Creation of a WRF Ensemble for Improved Wind Forecasts at Turbine Height

Adam J. Deppe[1], Eugene S. Takle[1,2,3] and William A. Gallus, Jr.[1]
Department of Geological and Atmospheric Sciences, Iowa State University[1]
Department of Agronomy, Iowa State University[2]
Ames Laboratory, Iowa State University[3]

1. Introduction

In recent years, wind energy production has undergone rapid growth, and the U.S. Department of Energy goal of having 20% of the nation's electrical energy from wind by 2030 will require continued growth. Wind, unlike other sources of energy, varies substantially over both space and time. Therefore, the production rates of wind energy fluctuate more strongly than other traditional fossil fuel sources of energy generation. To optimize wind for power generation, accurate forecasts are needed.

Unfortunately, there have been very few evaluations of model forecasts of winds at 80m, a height where the influence of friction from the earth's surface can vary greatly depending on the time of day, season, and vertical temperature stratification of the atmosphere. Meteorologists traditionally have focused wind forecasts at the 10m level, a height at which official wind observations are taken and a level at which winds are strongly influenced by surface friction. Prior wind forecasting research in the western United States has focused on flow in complex terrain like mountains (Wood 2000) and is therefore not applicable in Iowa where boundary layer stratification, low-level jets, and changing surface conditions are likely to be the dominant factors providing uncertainty in short-term forecasts at 80m. Other modeling studies have taken a more statistical approach to predict wind speed at different levels (Huang et al. 1996); however, none have been done over the state of Iowa.

Corresponding author address: Adam J. Deppe, ISU, 3010 Agronomy Hall, Ames, IA. 50011 Email: ajdeppe@iastate.edu.

In this study, the ability of the Weather Research and Forecasting (WRF) model to accurately reproduce the 80m wind speed was evaluated by comparing WRF simulations using six different planetary boundary layer (PBL) schemes to observations of 80m wind speed gathered at the Pomeroy, Iowa wind farm site.

2. Data and Methodology

The WRF model with 10-km horizontal resolution and domain compromising much of the Great Plains was used to explore improvements in wind speed forecasts at hub height (80m). The following PBL shemes were evaluated: the Yonsei University scheme (YSU), the Mellor-Yamada-Janjic scheme (MYJ), the quasi-normal scale elimination PBL scheme (QNSE), the Mellor-Yamada Nakanishi and Niino level 2.5 PBL scheme (MYNN 2.5), the Mellor-Yamada Nakanishi and Niino level 3.0 PBL scheme (MYNN 3.0), and the Pleim PBL scheme (also called the asymmetric convective model [ACM2]). The model configurations above were run using both the Global Forecast System (GFS) and North American Model (NAM) analyses for initial and lateral boundary conditions. Observed data was taken from a 80m meteorological tower on the southwest side of the Pomeroy, Iowa wind farm.

Prior research has shown that the ensemble mean forecast is almost always more accurate than any single member forecast (Arribas et al. 2005, Bowler et al. 2008). We tested this assumption by comparing both the ensemble mean and individual PBL scheme performance to the actual observations of the wind speed over the wind farm. Because at any one time a

PBL forecast model can either under-predict or over-predict the observed wind speed, mean absolute error (MAE) was used to compare the different PBL schemes and the ensemble average. When computing MAE, lower values indicate better skill. Model and ensemble biases were also analyzed.

3. Results

a.) Pre-Processing

Little model spread was present among all six PBL schemes at the same initialization time. With small model spread, the models either predicted the wind speed correctly, or the models were all wrong. Therefore, two pre-processing techniques were investigated to improve model scheme spread and skill over 10 cases during January 2010.

The first attempt to improve model skill involved using different perturbations of the initial and lateral boundary conditions for the GFS model. Three perturbations were picked at random and run for 10 cases in January 2010 using the YSU and MYNN 3.0 PBL schemes. The results of this trial increased model spread; however, model skill was decreased (Table 1).

The second attempt to improve model skill involved changing the time initialization. PBL schemes were initialized at 18Z, 00Z, and 06Z over a 10-day period in January to determine skill. Again, the YSU and MYNN 3.0 PBL schemes were used for this test. The 00Z and 18Z time initializations showed the best skill while the 06Z initialization, the initialization closest to the forecast period, showed the lowest model skill (Table 2). Because the time initialization results showed approximately the same model spread as seen in the perturbation cases, it was determined that the best ensemble would be made up of both different PBL schemes and 00Z and 18Z time initializations.

