FORECASTING SOUTHERN PLAINS WIND RAMP EVENTS USING THE WRF MODEL AT 3-KM

Kristen T. Bradford [1,2], Dr. Richard L. Carpenter, Jr.[3], and Brent L. Shaw[3]

[1]National Weather Center Research Experiences for Undergraduates Program
and
[2]Valparaiso University
Valparaiso, Indiana

[3]Weather Decision Technologies, Inc.
Norman, Oklahoma

ABSTRACT

Wind ramp events—extreme and rapid changes in wind power output due to abrupt changes in wind speed—are a growing concern for the wind energy industry; therefore, precise forecasting of these phenomena is crucial to the advancement of wind power in the United States. Weather Decision Technologies, Inc., (WDT) is partnering with NanoWeather, Inc., to create a wind forecasting system, called WindPredictor™, in order to precisely predict winds (and, in turn, ramps) for the energy industry. WDT's contribution to WindPredictor will be a customized version of the Weather Research and Forecasting (WRF) model, which is currently being run on a 3-km grid. This paper assesses the 3-km WRF's performance regarding ramp event prediction. A comparison between surface wind forecasts and hourly METAR observations was utilized to assess its performance.

1. INTRODUCTION

1.1 The Importance of Wind Energy

Wind energy is a rapidly growing industry in the United States. In fact, the U.S. Department of Energy (DoE) declared that the United States has the ability to obtain 20% of its energy from wind power by 2030 (AWEA 2009). Should this be implemented, the year 2030 would experience a cumulative, decadal increase of 500,000 jobs, savings in electric costs of $128 billion in comparison to those of today (AWEA 2009), and greenhouse gas emissions reduced by more 825 million tons—the equivalent of taking 140 million of today's cars off the roads (Goggin 2009). Therefore, it is clear that the DoE's goal has the potential to significantly affect the nation, both in economical and environmental terms. However, in order for the country to eventually experience these benefits, the wind energy industry requires extremely precise wind forecasts to run efficiently.

[1]Kristen Bradford
address: 7844 Autumn Woods Dr.
Ada, MI 49301
e-mail: Kristen.Bradford@valpo.edu

One of the growing concerns of the wind energy industry is the occurrence of ramp events, which are extreme and rapid changes in wind power output due to abrupt changes in wind speed. If these incidents are properly predicted, a utility company can work to reallocate or balance the energy grid; on the other hand, if the utility company is caught off guard, significant energy management issues arise. For instance, the Electric Reliability Council of Texas (ERCOT) experienced a rapid decrease in power output on 26 February 2008 when the boundary layer quickly cooled and stabilized, resulting in a reduction of winds so large that it was forced to declare a system emergency (Francis 2008, Zack 2008). Incidents like this can prove to be extremely costly for wind farms and, consequently, the ability to foresee ramp events is quickly becoming a crucial priority of the industry (Francis 2008).

A Norman, OK based company named Weather Decision Technologies, Inc., (WDT) acknowledged this need wind forecasting system within energy industry and received an Oklahoma Economic Development Generating Excellence (EDGE) grant to support their endeavor. The wind forecasting system, called WindPredictor™, couples WDT's customized version of the Weather Research and Forecasting (WRF) model (which is currently being run on a 3-km grid) and the Uncoupled Surface Layer (USL) model, a microscale model developed by NanoWeather, Inc. The precision of WindPredictor™ enables forecasts specific not only to the wind farms as a whole, but to the individual wind turbines. WDT and NanoWeather also want to ensure that the model they develop correctly predicts significant changes in wind power. This paper will attempt to distinguish how accurately the 3-km WRF is forecasting the occurrence of ramp events.

1.2 Ramp Events

It is important to note that the wind energy industry has only recently begun to assess the nature of ramp events. Therefore, there is no universally accepted threshold employed to detect them. Still, most classifications are similar to that of this paper, which defines a ramp event as a change in wind power output of greater than or equal to 20% capacity in magnitude over a one-hour time period.

Ramp events are induced by a variety of meteorological phenomena. In their West Texas study, Freedman, Markus, and Penc (2008) categorized ramps into ramp-up (an increase in wind power) and ramp-down (a decrease in wind power) events. They found that increases in power output tended to accompany frontal systems, dry lines, convection, and the low-level jet, while decreases were usually associated with a quickly weakening pressure gradient. Zack (2008) also noted the affect that shallow, cold air masses and turbulent mixing can have on the initiation of a ramp event. However, ramp-down events can also occur during high wind events, as wind turbines cannot withstand extremely high wind speeds and must be turned off to preserve the machinery. High-speed shutdown usually occurs between 22 and 25 m s^{-1}, depending on the wind turbine (Freedman, et al. 2008).

As ramp events have not been extensively studied, very little is known about their climatology. Freedman, et al. (2008) compiled a basic climatology of the ramp events observed between 2005 and 2006. Of the 59 ramps observed, roughly 60% were ramp-up events and the remaining 40% were ramp-down events. While ramp-down events did not show any distinct climatological pattern, ramp-ups exhibited a broad annual maximum spanning from late winter through summer. Ramp-up events also tended to occur during the evening (the mode occurring at around 2300 UTC). Both the diurnal and annual climatology supported the researchers' conviction that ramp-up events were primarily caused by convective features.

2. METHODOLOGY

2.1 Domain of Study

Figure 1 depicts the sector of the Southern Plains upon which this study focuses. A total of 34 locations were chosen due to the fact that they are also METAR observation sites, thus making surface wind observations readily available for comparison to the WRF forecasts.

Fig. 1: A map of the domain of study. Map created with Google maps

2.2 Data

For the EDGE WRF domain, a version 3.1 of the Advanced Research WRF was used (ARW; Skamarock et al. 2008). The horizontal domain consists of 512 x 512 grid points with a spacing of 3 km. In the vertical, there are 32 vertical levels from the surface to 50 mb, with higher resolution in the boundary layer. For maximum computational efficiency, the WRF adaptive time step is used but is not allowed to exceed 30 sec. Additionally, surface observations were incorporated during a 1-hour Four Dimensional Data Assimilation (FDDA; Liu et al. 2006) period.

