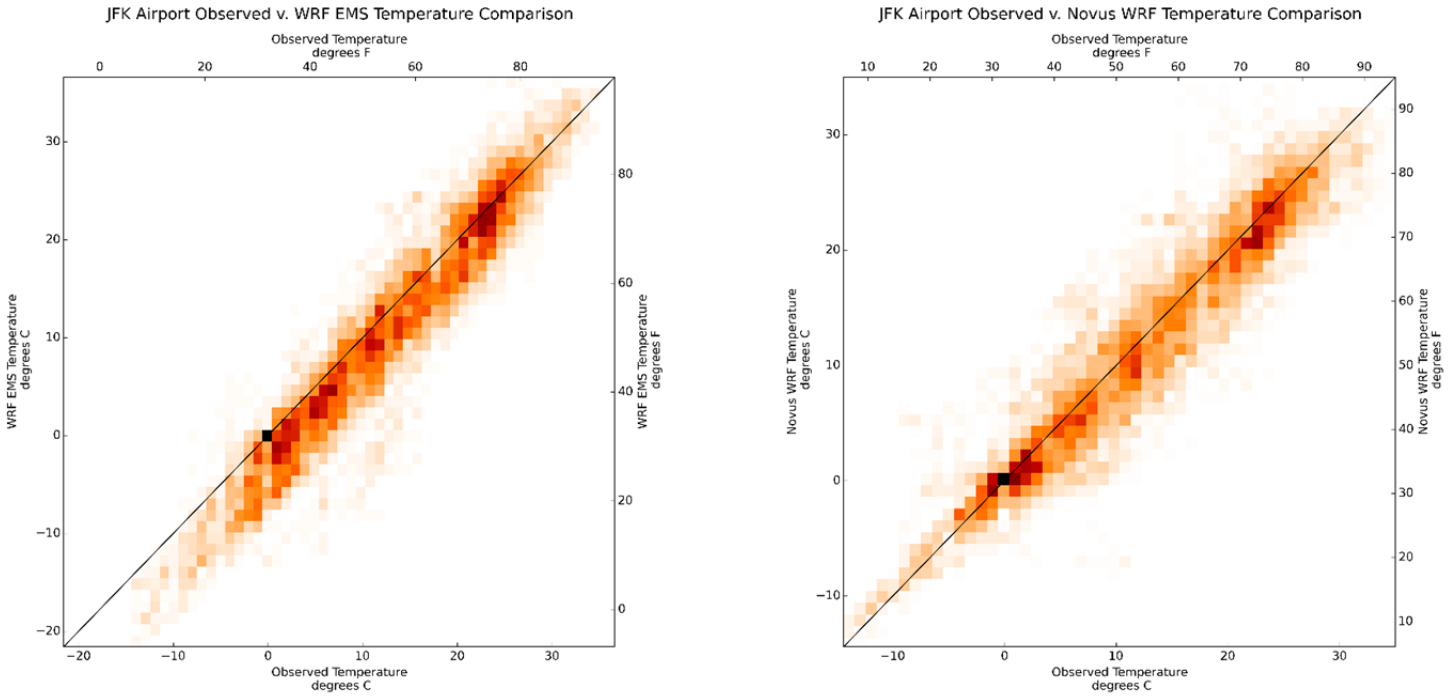

Figure 34 – Novus WRF & WRF EMS vs JFK Airport Observed Temperature

The above plots confirm that both models slightly under predict when compared to measured data within the modeling domain. As shown in the statistical comparison between the models, Novus WRF performs slightly better than WRF EMS, especially at lower temperature levels.

Figure 35, Figure 36, and Figure 37 show a similar comparison between the modeled and measured datasets for wind speed. It can be seen that both models under predict measured values with WRF EMS doing a better job, especially at higher wind speeds.

