JGR Atmospheres

RESEARCH ARTICLE
10.1029/2018JD029691

Parametric and Structural Sensitivities of Turbine-Height Wind Speeds in the Boundary Layer Parameterizations in the Weather Research and Forecasting Model

Ben Yang[1,2], Larry K. Berg[2], Yun Qian[2], Chen Wang[2,3], Zhangshuan Hou[2], Ying Liu[2], Hyeyum Hailey Shin[4], Songyou Hong[5], and Mikhail Pekour[2]

[1]CMA-NJU Joint Laboratory for Climate Prediction Studies, School of Atmospheric Sciences, Nanjing University, Nanjing, China, [2]Pacific Northwest National Laboratory, Richland, WA, USA, [3]College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [4]UCAR/Cooperative Programs for the Advancement of Earth System Science and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [5]Korea Institute of Atmospheric Prediction Systems, Seoul, South Korea

Key Points:
- MYNN parameters induce relatively weaker (stronger) impacts than YSU parameters on wind speeds during daytime (nighttime)
- YSU scheme has a potential to better match daytime wind observations, while MYNN generally better matches the observations at night
- Incompatible biases exist in different wind strengths, time of day, and locations due to structural inadequacy in both PBL schemes

Correspondence to:
L. K. Berg and Y. Qian,
larry.berg@pnnl.gov;
yun.qian@pnnl.gov

Citation:
Yang, B., Berg, L. K., Qian, Y., Wang, C., Hou, Z., Liu, Y., et al. (2019). Parametric and structural sensitivities of turbine-height wind speeds in the boundary layer parameterizations in the Weather Research and Forecasting model. *Journal of Geophysical Research: Atmospheres*, *124*, 5951–5969. https://doi.org/10.1029/2018JD029691

Received 21 SEP 2018
Accepted 27 MAR 2019
Accepted article online 2 APR 2019
Published online 17 JUN 2019

Abstract Structural and parametric problems associated with physical parameterizations are often tied together in weather and climate models. This study examines the sensitivities of turbine-height wind speeds to structural and parametric uncertainties associated with the planetary boundary layer (PBL) parameterizations in the Weather Research and Forecasting model over an area of complex terrain. The sensitivity analysis is based on experiments from two perturbed parameter ensembles using the Mellor-Yamada-Nakanishi-Niino (MYNN) and Yonsei University (YSU) PBL schemes, respectively. In each scheme, most of the intermember variances can be explained by a few parameters. Compared to the YSU parameters, the MYNN parameters induce relatively weaker (stronger) impacts on wind speeds during daytime (nighttime). The two schemes can overall reproduce the observed diurnal features of turbine-height wind speeds. Differences in the daytime wind speeds are evident between the two ensembles. The daytime biases exist even with well-tuned parameter values in MYNN, indicating the structural error. The YSU scheme better matches monthly mean daytime observations, partly due to the compensation among the biases in different wind strengths. Compared to YSU, MYNN generally better agrees with observations in both weak and strong wind conditions. However, the improvements accomplished for one condition by parameter tuning may degrade model performances for others, suggesting the relationships that link different conditions are not accurately represented in the parameterizations. Simulated biases due to structural errors are further identified by evaluating them for different time of day and locations. Ultimately, this study improves understanding of structural limitations in the PBL schemes and provides insights on further parameterization development.

1. Introduction

Currently, it is still computationally unaffordable to use weather or climate models at resolutions approaching the order of turbulent scale motions (i.e., 10–100 m) for operational applications. Therefore, subgrid parameterizations, including those for the planetary boundary layer (PBL) and surface layer, must be used to represent the statistical features of the turbulence processes and their mean effects on the grid-scale fields.

