Evaluation of WRF-Predicted Near-Hub-Height Winds and Ramp Events over a Pacific Northwest Site with Complex Terrain

QING YANG, LARRY K. BERG, MIKHAIL PEKOUR, JEROME D. FAST, AND ROB K. NEWSOM
Pacific Northwest National Laboratory, Richland, Washington

MARK STOELINGA
3TIER, Inc., Seattle, Washington

CATHERINE FINLEY
WindLogics, Inc., Grand Rapids, Minnesota

(Manuscript received 2 October 2012, in final form 19 March 2013)

ABSTRACT

One challenge with wind-power forecasts is the accurate prediction of rapid changes in wind speed (ramps). To evaluate the Weather Research and Forecasting (WRF) model’s ability to predict such events, model simulations, conducted over an area of complex terrain in May 2011, are used. The sensitivity of the model’s performance to the choice among three planetary boundary layer (PBL) schemes [Mellor–Yamada–Janjić (MYJ), University of Washington (UW), and Yonsei University (YSU)] is investigated. The simulated near-hub-height winds (62 m), vertical wind speed profiles, and ramps are evaluated against measurements obtained from tower-mounted anemometers, a Doppler sodar, and a radar wind profiler deployed during the Columbia Basin Wind Energy Study (CBWES). The predicted winds at near–hub height have nonnegligible biases in monthly mean under stable conditions. Under stable conditions, the simulation with the UW scheme better predicts upward ramps and the MYJ scheme is the most successful in simulating downward ramps. Under unstable conditions, simulations using the YSU and UW schemes show good performance in predicting upward ramps and downward ramps, with the YSU scheme being slightly better at predicting ramps with durations longer than 1 h. The largest differences in mean wind speed profiles among simulations using the three PBL schemes occur during upward ramps under stable conditions, which were frequently associated with low-level jets. The UW scheme has the best overall performance in ramp prediction over the CBWES site when evaluated using prediction accuracy and capture-rate statistics, but no single PBL parameterization is clearly superior to the others when all atmospheric conditions are considered.

1. Introduction

One of the challenges with wind-power forecasts is the accurate prediction of ramps, which are rapid increases or decreases (for the sake of brevity hereinafter referred to as up-ramps and down-ramps, respectively) in generated power. Wind ramps pose challenges to power-system operators for maintaining grid reliability (Wan 2011), and large wind ramps also are important in managing the electric market (Cutler et al. 2007). Ramps with longer durations (from 1 h to several hours) have a larger impact on electrical-system operations than do those with shorter durations (Wan 2011). Despite their importance to the industry, ramps are notoriously difficult to predict.

This study evaluates the Weather Research and Forecasting (WRF; Skamarock et al. 2008) model’s ability to predict wind ramps over a site with complex terrain using in situ and ground-based remote sensing data from the Columbia Basin Wind Energy Study (CBWES; Berg et al. 2012). In WRF, planetary boundary layer (PBL) parameterizations play an important role in simulating low-level winds and PBL structures (Berg and Zhong...
and are used to parameterize subgrid vertical fluxes of momentum, heat, and moisture. Three PBL parameterization schemes have been selected to test the sensitivity of the WRF model’s performance to the choice of PBL schemes: Mellor–Yamada–Janjić (MYJ; Janjić 1994), University of Washington (UW; Bretherton and Park 2009), and Yonsei University (YSU; Hong et al. 2006). The MYJ and YSU PBL schemes were selected because they are popular turbulent kinetic energy (TKE)-based and non-TKE-based PBL schemes, respectively, and are widely used within the WRF community for simulating boundary layer processes [e.g., Berg et al. (2013) for MYJ and Yang et al. (2012) for YSU]. The YSU PBL scheme is a first-order nonlocal scheme that is characterized by nonlocal mixing within the PBL, local mixing at the PBL top, and explicit treatment of an entrainment layer at the PBL top (Hong et al. 2006). The MYJ scheme uses local closure with prognostic TKE. The UW scheme recently has been implemented in regional (e.g., version 3.3.1 of WRF) and global (e.g., version 5 of the Community Atmosphere Model) models. This scheme is TKE based, and it is characterized by the use of moist-conserved variables, an explicit entrainment closure, downgradient diffusion of momentum, and conserved scalars within turbulent layers (Bretherton and Park 2009). The new features in the UW scheme are likely to improve boundary layer wind predictions.