Data for this study was obtained between 4 June and 8 July 2009. Point forecasts were taken from the first 24 hours of the 0900 UTC 3-km WRF run for each of the 34 sites. Only the 2-m hourly wind forecasts were used for the study. These were then compared to their corresponding METAR observations. The METAR data was filtered to only include observations from 10 minutes before to 15 minutes after the hour. If more than one observation occurred during the 25-minute window, only the closest one to the top of the hour was recorded. Due to a combination of model complications and missing METAR observations, some of the data had to be disregarded, resulting in roughly 80% of the forecast hours that could be analyzed.

2.3 Determining Ramp Events

Wind power *P* is calculated as,

$$P = .5 \times A \times \rho \times v^3, \quad (1)$$

where *A* is the rotor swept area, *ρ* is the air density, and *v* is the wind speed. Because *A* varies from wind farm to wind farm, an exact power value could not be calculated from the collected data. However, due to the fact that *A* is constant for a given wind farm, it can be concluded that power is proportional to the cube of the wind speed, assuming density also remains constant.

The actual power output of a wind turbine, however, is much more complicated than a cubic function due to the characteristics of its engineering (Figure 2). For instance, a wind less than or equal to 3 m s^{-1} will, according to (1), result in some power output; in reality, wind speeds this low cannot turn the turbines, thus resulting in no power output. On the other hand, once winds reach speeds of about 15 m s^{-1}, the wind turbines begin to perform at their maximum capability (100% capacity). Then, as speeds begin to approach 25 m s^{-1}, the turbines experience high-speed shutdown, causing power output to plummet to 0. For this study, the typical turbine power curve was approximated as linear sections, shown by Figure 3.

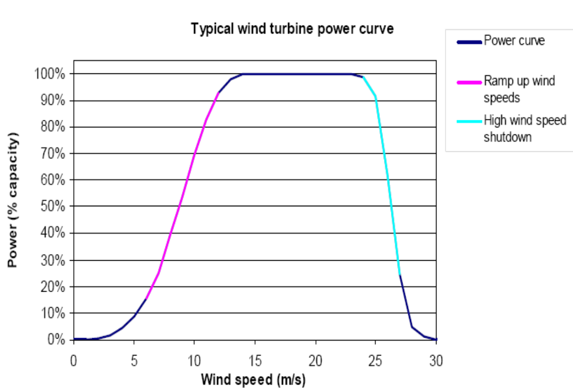


Fig. 2: A typical wind turbine power curve. Figure from Greaves, et al. 2008

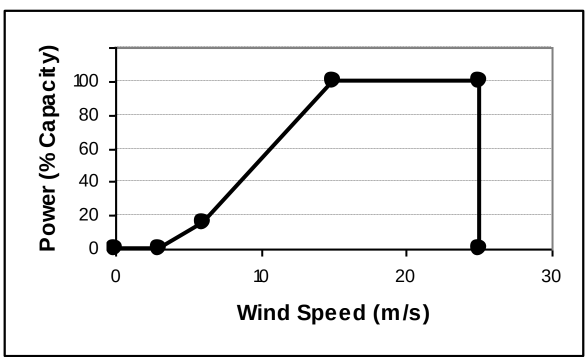


Fig. 3: An estimation of a typical wind turbine power curve.

The power percent capacity (PPC), defined as

$$PPC = \frac{\text{turbine power output}}{\text{maximum turbine power output possible}} \times 100, \quad (2)$$

was calculated for every forecast hour. Differences in PPC for consecutive hours greater than or equal to 20% in magnitude were reported as a ramp event. A contingency table was employed to divide the forecasted and observed ramp events into hits (correct forecasts), false alarms (forecasted ramp did not occur), misses (no forecast but ramp occurred), and correct nulls (no ramp event was forecasted or occurred). A ramp event was considered a hit if it occurred one hour before or after the forecast time. In future, the authors would like to divide the forecasted and observed ramps into ramp-up and ramp-down events; however, due to the data methodology used here (see Section 4), this technique would not have resulted any additional knowledge, as the model had enough trouble detecting the magnitude of the change, let alone whether or not it was an increase or decrease.

3. RESULTS

3.1 Ramp Climatology

A basic climatology of ramp events was compiled from the collected data. Ramps were most frequent in the Texas and Oklahoma panhandles, with the maximum occurring in Guymon (103 ramps), Amarillo (96), Liberal (95), Perryton (93), Pampa (92), and Gage (92). The diurnal variation of the observed ramp events can be seen in Figure 4a. It illustrates that 3 primary peaks in ramps occur throughout the day (at 0800 UTC, 1400 UTC, and 1900 UTC) with a relative lull between 0000 UTC and 0600 UTC. This differs from the study by Freedman, et al. (2008), which stated that these events usually arise during the early evening, corresponding to hours most commonly associated with convection. This discrepancy could be due to the fact that the amount of convective activity in the Southern Plains between 4 June and 8 July 2009 was unusually low. Perhaps more ramp events would have occurred during the early evening had this study been conducted during a more convectively active season. It is also noteworthy that the WRF forecasts exhibited a diurnal peak at 0200 UTC, which appears to be more compatible with Freedman's observations than those of this study (Figure 4b). The extreme difference between the forecasts and observations will be discussed for the remainder of Section 3 and 4.