Many PBL schemes have been developed over the past decades, with some of them being routinely updated (e.g., Hong, 2010; Hong et al., 2006; Hong & Pan, 1996; Mellor & Yamada, 1974; Nakanishi & Niino, 2009; Pleim, 2007; Troen & Mahrt, 1986). These schemes feature different levels of complexity and can be based on structurally different assumptions or formulas regarding the representation of PBL processes. As a result, structural uncertainty in the physical parameterizations contributes to a large portion of the total model uncertainty (e.g., Giorgi & Mearns, 2002; Gupta et al., 2012; Hawkins & Sutton, 2009; Huang et al., 2013). Furthermore, the PBL schemes usually apply a large number of tunable parameters (e.g., Hong et al., 2006; Nakanishi & Niino, 2009; Nielsen-Gammon et al., 2010; Yang et al., 2017), whose values are typically calibrated based on limited measured fields or idealized large eddy simulations. Due to the existences of numerous uncertain parameters and the potential nonlinear interactions among them, sensitivity analysis frameworks that consist of a multiple-dimension sampling approach and/or surrogate-based variance decomposition have been applied to quantify the parameter uncertainty (e.g., Boyle et al., 2015; Guo et al.,

2014; Hou et al., 2012; Posselt et al., 2016; Qian et al., 2015; Yan et al., 2015; Yang et al., 2012, 2015; Zhao et al., 2013). In addition to improved understanding of uncertainty, these approaches can also be used to assist with the planning of new field studies or model developments (Qian et al., 2016).

There has recently been a huge growth in the wind-generated electricity around the world (Pimentel et al., 2002; National Renewable Energy Laboratory, 2008; Lu et al., 2009; Musgrove, 2010). Operational wind forecasting and wind resource characterization highly rely on the ability of mesoscale atmospheric models to represent the fine-scale flow features (Mahoney et al., 2012; Marjanovic et al., 2014; Pichugina et al., 2017; Siuta et al., 2017). It remains a challenge to accurately predict the winds at different forecast lead times due to the significant variability in the wind flows in the PBL, which ultimately influences forecasts of wind power to be integrated into the electric grid that can potentially have serious financial consequences (Banta et al., 2013, 2018; Marquis et al., 2011).

The impacts of tuning PBL parameter values or switching PBL scheme options on the simulated wind flow features have been separately examined by previous studies (e.g., Carvalho et al., 2012; Draxl et al., 2014; Nielsen-Gammon et al., 2010; Siuta et al., 2017; Yang et al., 2013; Yang et al., 2017). Structural and parametric problems, however, are often linked, which means the discrepancies in model outputs using different schemes are jointly contributed by the structural differences among them and parameter uncertainties within each scheme (Dai et al., 2017; McNeall et al., 2016). Therefore, conclusions regarding the strengths or limitations of different schemes may not be the same when different parameter values are applied in the schemes (e.g., Clark et al., 2008; Leung et al., 1999; Williamson et al., 2015).

Relative to the parametric issue, the structural problem is usually difficult to represent in a straightforward way because two schemes can differ from each other physically, mathematically, or computationally (Gupta et al., 2012). Model bias that is evident even when all parameter values have been optimized confirms the existence of structural errors in the scheme that cannot be explained by parameter uncertainties (Williamson et al., 2015). It is also common that all the model aspects can be tuned close to a specific observational target, but the improvement related to one target often corresponds to degraded performances in others, exhibiting a diversity of incompatible biases such as those of different variables or at different times or locations (Gupta et al., 2012; McNeall et al., 2016). This kind of structural error can occur when different model components are unrealistically linked and their relationships are structurally fixed, hardcoded, or only weakly sensitive to parameter perturbations, though each of these components individually could have a strong sensitivity and has the potential to be optimized.

In this study, we examine the sensitivities of simulated turbine-height wind speed to structural differences and uncertain parameters in the PBL parameterizations based on experiments using two perturbed parameter ensembles (PPEs) for two different PBL schemes. Our main purposes are to (1) compare the parametric sensitivities of turbine-height wind speeds in the two PBL schemes, (2) examine the structural sensitivity to the PBL parameterizations, and (3) identify the biases in the simulated wind speeds due to structural limitations in the schemes that cannot be explained by parameter uncertainty alone. Simulated biases related to structural limitations will also be explored when considering different wind strengths, time of day, and locations as the improvements for one target by parameter tuning may degrade model performances for others.