The goal of this research is to characterize the ramp occurrence over the CBWES site that is within an operating wind farm, to evaluate the WRF model’s capability in ramp prediction, and to test the sensitivity of the model’s performance to the choice of PBL schemes. Results of this study are also intended to provide recommendations to the wind-energy community regarding the choice of PBL schemes in WRF in areas of complex terrain.

2. Data and model descriptions

a. Observations

The CBWES was designed to help to elucidate complex flow patterns by providing measurements of winds within several hundred meters above the surface. The CBWES site (45.955°N, 118.688°W; Fig. 1) is located in an area of complex terrain on a northeast-facing slope of a long ridgeline near the Washington–Oregon border within the Stateline Wind Energy Center. In addition to an existing tower with propeller and vane anemometers [31, 44, and 62 m above ground level (AGL)] that were installed and maintained by the Bonneville Power Administration, instruments deployed during the CBWES included a Vaisala, Inc., 915-MHz radar wind profiler (RWP), a Scintec AG Doppler sodar, and Applied Technologies, Inc., SATI/3K ultrasonic anemometers.
b. Simulations

The RWP provided measurements from 146 to 1462 m in height and was configured to use only its low-power setting to apply a range-gate spacing of 57 m. Because of the potential of ground clutter from the radio tower, the RWP was operated in three-beam mode with the following beam orientations: 1) vertical, 2) tilted 23.6° from vertical toward direction 275°, and 3) tilted 23.4° from vertical toward direction 5°. Note that the lowest range gate (89 m AGL) was contaminated and was therefore discarded. No corrections were applied to the data to account for the complex terrain in the site’s vicinity. The sodar was operated using 10-m range-gate spacing and a maximum range of 400 m, and an averaging period of 15 min was used. Only a very small portion of the sodar data was used in this study because of the relatively poor data quality, most likely caused by ambient noise at the site and/or related to technical issues with the sodar itself. Two sonics were mounted near the top of the tower (62 m) and at 30 m AGL, respectively, providing measurements of heat flux every 30 min. Two corrections were applied to the sonic measurements: 1) crosswind contamination and 2) a two-dimensional coordinate rotation. The sonic heat flux data were not corrected to account for humidity. Therefore, they are only rough estimates of the sensible heat flux (SHF). More details about the CBWES, the instrumentation, and data processing can be found in Berg et al. (2012).

Composite wind profiles with a temporal resolution of 30 min were created by combining wind measurements from tower anemometers (averaged from 10-min data), sodar, and RWP. Because of frequent periods of missing sodar data, wind speed measurements from the tower were used for the ramp analysis. SHF measurements from the sonic anemometer at the 62-m height were used to determine static stability (SHF < −5 W m−2 is considered to be stable, and SHF > 10 W m−2 is considered to be unstable). As expected, in both model simulations and observations, stable conditions occur primarily at night and unstable conditions occur primarily during the day.

b. Simulations

The WRF model was configured to use the Goddard shortwave radiation scheme (Chou and Suarez 1994) and the Rapid Radiative Transfer Model for longwave radiation (Mlawer et al. 1997). For the current version of WRF, the surface-layer options generally are tied to particular boundary layer schemes. The MYJ and UW PBL schemes were run with the Eta Model (Janjic 1996, 2002) surface-layer scheme, and the YSU scheme was run with the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) similarity-theory surface-layer scheme (Skamarock et al. 2008). Three sets of simulations, denoted by SYJ, SUW, and SYSU, were conducted using the MYJ, UW, and YSU PBL schemes. WRF, version 3.3.1, was initially used for all simulations. There were significant modifications to the YSU scheme in the recently released version 3.4.1 of WRF that are likely to be relevant to the results in this study. Therefore, the original SYSU set was replaced with a new set of simulations using WRF, version 3.4.1, with the newly modified YSU scheme. The modifications to the YSU scheme from version 3.3.1 include 1) an increase in Prandtl number for unstable conditions, which will result in weaker mixing during daytime, and 2) weakened mixing in stable conditions during nighttime by reducing mixing length scale αs and allowing it to decrease with height (Jiménez et al. 2012).

