Evaluating WRF Meteorological Downscaling Performance for Use in Air Quality Dispersion Modeling Studies

Gary Moore • William Leatham • Jeffery Connors • Robert Paine (AECOM, Chelmsford, MA)

Introduction

Traditional model performance evaluation statistics used to evaluate prognostic meteorological models generally are not suited to local near source air quality modeling. Air quality modeling exercise protocols require local scale meteorology (100's of meters). Dynamical downscaling using a multi-level application of a prognostic model like the Weather Research Forecast (WRF) model is handed off to a diagnostic or even a CFD model.

Additional information about the suitability of the WRF model output for the surface meteorology in the local modeling domain needs to be generated. Several additional diagnostic statistical measures and graphics were developed to examine a WRF model output’s suitability for input to the diagnostic meteorological model. These added performance tools are described and discussed in this presentation.

Motivation and Approach

Motivation: Improve air quality met model evaluation products

Approach: Add diagnostics to traditional WRF evaluation exercises

TRADEITIONAL STATISTICS

• Mean bias
• Mean square error or its square root (MSE or RMSE)
• (absolute) Gross Error
• Index of Agreement (IA)
• Normalization (e.g. differences over averages or division by observed)
• A priori performance targets

VALUE ADDED EVALUATION ATTRIBUTES

• Confidence intervals - robust, flexible, and meaningful
• Stratification of residuals focus on conditions of interest
• Directionality - bias in trajectory heading is important
• Persistence - degree of hourly independence of residuals
• Accumulation metrics – duration and build up of biases
• Episodic summaries - frequency of residuals
• Classification of residuals - are there useful principle components?

EXERCISE SUMMARY

• Nested 36, 12, and 4 km domains with 37 vertical layers.
• Two-way nesting
• National Center for Environmental Protection (NCEP) Global Forecast System (GFS) for initialization conditions, boundary conditions and sea surface temperature (SST)
• Four Dimensional Data Assimilation (FDDA) analysis nudging is performed on the 36 and 12 km domains, while FDDA observational nudging (OBSERGED) is performed on the 4 km domain (old Tru)
• Based on the University of Alaska Fairbanks (UAF) weather forecast model configuration

WRF Use: Dynamical Downscaling of Hourly Meteorology into CALMET Centered on Alder Met Tower

Summary of Traditional Evaluation Statistics

• Passing “grade” for all variables except wind speeds – always an over prediction
• 2012-2013 appears to be overall best predicted of the 3 years (had most observed data)
• Not much physical information to judge adequacy for air quality modeling
• 6 different WRF configurations were exercised as sensitivity runs with none showing significantly better performance
• Winter does better than summer for low wind speeds – opposite is true for higher wind speeds
• Summer residual biases are often not significant from zero for higher wind speeds
• Winter wind residuals show the greatest variability and bias

Closer Look at Wind Speed (‘Brickplot’)

Using 96% Confidence Intervals for 2009

• Winter month wind speeds and better predicted with long runs of days of “passing”
• Blue has 29% days “passing”
• RMSE has 38% days “passing”

Cross Correlation/ Factor Analysis

• Correlations of residuals with observations or predictions or each other are generally rather small
• Residual variable factor analyses does not find one or two really dominant components (only 20-30% of residual variance each)

Footnotes

• Only correlations >0 passing Fisher Z at 95% confidence level are included
• (b) = bias
• (m) = mean = observed
• (p) = predicted

Bakergrams for Daily Wind Speed Statistics for 2012-2013

• Summer month wind speeds are better predicted with long runs of days of “passing”
• Blue has 29% days “passing”
• RMSE has 38% days “passing”

EXERCISE SUMMARY

• Nested 36, 12, and 4 km domains with 37 vertical layers.
• Two-way nesting
• National Center for Environmental Protection (NCEP) Global Forecast System (GFS) for initialization conditions, boundary conditions and sea surface temperature (SST)
• Four Dimensional Data Assimilation (FDDA) analysis nudging is performed on the 36 and 12 km domains, while FDDA observational nudging (OBSERGED) is performed on the 4 km domain (old Tru)
• Based on the University of Alaska Fairbanks (UAF) weather forecast model configuration

Monthly Station Ensemble Wind Rose Comparisons (2010) — High Degree of Agreement (Some Over Prediction)

• K-S test indicates summer and winter wind speed residuals are significantly different
• Only December 2009 and 2012 wind speed residuals are similar and pass K-S test
• Largest failures occur near the average residual point for both speed and direction

Conclusions

• The degree to which multi-day episodes of similar residuals can be related to specific weather conditions
• What stations/locations may be most difficult to apply further downscaling modifications

These and other findings provide an indication of how the further downscaling of the meteorology for air quality worst case conditions should be conducted. In the current example the 4 km WRF output appears suitable to drive high resolution diagnostic modeling during the years 2009, 2010, and 2012-2013 when more refined terrain and roughness are considered.