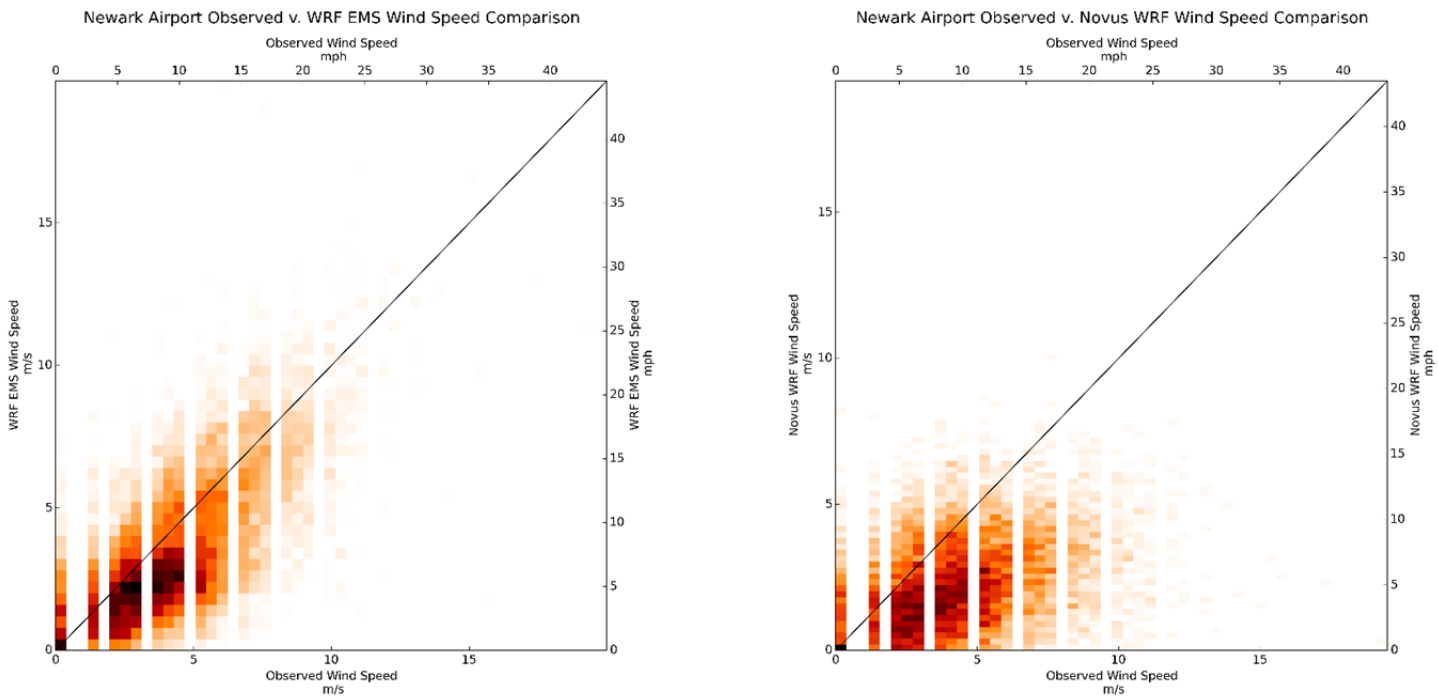

Figure 35 – Novus WRF & WRF EMS vs Newark Airport Observed Wind Speed

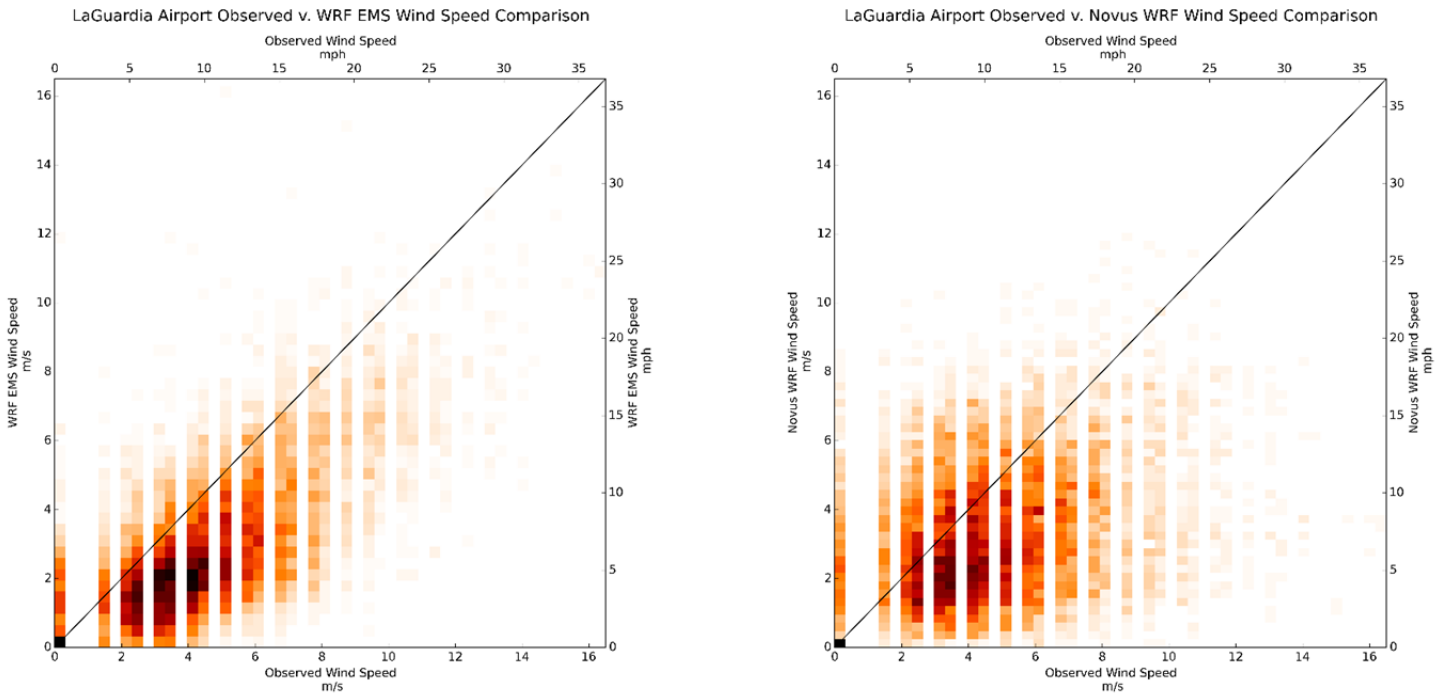

Figure 36 – Novus WRF & WRF EMS vs LaGuardia Airport Observed Wind Speed

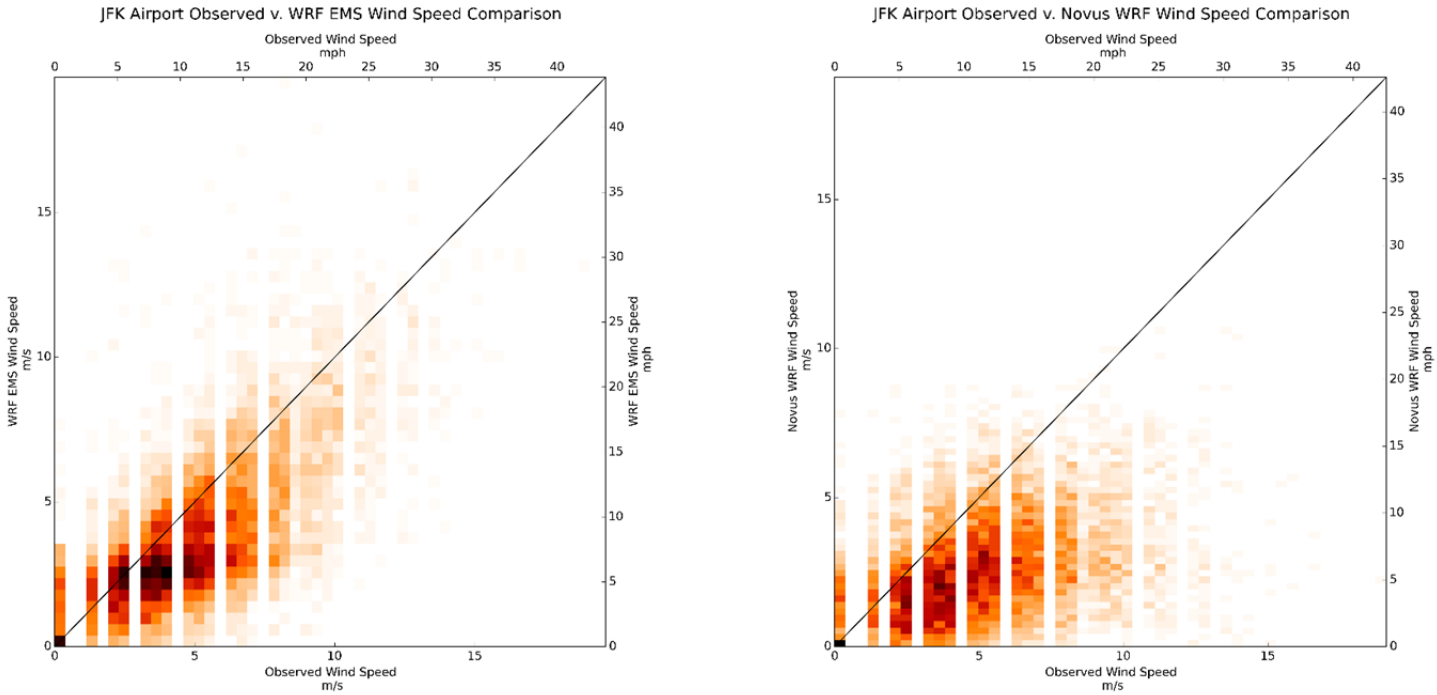

Figure 37 – Novus WRF & WRF EMS vs JFK Airport Observed Wind Speed

These comparisons show that for the New York City domain, WRF EMS performs to the same standards as Novus WRF, and is suitable for use as a simplified model for ASHRAE members.

5.3 On-line Procedure

This WRF EMS methodology is provided to the ASHRAE community through a simplified procedure based on a set of instructions and computer scripts located on-line (http://klimaat.ca/emspy). The instructions describe how to setup a modeling domain through a graphical interface (or through a simple script), simply from knowledge of a site's latitude and longitude. A script is available to download data as required, make the necessary changes