Land surface models (LSMs) are used in WRF to compute heat and moisture fluxes over the land surface. As such, they also influence the low-level wind prediction in the model. The four-layer “Noah” LSM (Chen and Dudhia 2001) was used for the three main simulations. The effects of changing the LSM were tested with 10-day sensitivity simulations using the six-layer Rapid Update Cycle (RUC; Skamarock et al. 2008) LSM and the two-layer Pleim–Xiu (PX) LSM (Pleim and Xiu 1995; Xiu and Pleim 2001) with selected PBL schemes. The model domain (Fig. 2) was centered on the CBWES site. The three domains had respective horizontal spacings of 12, 4, and 1.3 km. The model was configured with 55 vertical layers, providing vertical grid spacing of approximately 15 m in the lowest 200 m AGL. The terrestrial data used for the two nested-domain simulations have a resolution of 30 arc s (~0.9 km). The horizontal resolution of the finest domain was chosen as a compromise between wanting to capture as much terrain complexity as possible and honoring the assumption in PBL schemes that boundary layer eddies are unresolved on the grid. WRF was initialized at 0000 UTC each day and employed a 36-h simulation period. The first 12 h of each simulation were spinup periods and were excluded from the analyses. The model output frequency is every 30 min. Lateral boundary conditions for the outer domain and initial conditions were obtained from the North American Regional Reanalysis dataset with 3-h temporal resolution and 32-km horizontal resolution. Simulations were conducted for May of 2011, which was the only period during which both RWP and sonic measurements on the 62-m-tall tower were available. In addition, May is one of the months with the highest ramp counts (shown later) over the site.
c. Wind-power calculation

The power $P$ generated from the wind is theoretically proportional to the swept area of the rotor, air density, and third power of the wind speed. Actual turbines have a “power curve” (output power as a function of wind speed $V$) that does not follow a simple $V^3$ dependency, however. In this study, we employed a typical wind turbine power curve, as illustrated in Fig. 3. When wind speed is lower than the cut-in wind speed of $3.0 \text{ m s}^{-1}$, power production is zero. Above the cut-in speed, power production increases at nearly $V^3$ dependency until reaching full capacity at a rated wind speed of $13 \text{ m s}^{-1}$. When the wind speed exceeds the cut-out speed of $22 \text{ m s}^{-1}$, the turbine is shut down. For simplicity, we assume a constant air density. Consistent with industry practice, we also express power in terms of power capacity factor (PCF), which is power normalized by the turbine’s full-rated power capacity $P_c$; that is, $\text{PCF} = P/P_c$.

d. Ramp definition and statistical evaluation metrics

A range of ramp definitions has appeared in the literature (Bossavy et al. 2012; Gallego et al. 2012). In this study, two ramp definitions are selected, and they are defined below.

1) DEFINITION 1: 2-HOURLY RAMPS

Definition 1 is marked by a decrease or increase in power within a 2-h span and larger than 50% of the power at the beginning of the period, or

\[
\frac{\max[\text{PCF}(t+2) - \text{PCF}(t)]}{\text{PCF}(t)} \geq 50\%
\]

where $t$ is the time in hours. When the initial power is small, the percentage change in power could be large—even with a small net change in the power. Thus, additional constraints are used. When the PCF is less than 0.2 at the starting hour, an increase in PCF of 0.3 or larger is required to be classified as an up-ramp. It is required that a down-ramp event have an initial power PCF > 0.3. If both hours from $t$ to $t + 2$ and from $t + 1$ to $t + 3$ are identified as up-ramps by definition, then the period...
from $t$ to $t + 3$ is considered to have a single up-ramp event occurring that lasts for 3 h. Similar rules are applied to the down-ramps. Thus, ramp duration is not limited to 2 h.

2) DEFINITION 2: HOURLY RAMPS

Hourly step changes are changes in power between two consecutive hours. Ramps are defined as hourly step changes of 15% or more of the total capacity:

$$|PCF(t + 1) - PCF(t)| \geq 0.15.$$ 

\textbf{e. Statistical evaluation metrics}

Two statistics, prediction accuracy (PA) and capture rate (CRt), defined in a way that is similar to those in Greaves et al. (2009), are used to assess the WRF simulations:

$$PA = \frac{N_{\text{hit}}}{N_{\text{pred}}} \quad \text{and} \quad CRt = \frac{N_{\text{hit}}}{N_{\text{obs}}},$$

where $N_{\text{obs}}$, $N_{\text{pred}}$, and $N_{\text{hit}}$ are the numbers of observed, predicted, and correctly predicted (hit) ramps. If the time periods of the predicted and the observed ramps overlap each other, it is considered a hit. Note that in the literature, the PA and CRt sometimes are called the success ratio and probability of detection, respectively, and PA can be used to infer the false-alarm rate $(1 - PA)$.