2. Data and Methodology

The model configuration and framework for parametric sensitivity analysis applied here are the same as those described by Yang et al. (2017) and Berg et al. (2019). Additional analyses are presented to investigate structural sensitivity and simulated biases due to structural error in the schemes.

2.1. WRF Model Configuration and Wind Field Sites

The advanced Weather Research and Forecasting (WRF) model version 3.6 is used (Skamarock et al., 2008) in this study. The WRF domain (Figure 1) focuses on a complex terrain region, which is centered at the Columbia Basin Wind Energy Study (CBWES; Berg et al., 2012) field site and including measurements of the subsequent second Wind Forecast Improvement Project. The horizontal grid spacing is 3.3 km in the horizontal and approximately 15 m in the vertical within the lower part of PBL (i.e., from surface to around 200 m). Initial and boundary conditions are derived from the 10-km WRF outputs that are dynamically

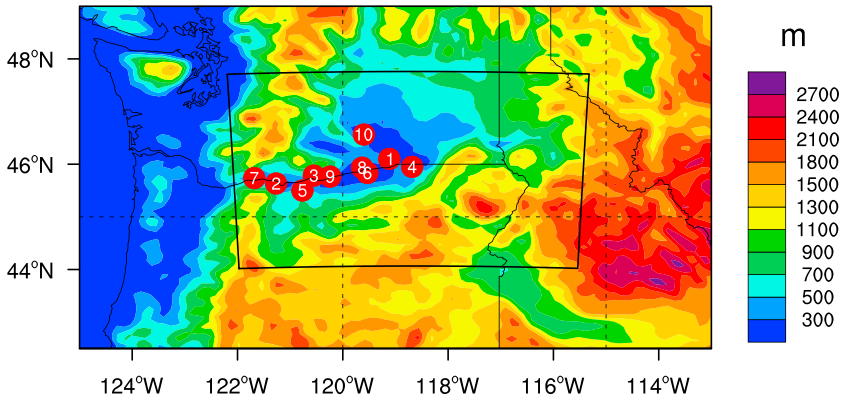

Figure 1. Weather Research and Forecasting domain (black box) with 3.3-km grid spacing. Terrain height is indicated by color (see legend on the right). The red solid circles denote the locations of field measurements or wind plants used in this study. The names of sites and corresponding heights of wind measurements are as follows: 1. Kennewick (36.6 m), 2. Seven Mile Hill (30.5 m), 3. Goodnoe Hills (59.4 m), 4. Butler Grade (62.5 m), 5. Wasco (30.5 m), 6. Chinook (50.0 m), 7. Augspurger (21.3 m), 8. Horse Heaven (21.3 m), 9. Roosevelt (21.3 m), and 10. Hanford (15.2 m). Columbia Basin Wind Energy Study site is collocated with the Butler Grade site (no. 4).

downscaled from the North American Regional Reanalysis (Mesinger et al., 2006) with grid nudging. The WRF model is integrated from 28 April to 31 May 2011, with the results during the month of May used for the analysis. For more details about the model configuration please refer to Yang et al. (2017).

To explore the structural errors in the schemes, we apply the wind speed measurements at 10 field sites in addition to the CBWES site, including nine stations that were part of the Bonneville Power Administration's Anemometer Network and the Department of Energy's Hanford site (see Figure 1 for site locations and heights of measurements above surface). The Butler Grade site (no. 4 in Figure 1) is colocated with the CBWES site near the center of the study region. At this site, the Bonneville Power Administration observations of standard meteorological variables made on a communications tower were augmented by sonic anemometers at two levels (30 and 62.5 m above ground). In addition, profiles of wind speed and wind direction were measured using a Doppler sodar and radar wind profiler during CBWES. Therefore, the model results at Butler Grade are a major focus of our analysis.