to the physics options, split up the simulation into chunks, run each chunk in sequence, and finally post-processes the results to generate a time series at a desired location (not necessarily the original central latitude/longitude). As the scripts are simply text files written in a high-level language (Python), the user is free to modify to suit. Additionally, the on-line procedure will be maintained to respond to future WRF EMS developments and provide additional functionality.

6 Discussions and Conclusions

Advanced weather forecasting models such as the Weather Research and Forecasting model (WRF) are capable of integrating meteorological data from a large number of observational sources (surface and upper air stations, radar, satellite and modeling data) in order to provide high-resolution climate data for a location where observed data is not available. This study aimed to achieve ASHRAE RP-1561's two overarching goals: 1) to develop a methodology for ASHRAE members to harness the power of modern mesoscale modeling techniques in order to derive meteorological conditions specific to their site; and, 2) to evaluate this methodology against available observational data in a variety of geographic categories and categories.

This document presents a final report on the results of this project. We provided an overview of the modeling methodology developed and then evaluated the performance of the methodology against Mesonet observed data in estimating hourly meteorology, i.e. the "weather", within climate zone. Modeling results indicate that temperature predictions are among the best for most of the geographic categories. Humidity is also predicted well within the evaluation criterion. The model performs reasonable well in wind speed and direction with larger bias than temperatures. It is believed that Mesonet wind data are likely affected by nearly buildings and obstacles that result in poor measurement performance.