3. Results

\textbf{a. Evaluation of mean wind predictions}

The assessment of model-predicted mean wind speeds and directions near the hub height is based on measurements from tower anemometers at 62 m during May of 2011. Sixty-two meters is the closest measurement to the typical hub heights of the wind turbines in the area. During May of 2011, the observed mean wind speed at the 62-m height was 7.9 m s$^{-1}$. Predicted wind speeds at this height have nonnegligible negative mean biases under stable conditions. The biases and ±2σ bounds (95% confidence intervals; σ is standard deviation) are $-0.9 \pm 0.18$, $-1.0 \pm 0.18$, and $-0.8 \pm 0.18$ m s$^{-1}$ for SMYJ, SUW, and SYSU, respectively. Under unstable conditions, however, the biases are either very small ($0.2 \pm 0.18$ m s$^{-1}$ for SMYJ) or not statistically significant at the 5% significance level ($-0.1 \pm 0.17$ and $-0.0 \pm 0.17$ m s$^{-1}$ for SUW and SYSU, respectively). As shown in Fig. 4, predicted wind speed biases strongly depend on wind speeds. For simulations with all three PBL schemes, low wind speeds (<3 m s$^{-1}$) tend to be overpredicted, and wind speeds up to ~13 and ~9 m s$^{-1}$ at the 62-m height have relatively small prediction biases under stable and unstable conditions, respectively. Increasing underpredictions are seen at high wind speeds with more negative biases under unstable compared to stable conditions. As shown in the histograms (top panel of Fig. 4), the wind speeds are much higher under unstable conditions than under stable conditions, with their peak frequencies located at 13 and 3 m s$^{-1}$, respectively. The correlation coefficients between the observed and predicted wind speeds are 0.69, 0.71, and 0.71 for SMYJ, SUW, and SYSU, respectively, for May of 2011. At the CBWES site, the predominant wind direction at 62-m altitude is south-westerly (at 84% and 55% frequencies under unstable and stable conditions, respectively; bottom panel of Fig. 4), and mean winds from the southwest are much stronger (9.9 m s$^{-1}$) than winds from other directions.
Simulated wind directions have overall mean biases of $\sim$20° with all three schemes. The variation of biases in predicted wind directions with observed wind direction is likely associated with terrain representation in the model.

The predicted mean diurnal cycles of wind speed also are assessed by comparing with measurements made at 62 (anemometers) and 200 (RWP) m AGL. As shown in Fig. 5, wind speeds at both heights have distinct diurnal cycles with low wind speeds during the day and high wind speeds at night, peaking at around 0400–0500 local standard time (LST). In simulations with all three schemes from afternoon to several hours after sunset ($\sim$1200–0200 LST), there is particularly good agreement with observations as the predicted means are generally bounded by the 95% confidence interval of the observed values. The more distinct underprediction is seen at 62-m height for SYSU and SUW. Simulation SYSU underpredicts wind speeds in the early morning (0400–0800 LST at 62 m), and the SUW underpredicts over a much longer period (0600–1200 LST) in the morning at 62 m. The rapid changes in wind speed around sunrise and sunset also correspond to respective peak down- and up-ramp occurrences during a day (not shown).

b. Ramp statistics

Figure 6 depicts the annual cycle of ramp occurrences using tower observations (propeller and vane anemometers) from July of 2006 to June of 2011 and the 2-hourly ramp definition (definition 1). Ramps occur most frequently in spring with 21–24 up-ramp events per month as compared with only 11 and 13 per month in January and February, respectively. The down-ramp frequencies follow a similar pattern, with a slightly higher (by 2.6 per month) frequency. The larger ramp frequency in March–June when compared with winter likely results from ramps associated with non-synoptic-scale forcing, such as lower-atmosphere or local-surface thermal heterogeneity (Kang et al. 2012).