3.2 Ramps Across the Domain of Study

Between 0900 UTC 4 June and 0900 UTC 8 July 2009, 283 ramp events were forecast, while a total of 2,252 occurred (Figure 5). The graph indicates that the model significantly under-forecasted the quantity of

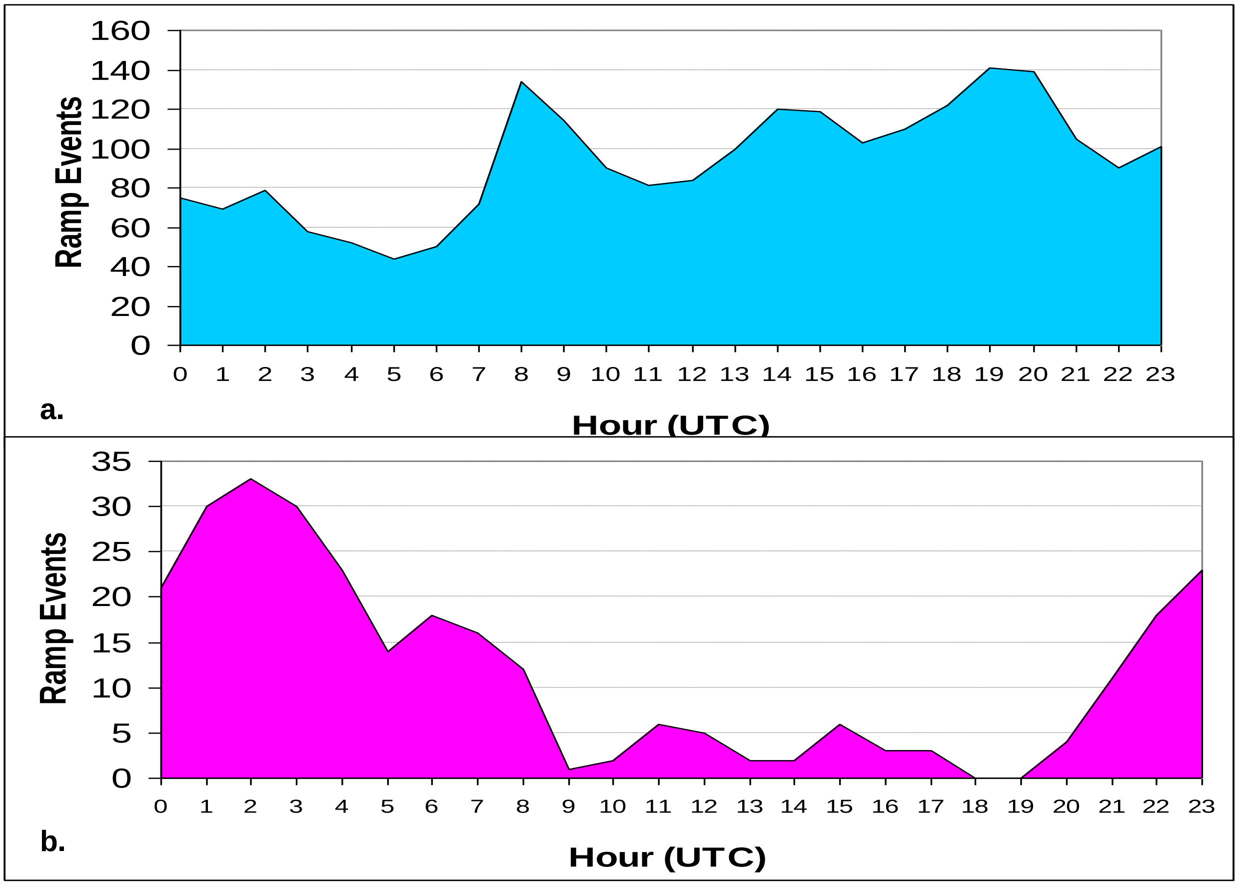


Fig. 4a,b: Diurnal variations in ramp occurrence. 4a depicts the observed ramps, while 4b is the forecasted.

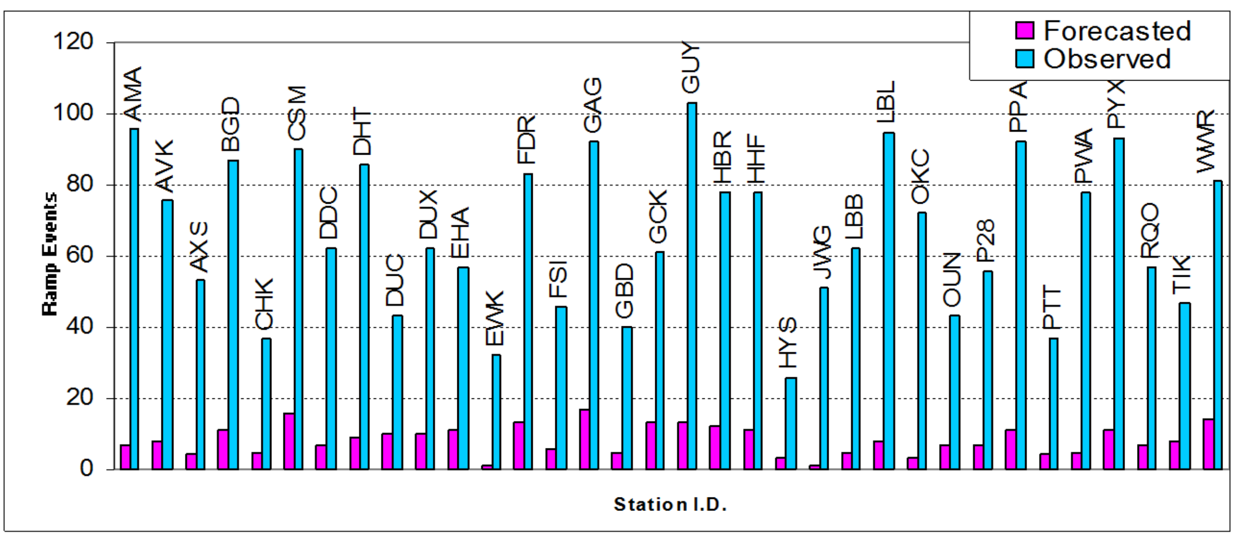


Fig. 5: Total forecasted and observed ramp events for each site.

ramps that occurred, though it performed relatively better at some sites in comparison to others. For example, Amarillo, TX (AMA) experienced 96 ramp events, yet only 7 were forecasted. Gage, OK (GAG), on the other hand, recorded slightly fewer ramp observations (92) but 11 were forecasted. Contingency tables were therefore utilized in order to evaluate the accuracy of each site's forecasts.