Table 1
Adjustable Parameters and Their Perturbed Ranges in the Mellor-Yamada-Nakanishi-Niino Planetary Boundary Layer Scheme

Parameter	Description	Default	Range
B_1	Constant for dissipation rate of TKE	24	12–36
D_f (*sqfac*)	TKE diffusion factor	2	1.5–4.5
Pr	Turbulent Prandtl number	0.74	0.5–2
C_3	Closure constant	0.34	0.33–0.50
C_5	Closure constant	0.2	0.1–0.3
γ_1 (*g*1)	Closure constant	0.229	0.1768–0.2395
α_1 (*alp*1)	Constant for calculation of the turbulence length scale (L_T)	0.23	0.115–0.345
α_2 (*alp*2)	Constant for calculation of the turbulence length scale (L_B)	0.65	0.5–1.0
α_3 (*alp*3)	Constant for calculation of L_B	3	2.5–7.5
α_4 (*alp*4)	Constant for calculation of the turbulence length scale (L_S)	20	20–100
α_5 (*cns*)	Constant for calculation of L_S	2.1	1.35–4.05
β	Exponent on equation of L_S	0.2	0.1–0.3

Note. Texts in parentheses in the first column is the parameter name used in the WRF source code. TKE = turbulent kinetic energy.

2.2. PBL Schemes and Perturbed Parameters

We examine 12 and 15 parameters in the Mellor-Yamada-Nakanishi-Niino (MYNN) level 2.5 and Yonsei University (YSU) PBL schemes, respectively. We chose the parameters that can potentially influence the model results by reviewing the existing literature, examining the scheme formulas, and/or consulting the scheme developers' suggestions (e.g., Andren & Moeng, 1993; Grachev et al., 2007; Hong, 2010; Hong et al., 2006; Hong & Pan, 1996; Kim & Mahrt, 1992; Mellor & Yamada, 1982; Nakanishi, 2001; Nakanishi & Niino, 2004, 2006, 2009; Noh et al., 2003; Schumann & Gerz, 1995; Stull, 1988; Troen & Mahrt, 1986; Vickers & Mahrt, 2004). The parameter ranges are determined based on the guidance mentioned above or, in cases with a lack of such guidance, by assigning a range of ±50% from the original value of the individual parameters, following Yang et al. (2017). The ranges specified for each of the parameters in the two PBL schemes are given in Tables 1 and 2.

The MYNN level 2.5 scheme (Nakanishi, 2001; Nakanishi & Niino, 2004, 2006, 2009) is a turbulent kinetic energy (TKE) scheme based on Mellor-Yamada scheme (Mellor & Yamada, 1974, 1982) but with modified formulas for the turbulent length scales and reestimated values for closure constants relative to its precursor (Nakanishi, 2001). Twelve parameters in

Table 2
Same as Table 1 but for the Parameters in the Yonsei University Planetary Boundary Layer Scheme

Parameter	Description	Default	Range
α_s ($aphi5$)	Constant associated with $\psi_{h,m}$ at stable condition	5	4.5–7
$\alpha_{u,h}$ ($aphi16$)	Constant associated with ψ_h at unstable condition	16	11.6–20.4
$\alpha_{u,m}$ ($aphi16$)	Constant associated with ψ_m at unstable condition	16	11.6–20.4
$afac$	Constant associated with temperature perturbation due to surface buoyancy flux	6.8	3.4–13.6
$bfac$	Constant associated with nonlocal transport of heat	6.8	3.4–13.6
S_m	Constant associated with nonlocal transport of momentum	15.9	7.95–31.8
$pfac$	Exponent associated with momentum diffusivity	2.0	1.0–4.0
b	Constant associated with velocity scale for heat flux calculation at inversion layer	5	2.5–10
c	Constant associated with heat flux calculation at inversion layer	0.15	0.075–0.3
Pr_h	Prandtl number at inversion layer	1	0.5–2
a_4	Constant associated with Prandtl number calculation	3.0	1.5–6.0
Ri_{land}	Critical bulk Richardson number over land	0.25	0.2–0.3
$Ri_{f,\text{ocean}}$	Factor associated with critical bulk Richardson number over ocean	1	0.5–2
λ_0	Asymptotic length scale for free atmosphere diffusion	30	15–60
k	Von Kármán constant	0.4	0.35–0.4

the MYNN schemes are investigated (Table 1, for additional details please refer to Yang et al., 2017), including those related to the dissipation and diffusion of TKE (B_1 and D_f), the turbulent Prandtl number (Pr), closure constants (C_2, C_5, and γ_1), and coefficients associated with the turbulent length scales (α_1–α_5 and β).