The WRF’s ability to predict ramps is first assessed using definition 1 (Fig. 7). In May 2011, there were 24

![Fig. 5.](image_url) The diurnal variations of the observed and simulated wind speeds at heights of (bottom) 62 and (top) 200 m. The observed wind speeds (obs; black lines) at 200- and 62-m heights are from RWP measurements and tower anemometers, respectively. The gray-shaded area shows the 95% confidence intervals of the mean wind speed $\left[\pm 2\sigma_{\text{mean}} = 2\left(\sigma^2/n\right)^{1/2}\right]$, where $\sigma^2$ is the variance and $n$ is the number of data included for the averaging.}

![Fig. 6.](image_url) Annual cycle of monthly numbers of up-ramps and downramps from tower observations from July 2006 to June 2011 and ramp definition 1 at 62 m over the CBWES site. The error bars indicate ±1 std dev.
up-ramps and 31 down-ramps observed. The SUW is markedly better in up-ramp predictions under both stable and unstable conditions as measured by the PA and CRt scores. They are higher by 12% and 6% in PA and by 21% and 9% in CRt when compared with SMYJ and SYSU, respectively. During down-ramps, there is no one PBL scheme that consistently has better ramp predictions across all classification categories. During down-ramps, the SUW has the highest overall CRt score (higher by ~7% and 10% relative to SMYJ and SYSU, respectively), and SYSU has the highest overall PA score (higher by ~5%–9%). The SMYJ performs the best for down-ramp predictions under stable conditions, with the highest PA (higher by 7%–9%) and CRt (tied with SUW at 80%) scores in this category. The SYSU best predicts the magnitude of up-ramps (91% of the observed magnitude as compared with 87% and 77% in SUW and SMYJ, respectively), and SUW better predicts down-ramps (97% of observed strength vs 91% and 93% for SMYJ and SYSU, respectively). When WRF’s performance is considered in terms of stability categories, the SUW has the best agreement with observations in the magnitude of the both up-ramp (88% of observed magnitude vs 77% and 84% for SMYJ and SYSU, respectively) and down-ramp (101% vs 92% and 93% for SMYJ and SYSU, respectively) under stable conditions. Under unstable conditions, SMYJ agrees the best with the observations in predicting the strength of down-ramps (90.8% of observed magnitude vs 84.8% and 77% for SMYJ and SUW, respectively), and the strength of up-ramps is better predicted by SYSU (by 5%–18%) under the same stability category. The SUW outperformed SMYJ and SYSU in up-ramp predictions for both stability categories, and it is also the most consistent in overall performance across ramp and stability classifications.

There are also variations in the simulated ramps that are due to the use of different LSMS. The sensitivity simulations are denoted as SUW-Noah (a subset of SUW), SUW-RUC, SUW-PX, SMYJ-Noah (a subset of SMYJ), and SMYJ-RUC with the combination of PBL schemes and LSMS shown as subscripts. The SUW-Noah simulation performs much better (higher by ~40% in PA and CRt) than SMYJ-Noah in up-ramp predictions, and changing the LSM does not reverse that finding. When using the RUC LSM, SUW-RUC is higher by ~16% and ~14% in PA and CRt, respectively, during up-ramps when compared with SMYJ-RUC. During down-ramps, SMYJ-Noah and SUW-Noah perform similarly with the same CRt (both 66%) and a 4% difference in PA (46% and 50% for SMYJ-Noah and SMYJ-Noah, respectively) for the 10-day period. With the RUC LSM, SUW-RUC is slightly higher in PA and CRt (by 6% and 11%, respectively) in comparison with SMYJ-RUC during down-ramps. Therefore, our sensitivity tests using different LSMS essentially yield consistent results in that SUW performs better than SMYJ for ramp predictions measured by CRt and PA, especially in the...
Table 1. Observed and simulated magnitudes of hourly step changes defined as the difference in power capacity factor in two consecutive hours, along with the PA and CRt in terms of hourly ramps. The simulations are conducted with the MYJ, UW, and YSU PBL schemes. The step changes have units of “per hour.”