Table 1 depicts the overall contingency table for all 34 locations. It shows that of the 283 forecasts, only 82 were correct, indicating that 2,170 ramp events were missed by the 3-km WRF. Several statistics were calculated from this table, the most important being Peirce's Skill Score (3) and what Greaves, et al. (2008) deems ramp capture (4), which statisticians more commonly refer to as the hit rate (Jolliffe and Stephenson, 2003). Both of these values should equal 1 for a perfect forecast record and will be closer to 0 as the accuracy decreases. PSS can reach negative values for particularly poor forecast records.

$$\text{PPS} = \frac{((\text{hits} \times \text{correct nulls}) - (\text{false alarms} \times \text{misses}))}{((\text{hits} + \text{misses})(\text{false alarms} + \text{correct nulls}))} \quad (3)$$

$$\text{ramp capture} = \frac{\text{hits}}{(\text{hits} + \text{misses})} \quad (4)$$

		Observation	
		Yes	No
Forecast	Yes	Hits	False Alarms
		82	**201**
	No	Misses	Correct Nulls
		2170	**18365**

Table 1: Contingency table representing ramp event forecasts for the 34 sites across the Southern Plains.

For the overall domain of study, both PSS and ramp capture were very low, with values of .026 and .036 respectively. However, the forecast performance varied between sites, as can be observed in Figures 6 and 7. Out of the 34 locations monitored, Gage, OK and Woodward, OK (WWR)—which are only about 40-km apart—exhibited the highest statistics, while Tinker Air Force Base, OK (TIK) performed the worst with a PSS of -.033.

3.3 High-Speed Shutdown Cases

Another concern of utility companies is ramp-down events induced by wind speeds greater than 22-25 m s^{-1} (depending on the type of wind turbine). As this study examined a relatively brief period of time, only one high wind event occurred (Liberal, KS [LBL] on 10 June at 2000 UTC). In order to obtain a more complete understanding of the model's performances in terms of high-speed shutdown forecasting, the threshold was lowered to events that exhibited wind speeds of greater than or equal to 17 m s^{-1}. A total of 12 events fit this new criteria and were subject to further study.

Of the 12 cases, 11 of were observed and 1 was forecasted. None of the cases were validated, as the model significantly under-forecast the wind speeds for every observed high wind case and overestimated the hour's wind speed for the one forecasted case (Figure 8). It is interesting to note that 3 of the cases (DUC, GAG [6/14], and PYX) did correspond to ramp event hits, thus indicating that, while the model can sometimes recognize a rapid change in wind speeds associated with these high-speed shutdown cases, it is still considerably under-forecasting their magnitude.

3.4 Case Study: 18 June – 20 June

The forecast period between 0900 UTC 18 June and 0900 UTC 20 June was selected due to the

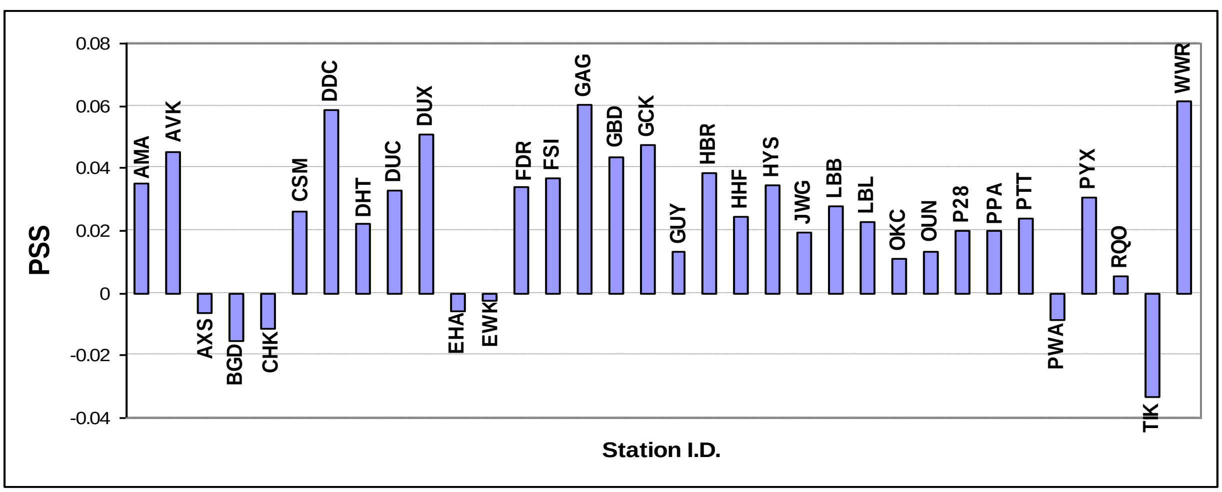


Fig 6: Peirce's Skill Score for each of the 34 sites.