Perturbation Number	MYNN 3.0 MAE (m/s)	YSU MAE (m/s)	Ensemble MAE (m/s)
2	2.34	2.06	2.05
4	2.18	2.04	1.98
15	2.27	2.18	2.08

Table 1: MAE from three different GFS perturbations using the YSU and MYNN3.0 PBL schemes. The two member ensemble average is also listed.

Time Initialization	MYNN 3.0 MAE (m/s)	YSU MAE (m/s)	Ensemble MAE (m/s)
18Z	1.88	1.78	1.69
00Z	1.82	1.74	1.63
06Z	1.83	2.07	1.73

Table 2: MAE from three different initialization times.

b.) Post-Processing

After investigating the pre-processing techniques, we investigated three post-processing techniques which included training the model, the neighborhood approach, and bias correction of the wind speed.

The first post-processing technique involved improving day 2 forecasts (hours 30-54) by training the model. In this method, day 1 forecasts (hours 6-30) were analyzed. Based on these results, the three most accurate PBL schemes (lowest MAE) were chosen and a Picked ensemble was developed to forecast day 2 wind speeds. The three least accurate PBL schemes (highest MAE) made up the Non-Picked ensemble (Table 3). Results of this technique show that the most accurate day 1 forecasts do not give the most accurate day 2 forecasts. From this 15 member case study, the Non-Picked ensemble showed the highest skill 4 out of the 15 times (27%), the Picked ensemble showed the highest skill 5 out of the 15 times (33%), and the ensemble, incorporating all six model members, showed the highest skill 6 out of the 15 times (40%). Therefore, the training approach is not a reliable method to predict wind speed as conditions change from day to day.

Model Number	Day 1 MAE	Times Picked
00Z MYJ GFS with a 10km grid spacing	2.51	5
00Z MYJ NAM with a 10km grid spacing	2.61	6
00Z Pleim NAM with a 10km grid spacing	2.58	4
00Z Pleim GFS with a 10km grid spacing	2.36	9
00Z YSU NAM with a 10km grid spacing	2.32	11
00Z YSU GFS with a 10km grid spacing	2.37	10
Ensemble Mean	1.97	
Day 2 Picked ensemble best MAE	**Day 2 Non-Picked ensemble best MAE**	**Day 2 All Member Ensemble best MAE**
5/15	4/15	6/15

Table 3: MAE calculated for the first 24 hour period. The three PBL schemes with the greatest skill were chosen, making up the Day 2 Picked ensemble. Times Picked indicates the number of times a model is chosen as a member of the Day 2 ensemble. Non-picked ensemble incorporates least accurate model for the first 24 hour period. Day 2 All Member Ensemble incorporates all six model members.

The second post-processing technique improved forecasts using the neighborhood approach. In most forecast models, the forecast is taken from the grid point closest to the location. However, in the neighborhood approach, the grid points around the location of interest are averaged, an approach used successfully for precipitation. The results of this test showed opposite results for different PBL schemes. The YSU scheme became more accurate with larger grid spacing while the MYNN 3.0 scheme became less accurate. Ensemble results from this approach show better model skill when an average was taken from a box consisting of 5 by 5 grid points, as model skill did not improve after this point (Table 4). However, the improvement is very small and not as large as that occurring from other methods.

Grid Averaging	MYNN 3.0 MAE (m/s)	YSU MAE (m/s)	Ensemble MAE (m/s)
Point	1.82	1.74	1.63
3x3	1.82	1.72	1.61
5x5	1.82	1.70	1.59
11x11	1.83	1.64	1.59
17x17	1.84	1.62	1.59
21x21	1.85	1.61	1.59

Table 4: MAE associated with the neighborhood approach.

The third post-processing technique improved forecasts based on biases observed in the PBL schemes. A bias in the model was computed by analyzing 30 random cases from all seasons between June 2008 and June 2009 (Figure 1). In all PBL schemes except the YSU, a diurnal pattern or cycle exists. We noticed that a negative bias, or under-prediction of the wind speed, occurs from hour 12 to hour 20 or from 6 a.m. to 2 p.m. Central Standard Time (CST), while a positive bias (over-prediction) occurs from hour 20 to hour 12 or from 2 p.m. to 6 a.m. CST. The same pattern exists in day 2 of the 54 hour forecast and was present in both the GFS and NAM initializations. Therefore, knowing that some models have inherent errors based on the diurnal cycle, we can develop a bias correction to improve forecasts.