<table>
<thead>
<tr>
<th>Effect/statistic/condition</th>
<th>Obs</th>
<th>SMYJ</th>
<th>SUW</th>
<th>SYSU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive steps: stable</td>
<td>$0.131 \pm 0.175$</td>
<td>$0.145 \pm 0.187$</td>
<td>$0.130 \pm 0.165$</td>
<td>$0.157 \pm 0.185$</td>
</tr>
<tr>
<td>Positive steps: unstable</td>
<td>$0.085 \pm 0.124$</td>
<td>$0.126 \pm 0.167$</td>
<td>$0.096 \pm 0.146$</td>
<td>$0.107 \pm 0.158$</td>
</tr>
<tr>
<td>Up-ramps: PA; stable</td>
<td>—</td>
<td>$14.3%$</td>
<td>$23.5%$</td>
<td>$10.0%$</td>
</tr>
<tr>
<td>Up-ramps: PA; unstable</td>
<td>—</td>
<td>$11.5%$</td>
<td>$11.1%$</td>
<td>$13.6%$</td>
</tr>
<tr>
<td>Up-ramps: CRt; stable</td>
<td>—</td>
<td>$16.1%$</td>
<td>$25.8%$</td>
<td>$9.7%$</td>
</tr>
<tr>
<td>Up-ramps: CRt; unstable</td>
<td>—</td>
<td>$21.4%$</td>
<td>$14.3%$</td>
<td>$21.4%$</td>
</tr>
<tr>
<td>Negative steps: stable</td>
<td>$-0.133 \pm 0.185$</td>
<td>$-0.129 \pm 0.167$</td>
<td>$-0.120 \pm 0.138$</td>
<td>$-0.118 \pm 0.165$</td>
</tr>
<tr>
<td>Negative steps: unstable</td>
<td>$-0.108 \pm 0.117$</td>
<td>$-0.100 \pm 0.099$</td>
<td>$-0.101 \pm 0.113$</td>
<td>$-0.089 \pm 0.090$</td>
</tr>
<tr>
<td>Down-ramps: PA; stable</td>
<td>—</td>
<td>$13.8%$</td>
<td>$27.6%$</td>
<td>$14.3%$</td>
</tr>
<tr>
<td>Down-ramps: PA; unstable</td>
<td>—</td>
<td>$30.3%$</td>
<td>$32.1%$</td>
<td>$26.6%$</td>
</tr>
<tr>
<td>Down-ramps: CRt; stable</td>
<td>—</td>
<td>$16.7%$</td>
<td>$33.3%$</td>
<td>$16.7%$</td>
</tr>
<tr>
<td>Down-ramps: CRt; unstable</td>
<td>—</td>
<td>$26.3%$</td>
<td>$23.7%$</td>
<td>$21.05%$</td>
</tr>
</tbody>
</table>

The individual scheme’s sensitivity to different LSMs also was assessed for the UW PBL scheme. During both up-ramps and down-ramps, the change in overall PA and CRt that is due to the use of different LSMs is not large (±10%) except for CRt during up-ramps, which is ~58%, 30%, and 45% for SUW-NOAH, SUW-RUC, and SUW-PX, respectively.

Model performance may vary with ramp definition, so simulations were also assessed using an hourly ramp (definition 2). Step changes provide a simple measure and the first-order estimation of wind-power ramps (Wan 2011). For positive step changes, $S_{UW}$ has the smallest errors in mean and standard deviation of the magnitudes of changes, and its PA and CRt are ~10% higher than those of $S_{MYJ}$ and $S_{SYSU}$ under stable conditions (Table 1)—consistent with the results using definition 1. The $S_{SYSU}$ also consistently has higher CRt associated with up-ramp predictions during unstable conditions. For $S_{MYJ}$, the better prediction of negative step-change strength and variation under stable conditions is consistent with the relatively good down-ramp predictions under stable conditions with definition 1. Under unstable conditions, the $S_{UW}$ and $S_{MYJ}$ have the highest PA and CRt, respectively, for down-ramp predictions. The $S_{UW}$ also has higher CRt associated with down-ramp predictions under stable conditions. The $S_{UW}$’s good overall performance as measured with step changes indicates that, relative to $S_{MYJ}$ and $S_{SYSU}$, $S_{UW}$ better predicts hourly-time-scale variations in wind speed.