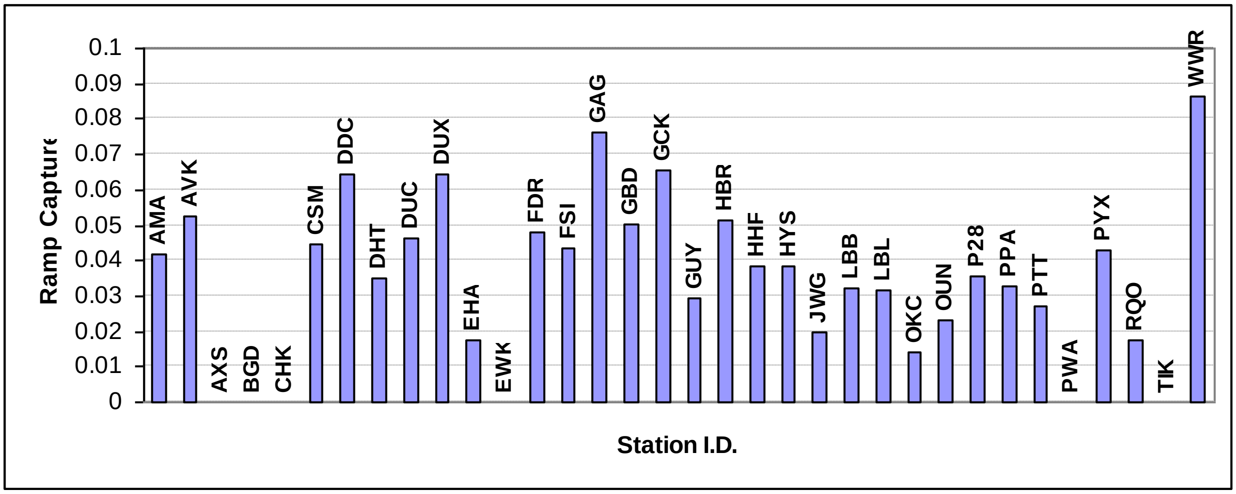


Fig. 7: Ramp capture for each of the 34 sites.

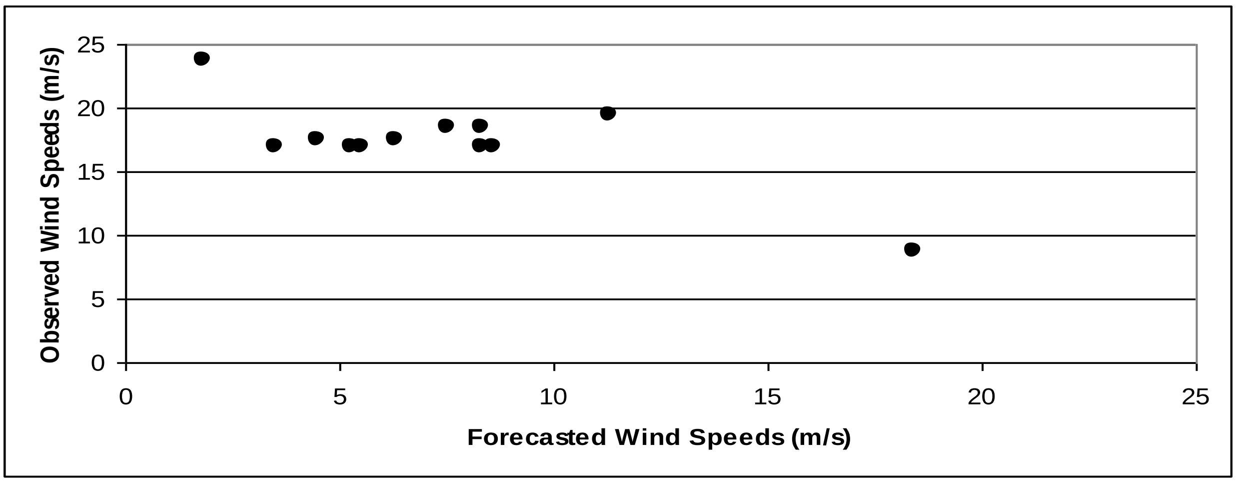


Fig. 7: Forecasted verses observed wind speeds for high-speed shutdown cases.

relatively large number of hits observed. The case study only includes only 13 of the 34 sites, which were selected if they possessed a ramp cluster. This paper defines a ramp cluster as a group of 3 or more ramps where there is no more than 2 hours between each event. Ramp clusters were employed in hopes of only identifying ramps caused by mesoscale phenomena, as opposed to those induced by sporadic changes in surface winds (see Section 5). The sites that fit this criterion were Amarillo (AMA), Alva (AVK), Borger (BGD), Duncan (DUC), Dumas (DUX), Elkhart (EHA), Frederick (FDR), Gage (GAG), Guymon (GUY), Pampa (PPA), Oklahoma City (PWA), Perryton (PYX), and Woodward (WWR). Norman (OUN) also exhibited a ramp cluster but was neglected by this particular case study as it occurred much earlier in the forecast period than the others.

The period of study was divided into two forecast periods for comparison purposes. During the first forecast period, a total of 75 ramp events occurred and 24 were forecasted, 12 of which were hits. This is drastically different from the following forecast period (0900 UTC 19 June – 0900 UTC 20 June), which experienced 49 ramps and 13 forecasts; none of the forecasts were hits. Table 2 shows the contingency table used to compile statistics for the first forecast period. This period exhibited a PSS of .107 and a ramp capture of .16. These values are approximately four times greater than the overall statistics presented in Section 4.2. On the other hand, the second forecast period (depicted in Table 3), had a PSS of -.050 and ramp capture of 0.

		Observation	
		Yes	No
Forecast	Yes	Hits	False Alarms
		12	12
	No	Misses	Correct Nulls
		63	203

Table 2: Contingency table representing ramp event forecasts for the 13 sites during the first forecast period of the case study.

		Observation	
		Yes	No
Forecast	Yes	Hits	False Alarms
		0	13
	No	Misses	Correct Nulls
		49	248

Table 3: Contingency table representing ramp event forecasts for the 13 sites during the second forecast period of the case study.

After observing such a large number of ramp event misses in the overall data, the authors began questioning what factors, if any, determined whether or not the model would be able to predict a large change in wind power output. This case study provided an excellent opportunity to attempt to discern this, as there was an extreme difference between the model's performance during the first and second forecast period. The graph in Figure 8 attempts to distinguish between hits and misses in terms of differences in percent capacity of power output between 0900 UTC 18 June and 0900 UTC 20 June. However, the significant overlap between the two box-and-whisker plots indicates that there is no apparent difference between the magnitudes of ramp hits and ramp misses.