We also performed an 8-year modeling evaluation to address common design elements such as the 99% heating dry-bulb temperature, i.e. the "climate". Detailed modeling performance evaluations from those locations suggested that WRF is able to generate gridded localized climate information, with varying accuracy, as a supplemental tools to the Handbook of Fundamentals data. In particular temperature, and to a lesser extent humidity and wind, demonstrate great predictive skill. However, the methodology struggles to predict precipitation or solar radiation with any accuracy given the general failure of meteorological models in predicting cloud cover.

Lastly, we prepared a simplified WRF modeling method based on the WRF EMS system.

Limited model evaluations suggested this method can generate data with comparable quality to the Novus WRF method. A step-by-step recipe for a simplified version of the methodology has been made available on-line.

The quantity and spatial distribution of wind-driven rain are important considerations for building envelope design. However, a good wind-driven rain index requires both wind and rain data be either measured or modelled at the location simultaneously. It was speculated that the current modeling procedure could be utilized. Unfortunately, the current WRF modeling grid spacing is insufficient to provide the required level of detail. Furthermore, the WRF model, like any other weather/climate model, cannot replicate all weather patterns particularly under precipitation conditions. Hence, we do not recommend the use of this procedure for wind-driven rain analysis.

However, there are numerous other potential applications for these modeling results. While the focus of the procedure is the generation of a time series of data for a single location, it is often useful to examine the variability of a given climate parameter within a region. For example, as shown in Appendix A, large variations are observed in e.g. temperature in San Francisco of 5 °C (9 °F) within 4 km (2.5 mi). Generally, this level of detail cannot be determined from either observational datasets or re-analysis datasets at the resolutions currently available or planned. Another application of the modeling data is in selecting representative data for a study location. This can also be seen in the results from the San Francisco modeling domain. With a number of monitoring stations to choose from, it can be difficult to identify which monitoring station is the most representative for a specific location. Engineers and architects will often create a blended meteorological set in an attempt to represent the local meteorology. Mesoscale modeling can avoid this by providing site-specific meteorology which is determined by measured data and site parameters.

It is important to mention that this project developed an approach by implementing a mesoscale model as an advanced numerical tool to generate high-resolution climate data (on-site), particularly for where HOF climate data is unavailable. Further development is required to prepare and streamline the model output directly to a format suitable for ASHRAE members.

7 References

ASHRAE Technical Committee 4.2. (2009). Chapter 14 - Climatic Design Information. In *ASHRAE, 2009 Fundamentals*.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land-surface/ hydrology model with the Penn State/ NCAR MM5 modeling system. Part I: Model description and implementation. Mon. Wea. Rev., 129, 569–585.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, J. Atmos. Sci., 46, 3077–3107.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6), J. Korean Meteor. Soc., 42, 129–151.

Janjic, Z. I., 2000: "Comments on Development and Evaluation of a Convection Scheme for Use in Climate Models", J. Atmos. Sci., 57, p. 3686.

Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61 pp.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritcsh scheme, The representation of cumulus convection in numerical models, K. A. Emanuel and D.J. Raymond, Eds., Amer. Meteor. Soc., 246 pp.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res., 102 (D14), 16663–16682.

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

Stauffer D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited area mesoscale model. Part I: Experiments with synoptic-scale data. Mon. Wea. Rev., 118, 1250–1277.

Wu, S.-Y., Krishnan, S., Zhang, Y., Aneja, V., 2008. Modeling atmospheric transport and fate of ammonia in North Carolina, Part I. Evaluation of meteorological and chemical predictions. Atmospheric Environment 42, 3419e3436.

Zhang, Y., Liu, P., Pun, B., Seigneur, C., 2006. A comprehensive performance evaluation of MM5eCMAQ for summer 1999 southern oxidants study episode, Part-I. Evaluation protocols, databases and meteorological predictions. Atmospheric Environment 40, 4825e4838.