c. Vertical profiles during ramp events

Mean wind profiles (Fig. 8) are used to further understand ramp events. For up-ramps under stable conditions (UP-STAB; Fig. 8a), the observed composite profiles show a low-level jet with a maximum located at ~200 m AGL. Note that in this study low-level jet is loosely defined as low-level wind maxima at an altitude of 200–300 m with strong shears below the maxima (wind speed below the nose of the maximum wind speed is smaller by, at least, 2 m s$^{-1}$). The $S_{SYSU}$ agrees the best with the mean observed profile below 200 m, followed by $S_{SUW}$. Both $S_{SYSU}$ and $S_{SUW}$ overpredict the height (~300–350 m) of the low-level jet maximum. Although $S_{MYJ}$ predicts a more realistic height of low-level jets, it overestimates wind speeds (by ~3 m s$^{-1}$ in the jet core) in the lowest kilometer under UP-STAB conditions. Vertical mean profiles of potential temperature and TKE (not shown) indicate that, under UP-STAB conditions, $S_{MYJ}$ has large TKE production near the surface relative to that of $S_{SUW}$. The predicted strong stratification and colder PBL air (relative to $S_{SUW}$ and $S_{SYSU}$) lead to reduced transfer of momentum, resulting in a lower stable boundary layer and the buildup of strong, low-level winds.

Under stable conditions, the observed up-ramps and low-level jets typically occur at night during periods of strong surface cooling and weak synoptic forcing. Some of those ramps (29%) occur around sunset (between 1800 and 2000 LST) when the surface undergoes rapid radiative cooling, becomes stable, and decouples from the air above, allowing supergeostrophic winds to develop over a period of several hours. Some up-ramps (24%) develop later at night and are influenced by large wind speeds aloft. The terrain’s complexity also could have affected low-level jet occurrences and may have induced occasional pressure gradients that caused up-ramps under stable conditions during the study period.

In past decades, progress in understanding and modeling the stable boundary layer has been extremely slow because of its associated small-scale features, caused by the suppressed mixing (Banta 2008). The large variability among profiles in $S_{MYJ}$, $S_{SUW}$, and $S_{SYSU}$ reflects the challenge in simulating stable boundary layers.

During up-ramps under unstable conditions (UP-UNSTAB; Fig. 8b), the observed and simulated mean
FIG. 8. Mean vertical profiles of wind speeds from observations and from WRF simulations with three PBL schemes for (left) stable and (right) unstable atmospheric stability conditions during (a),(b) up-ramps and (c),(d) down-ramps. The power ramps are identified using definition 2. Stable and unstable conditions are based on observed SHF from sonic measurements, with SHF > 10 W m$^{-2}$ for unstable conditions and SHF < −5 W m$^{-2}$ for stable conditions. The gray-shaded area shows the 95% confidence intervals of the mean profiles $[\pm 2\sigma_{\text{mean}} = 2(\sigma^2/n)^{1/2}]$, where $\sigma^2$ is the variance and $n$ is the number of profiles included for the averaging; $n = 24, 10, 4, \text{and } 35$ for (a)–(d), respectively. The dashed line marks the typical hub height of 80 m.
wind speeds increase with altitude until reaching a local maximum at a height of ~270 m. The $S_{\text{UW}}$ mean profile is the closest to the observed profile below the local maximum. The better agreement between the observed profiles using $S_{\text{UW}}$ and $S_{\text{YSU}}$ during up-ramps is consistent with the ramp statistics described in the previous section. Simulations with all three schemes fail to capture the decreased wind speed between the local maximum and the top of the profile during UP-UNSTAB conditions, however. It is likely related to the Noah LSM because $S_{\text{UW-RUC}}$ and $S_{\text{UW-PX}}$ show improvement in simulating this feature when compared with $S_{\text{UW-Noah}}$. Over the CBWES site, the up-ramps occurring under unstable conditions during the simulated period were sorted into a number of categories: 1) those influenced by a strong upper-level trough, 2) those with a low-level jet driven by synoptic forcing, 3) those associated with frontal passages, and 4) those influenced by small-scale variations that are probably thermally or terrain induced. Those ramps associated with strong large-scale forcings (such as a strong and moving trough) are generally predicted well, whereas ramps associated with small-scale variations are represented less well.