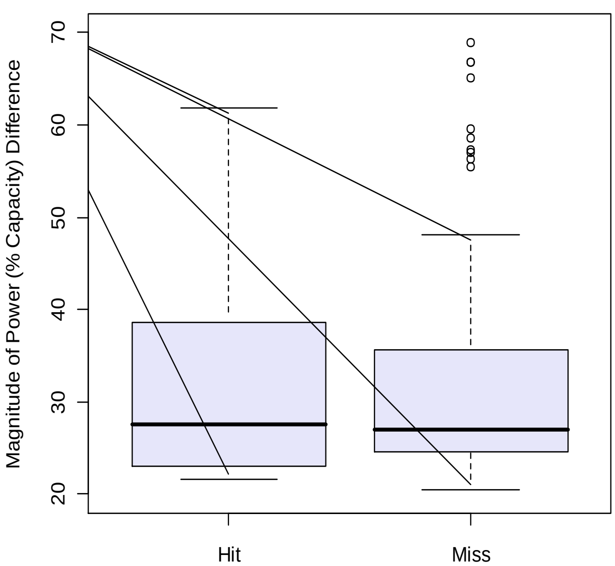


Fig. 8: Box-and-whiskers plots comparing the magnitudes of the power (% capacity) differences of the hits and the misses

The timing of each event was also considered, in order to ascertain if there was any sort of temporal pattern to the forecast hits in comparison to the misses. The scatter plot in Figure 9 shows that each of the 12 hits occurred during the 8-hr period spanning 23Z 18 June and 7Z 19 June. Figure 10 depicts the sites of each forecast hit. Archived weather maps accessed via the Internet from the Plymouth State School of Weathercasting, SPC, and NCEP, indicated that a cold front crossed through the western region of the domain of study (primarily the Texas and Oklahoma panhandles) during the early hours (UTC) of 19 June (Figure 11). The hits correspond to this frontal boundary, both in timing and location, leading to the conclusion that the presence of the cold front could have had some impact on the model's improved forecast accuracy during the first forecast period. Further study is required to determine whether the 3-km WRF resolves front-induced ramps better than other causes, or if this instance is simply a unique case.

4. LIMITATIONS OF METHODOLOGY

After dissecting this research's findings, the authors believe that these results do not necessarily indicate a problem with the 3-km WRF model, but rather illuminate limitations within some of the aspects of the methodology. These issues became clear toward the conclusion of the study. While there was not time to rectify these complications during the 10-week research project, they will no doubt be resolved in future model verification work.

Figure 12a suggests that there was no significant model wind speed bias; in other words, the model did not consistently over- or under-forecast hourly wind speeds. Instead, it was the lack of the model's variance in wind speeds that yielded the drastic difference between the number of forecasted and observed ramps

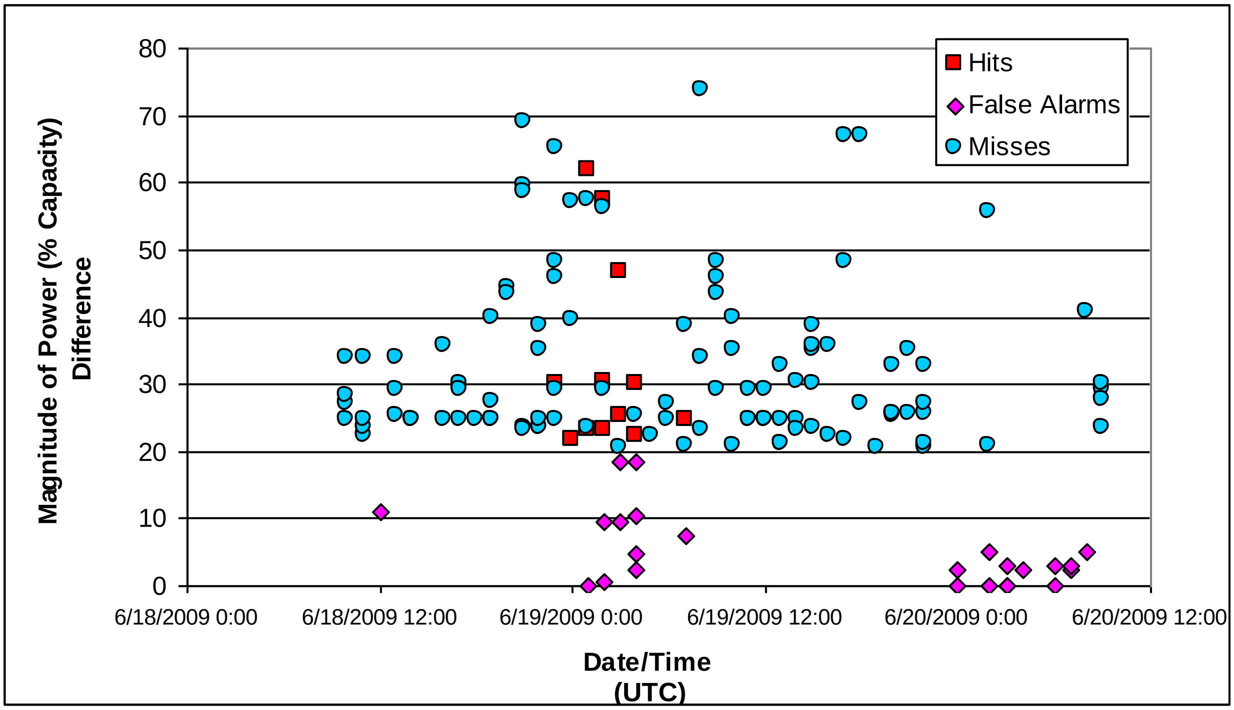


Fig. 9: Scatter plot comparing the hits, misses, and false alarms between 9Z 18 and 9Z 20 June.

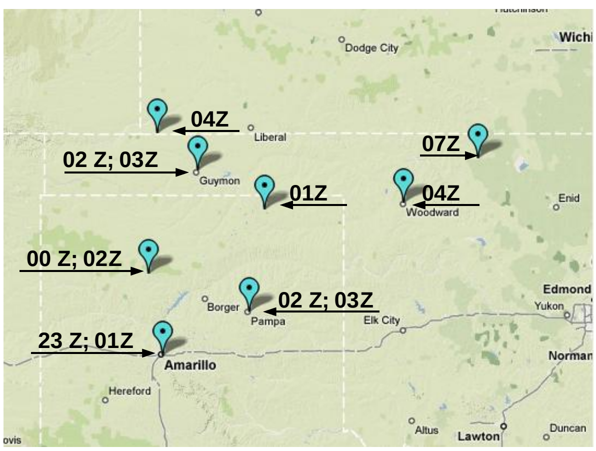


Fig. 10: A map of the sites that experienced hits between 23Z 18 June and 7Z 19 June. Map created with Google maps.