To understand if any larger biases existed in the PBL schemes, four possible biases were examined; the diurnal cycle, wind speed along with direction, wind speed only, and wind direction only. A bias correction for each PBL scheme, 00Z and 18Z time initialization, and the GFS and NAM initial boundary conditions was done over a one month period from Oct. 11, 2009 to Nov. 11, 2009. The GFS 00Z bias correction table is shown below (Table 4.).

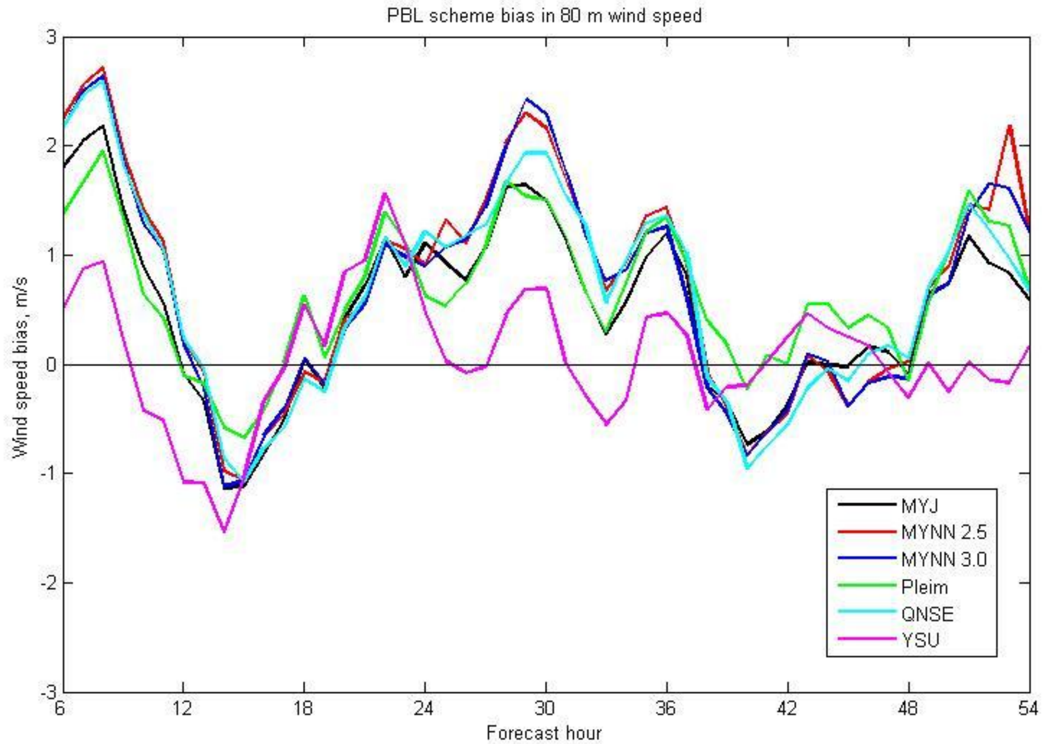

Figure 1: Composites of PBL biases by hour. Each line represents a different PBL scheme. Notice a diurnal bias feature present in the PBL schemes.

Bias Corrections	MYJ (m/s)	MYNN 2.5 (m/s)	MYNN 3.0 (m/s)	Pleim (m/s)	QNSE (m/s)	YSU (m/s)	Ensemble (m/s)
No Bias	2.34	2.49	2.41	2.36	2.45	2.28	2.27
Diurnal Cycle	2.29	2.33	2.28	2.27	2.30	2.21	2.18
Wind Direction	2.27	2.27	2.26	2.29	2.28	2.24	2.17
Wind Speed and Direction	2.15	2.16	2.14	2.17	2.17	2.10	2.05
Wind Speed	2.05	2.04	2.01	2.09	2.07	1.99	1.97
Best Improvement	.29 m/s – Wind Speed	.45 m/s – Wind Speed	.40 m/s – Wind Speed	.27 m/s – Wind Speed	.38 m/s – Wind Speed	.29 m/s – Wind Speed	.30 m/s – Wind Speed
% of Improvement	14.1%	22.1%	20.0%	13.0%	18.4%	14.6%	15.2%

Table 4 - MAE associated with different bias corrections developed for each PBL scheme for both 00Z and 18Z initialization times. This example is from the 00Z GFS run. The best improvement was seen with the wind speed bias correction. This case study was done from Oct. 11, 2009 to Nov. 11, 2009.