Down-ramps typically occur over the site in association with different scenarios: there are cases associated with solar heating near sunrise, weakening or shifting in position of synoptic systems, and decreasing of local pressure gradients. Among the 694 observed and simulated profiles under all conditions, 225 (38%) and 319 (53%) of them fall into the respective stable and unstable categories as based on measured SHF. The observed down-ramps occur 3 times more often under unstable conditions (6% of all observations) than under stable conditions (2% of all observations). For down-ramps that occur under stable conditions (Fig. 8c), the differences in simulated and predicted mean profiles are not statistically significant (at a significance level of 5%), with $S_{\text{UW}}$ being the closest to the observed mean profile. During down-ramps and unstable conditions (Fig. 8d), the mean profiles from $S_{\text{YSU}}$ and $S_{\text{MYJ}}$ agree well with observations. In contrast, the $S_{\text{UW}}$ predicts vertical variations of the mean profile reasonably well but systematically underestimates by ~1–2 m s$^{-1}$, which is accompanied by high TKE relative to $S_{\text{MYJ}}$ and slightly higher PBL air temperature and relatively small SHFs when compared with simulations using the other two PBL schemes.

The results shown in Fig. 8 are for periods with ramp events, and these cases represent a significant challenge for mesoscale models such as WRF. When all conditions (both ramps and no ramps) are considered, the agreement with observations is improved (figure not shown). Beyond the PBL scheme, the simulation of low-level winds in WRF is influenced by many model components, including the choice of LSM, dynamical core, radiation scheme, and boundary conditions from global reanalysis data. The complex terrain’s influence adds to the challenge in low-level wind simulations because there are biases due to small-scale, terrain-induced variations that cannot be resolved by the model and possible numerical errors associated with terrain-following coordinates in steep terrain (Lundquist et al. 2010). Another aspect worth noting is that the model output represents mean values in a model grid box, and, when one is doing comparisons with point measurements, there are potential biases contributed by the spatial variability within the grid box.

### 4. Conclusions

Although the predicted winds near hub height (62 m) have nonnegligible negative biases in monthly mean under stable conditions, the biases are very small under unstable conditions. Both observed and predicted low-level wind speeds over the site have distinct diurnal cycles. The predicted mean diurnal cycles show good agreement with observations in simulations with all three schemes from noon to ~1400 LST and noticeable underpredictions in both $S_{\text{YSU}}$ and $S_{\text{UW}}$ in the morning. Power ramps and hourly step changes consistently show that the simulation with the UW scheme better predicts up-ramps under stable conditions, with higher prediction accuracy and capture rates. The simulation with the MYJ scheme predicts down-ramps better than up-ramps. The $S_{\text{UW}}$ outperformed $S_{\text{MYJ}}$ and $S_{\text{MYJ}}$ in regard to both the frequency of events and their magnitude across ramp and stability classifications. Simulations with the three PBL schemes show the largest variability among them in the predicted mean wind speed profiles under stable conditions for ramp events. The high wind speed and large shear associated with low-level jets (~200 m in altitude) are frequently associated with up-ramps occurring under stable conditions. During up-ramps, all schemes overpredict the strength of the low-level winds, and simulations with UW and YSU overpredict the altitude of the low-level jet maximum as well. The overestimation in low-level jet strength with MYJ during up-ramps under stable conditions could be linked to its predicted higher surface-layer TKE production and stronger stratification. In simulations with all three schemes, better agreement with the mean observed wind speed profile is found during down-ramps. Although the best overall performance is obtained using the UW scheme, the results show that no single parameterization is clearly superior to the others when all atmospheric conditions are considered. The PBL scheme is but one aspect among many that
AUGUST 2013 YANG ET AL. 1763

affect low-level wind predictions for wind-energy applications within an entire regional modeling framework, and it requires further examination. Although the evaluated WRF performance is influenced by many model components within the regional framework, the large sensitivity of the ramp predictions to the choice of the PBL schemes in WRF under unstable conditions indicates a need for additional research into the development of PBL parameterizations targeting wind-energy applications.

Acknowledgments. We thank Julia Flaherty of Pacific Northwest National Laboratory (PNNL) for providing Fig. 1. We also thank Prof. Songyou Hong for the helpful discussion and two anonymous reviewers for their helpful comments and suggestions. This work was supported by the U.S. Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy. The RWP was provided by the DOE’s Atmospheric Radiation Measurement Program Climate Research Facility, and the tower anemometer data were provided by the Bonneville Power Administration. A portion of the research was performed using PNNL Institutional Computing. PNNL is operated by Battelle for the DOE under Contract DE-AC06-76RL01830.

REFERENCES