(Figure 12b). While surface winds are inherently erratic, the means by which the data was collected probably artificially increased this variability.

Although some stations recorded sub-hourly METAR observations (usually around 3 per hour), many only documented conditions around the top of the hour. To maintain consistency, only one reading was used per hour for each station. Still, wind readings from METAR data only represent a 2 minute average of the wind speed (OFCM, 2005). Thus, the study only sampled a small portion of the hour and then compared it to a WRF

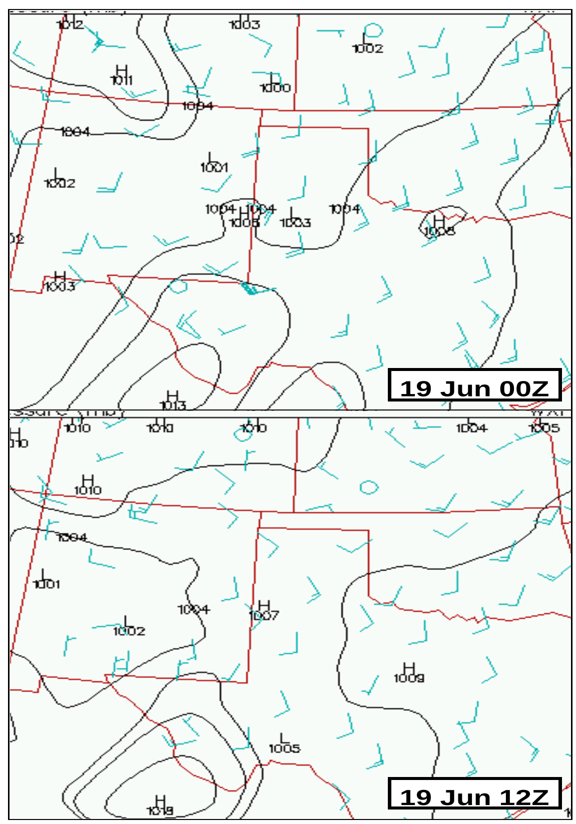
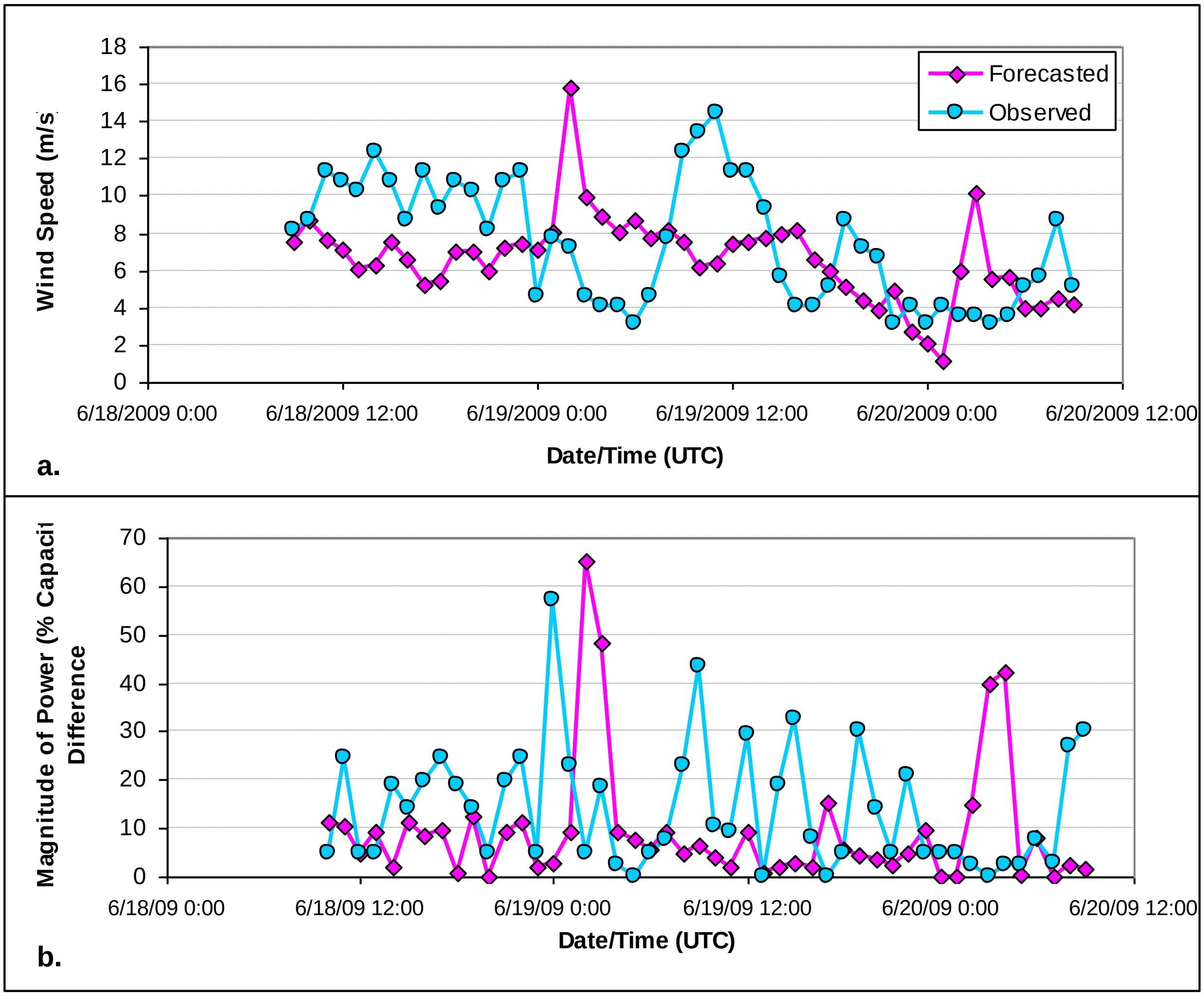


Fig. 11: Surface maps from 19 Jun 2009. Maps from Plymouth State Weather Center.