Based on the results from the bias corrections, a six member ensemble was created from the most accurate schemes using the wind speed bias correction. The six members found to have the most skill after the wind speed bias correction included 18Z Pleim GFS with a 10km grid spacing, 18Z Pleim NAM with a 10km grid spacing, 00Z Pleim GFS with a 10km grid spacing, 00Z YSU NAM with a 10km grid spacing, 00Z YSU GFS with a 10km grid spacing, and 00Z MYJ GFS with a 10km grid spacing. From these six members, an ensemble was created called "operational model". For comparison purposes, a deterministic forecast (a single PBL scheme that showed the best model skill) was compared against the other ensembles. The standard deviation was also calculated to determine model spread. To test the operational model ensemble, 25 random cases from the summer and fall of 2010 were used. Results showed that the six member operational model had the best model skill (lowest MAE) of any of the other six member ensembles tested, both before and after the wind speed bias correction (Table 5). The standard deviation of the operational model is also larger than that of any of the other ensembles, indicating a larger spread in the operational model which should be helpful in capturing outlier events.

Ensemble	MAE after Bias Correction (m/s)	Standard Deviation after Correction (m/s)	MAE Prior to Bias Correction (m/s)
GFS 00Z	1.67	1.99	0.74
GFS 18Z	1.66	2.05	0.80
NAM 00Z	1.68	1.91	0.67
NAM 18Z	1.70	1.93	0.73
Deterministic Forecast	1.70	1.77	- - -
Operational Model	1.52	1.67	0.98

Table 5: MAE of operation model ensemble after wind speed bias correction compared to other six member ensembles tested for 25 cases during the summer and fall of 2010. The deterministic forecast is the best individual model found from the period studied. Standard deviation (measure of model spread) for each ensemble is also calculated.

5. Summary and Conclusions

Understanding the biases and strengths of different PBL schemes will help to improve wind speed forecasts. From this study, we discovered that combinations of pre and post-processing techniques are required to improve wind speed forecasts at 80m. During pre-processing, it was found that perturbations of the GFS model give more spread in data than achieved with the six PBL schemes; however, a higher MAE is also created. Also, GFS initial and lateral boundary conditions showed higher model skill than the NAM. However, the most important discovery during the pre-processing stage was that different time initializations give equal spread, but better model skill.

During the post-processing stage, the first technique attempted was training the model based on day 1 results. It was found that using the training method to predict wind speed is not reliable as conditions change from day to day. The second technique attempted was the neighborhood approach, which increased the accuracy of the models, albeit just slightly. The post-processing technique that was the most successful was the bias correction. Many different bias corrections were tested however, the wind speed bias correction proved to yield the best results. From these results, a six member operational model ensemble was developed that outperformed other ensembles tested.

6. Acknowledgments

We would like to thank MidAmerican Energy Company for providing the 80m observations. Partial funding was supplied by NSF grant BCS0618823, DOE grant #13-450-141201, Ames Laboratory Project 290-25-09-02-0031, and ERPC grant #400-60-12.

7. References

Arribas A., K. B. Robertson, and K. R. Mylne, 2005: Test of a poor man's ensemble prediction system for short-range probability forecasting. *Mon. Wea. Rev.*, **133**, 1825–1839.

Bowler, N. E., A. Arribas, and K. R. Mylne, 2008: The benefits of multianalysis and poor man's ensembles. *Mon. Wea. Rev.*, **136**, 4113–4129.

Huang H. and Z. S. Chalabi, 1996 Use of time-series analysis to model and forecast wind speed. *J. Wind Eng. Ind. Aerodyn.,* **56**, 311–322.

Wood, N., 2000: Wind flow over complex terrain: A historical perspective and the prospect for large-eddy modeling. *Bound.-Layer Meteor.*, **96**, 11–32.