The YSU PBL scheme (Hong, 2010; Hong et al., 2006) is a first-order nonlocal *K*-profile scheme (Hong & Pan, 1996; Troen & Mahrt, 1986) that applies a countergradient term to represent the effects of large-scale eddies on the total flux. An explicit treatment for entrainment process at the PBL top has been included (Hong et al., 2006; Noh et al., 2003). Hong (2010) implemented a new vertical diffusion package in the YSU scheme for stable conditions, in which the PBL height is calculated based on the bulk Richardson number (Vickers & Mahrt, 2004). Here we focus on 15 YSU parameters (Table 2) such as those related to similarity functions (α_s, $\alpha_{u,h}$, and $\alpha_{u,m}$), constants associated with nonlocal transports ($bfac$, S_m), profile shape exponent for momentum diffusivity ($pfac$), constants related to the turbulent Prandtl number (Pr_h, a_4), and Von Kármán constant (k). In previous WRF versions, the YSU PBL scheme was tied with the MM5 (i.e., the fifth generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model, Grell et al., 1994) surface layer scheme and part of the calculations for similarity functions are connected between the two parts of the model. In the current WRF version, the similarity functions in the MM5 scheme have been updated (Jiménez et al., 2012) for both stable (Cheng & Brutsaert, 2005) and unstable conditions (Fairall et al., 1996). Thus, the two parts are not tied the same way as before. Note that the Von Kármán constant (k) is perturbed in the YSU experiments but fixed in the MYNN experiments. Our results suggest that this difference has only a small impact on the comparison between the two schemes. We can eliminate the effects of varying k via the surrogate model and resampling process as needed.

2.3. Sensitivity Analysis Approach

Here a total of 256 parameter sets for each PBL scheme are generated using the quasi-Monte Carlo sampling approach (Caflisch, 1998). Then the generated parameters sets are applied in WRF simulations using the configuration described in section 2.1. The generalized linear model (McCullagh & Nelder, 1989) is utilized to decompose the total variance of the WRF outputs (i.e., turbine-height wind speeds) into those contributed by different parameters.

The uncertainty in model results can be broken into two parts: parametric and structural. Parametric uncertainty is associated with the range of valid parameter values that are applied in the model. Structural uncertainty can be attributed to the structural differences among schemes if the simulation discrepancies cannot be reduced further by adjusting the parameter values given current knowledge of the parameter