Figs. 12a,b: Meteograms depicting wind speeds (a) and power (% capacity) differences (b) as a function of time. These values were recorded at Perryton, TX (PYX) come from the 0900 UTC 18 June – 0900 UTC 20 June case study.

forecast. This is most likely the source of much of the variability witnessed in the results. In future studies, WRF point forecasts will probably be run for locations that provide ASOS, mesonet, or 80-m tower data so that 5-minute wind readings can be used. These will enable the generation of hourly and sub-hourly averages that can be compared to forecasts.

Another issue is that this study employed 2-m wind forecasts and observations, while wind turbine hub heights are between 50 and 80-m (NYSERDA 2005). Surface winds are highly variable; therefore, ramp events occur quite often a few meters above the ground. Winds at 80-m, on the other hand, exhibit more consistent behavior in comparison to those at the surface. Therefore, the number of observed ramp events recorded in this study presumably over-estimates the amount that occurred at turbine hub heights.

The ideal observational data set to be applied in this study would be from 80-m towers across the domain of study. Readings are recorded every 5-minutes and depict more relevant wind speed values for wind turbines. However, at the time this study began, this data was not accessible to WDT researchers. Hopefully, 80-m tower observations will soon become available so that this study can be repeated and most likely yield more accurate results.

5. CONCLUSIONS

This study used the 3-km WRF model to forecast wind ramp events in the Southern Plains between 0900

UTC 4 June and 0900 UTC 8 July. These extreme changes in wind power output over a short period of time are of growing concern to the wind energy industry. Ramp events tended to peak 3 times each day (0800 UTC, 1400 UTC, and 1900 UTC), the maximum occurring in sites located in the Texas and Oklahoma panhandles. A case study that focused upon the period between 0900 UTC 18 June and 0900 20 June implies that the WRF may better resolve front-induced ramps as opposed to other causes; however, further study is required to validate this theory.

Still, due to several issues discovered with the methodology, it is quite difficult to arrive at any conclusive results for this study regarding the WRF's ability to detect ramp events. The study did not perceive any extreme model wind speed bias; instead, the difference between forecasted and observed ramps arose due to the significant variation in surface wind speeds observations. The primary conclusion that was attained is that the investigation needs to be repeated using a different source of observational data. Ideally, these observations would come from 80-m towers, although ASOS and/or mesonet data could also be sufficient. These observations (recorded at 5-minute intervals) could then be compiled into 15-minute averages and compared to 15-minute WRF forecasts. This method of research is anticipated to yield a more accurate depiction of the performance of the 3-km WRF.

6. FUTURE WORK

Once the methodological issues are resolved, several additional questions have the potential to be answered. For example, a more expansive data set (at least a year's worth) would provide a more complete climatology of ramp events. Research should continue to examine how these events vary spatially and temporally. Future studies should also investigate how well the WRF detects ramp-up verses ramp-down events (defined by Freedman, et al. [2008]), opposed to using the magnitude of the power change as in this study. By dividing ramps in these categories, a forecast would not be considered a hit unless it forecasted a ramp-up event and the observed power increased by 20% or it forecasted a ramp-down event and the observed power decreased by 20%. This method should provide researchers with a more accurate understanding of the model's performance.

Using the adjusted methodology and altered definition of a hit, future studies may also want to investigate the distinguishing factor between forecast hits, false alarms, and misses. For instance, does the model predict more extreme differences in power than those around 20%? Or does it better resolve ramps induced by a particular meteorological factor? The authors of this paper encourage future work to address these questions, among many others, in order to provide the wind energy industry with a complete understanding of ramp events and present it with a forecasting system that can precisely detect abrupt changes in wind power output.

7. AKNOWLEDGEMENTS

This material is based upon work done through the (OU/CAPS) Research Experiences for Undergraduates (REU) program, which is supported by the National Science Foundation under Grant No. ATM-0648566. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would also like to thank the Plymouth State Weather Center, the Storm Prediction Center, and the National Centers for Environmental Prediction for access to their online weather map archives, as well as the National Climatic Data Center for access to their online METAR observation archives.

8. REFERENCES

AWEA, 2009: 20% energy by 2030. Available at www.awea.org.

Francis, N., 2008: Predicting Sudden Changes in Wind Power Generation. *North American Windpower*.

Freedman, J., M. Markus, and R. Penc, 2008: Analysis of West Texas wind plant ramp-up and ramp-down events. *AMS Truewind Report*. 250-278.

Greaves, B., J. Collins, J. Parkes, and A. Tindal, 2008: Temporal forecast uncertainty for ramp events.

Goggin, M., 2009: 20% wind energy by 2030: wind, backup power, and emissions.

Liu. Y., A. Bourgeois, T. Warner, S. Swerdlin, and W. Yu, 2006: An Update on "Observation-Nudging"-Based FDDA for WRF-ARW: Verification using OSSEs and Performance of Real-Time Forecasts. *Extended Abstracts, 2006 WRF Workshop,* Boulder, CO, June 2006.

Jolliffe, I. and D. Stephenson, editors, 2003: *Forecast Verification: a Practitioner's Guide in Atmospheric Science*. 1st ed. West Sussex: John Wiley & Sons Ltd. 44-51.

NESERDA, 2005: Wind turbine technology overview.

OFCM, 2005: Federal Meteorological Handbook, No. 1: Surface Weather Observations and Reports. 5.1-5.3. Available from http://www.ofcm.gov/fmh-1/fmh1.htm.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X. Huang, W. Wang, and J. G. Powers, 2008: A description of the Advanced Research WRF Version 3. *NCAR Tech Notes-475+STR*.

Zack, J., 2008: Facilitating higher levels of wind power penetration.