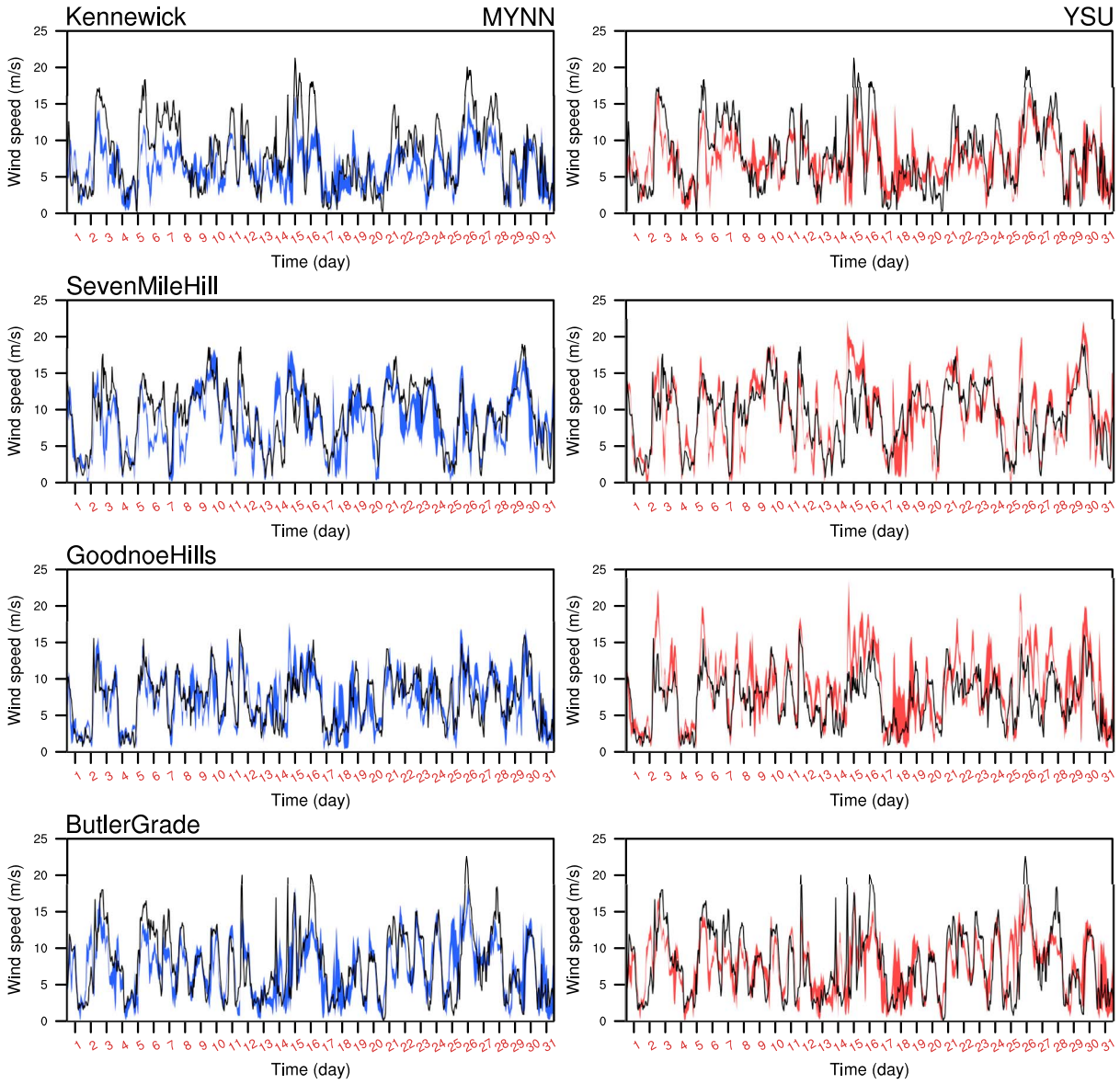

Figure 2. Time serials of wind speeds in observations (black lines) and simulations from the MYNN (blue belts in the left panel) and YSU (red belts in the right panel) ensembles at four selected sites (nos. 1–4 in Figure 1). The simulated wind speeds are interpolated to the corresponding height of measurement at each site. The colored belts denote the 5th to 95th percentiles in the 256 experiments. MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

uncertainties. Structural errors can be more complicated as the biases are possibly due to errors in other components or parameterizations (e.g., surface layer scheme) of the model. In this study, the term "structural error" is used to mean model error that cannot be explained by the uncertainties of the PBL parameters explored here. In practice, we need to examine the entire parameter spaces in all components of the model and then, the bias remaining could be attributed to model inadequacies, including structural errors in schemes other than PBL. However, because the WRF model is quite complex, here we focus on results from different PBL schemes as an example (assuming other physical components are perfect) to examine how to best characterize the structural sensitivity and structural error for a complex dynamical system.

3. Results

Our analysis starts with a comparison of the time series of the simulated and measured wind speeds. The simulated wind speeds are interpolated to the corresponding height of measurement at each site

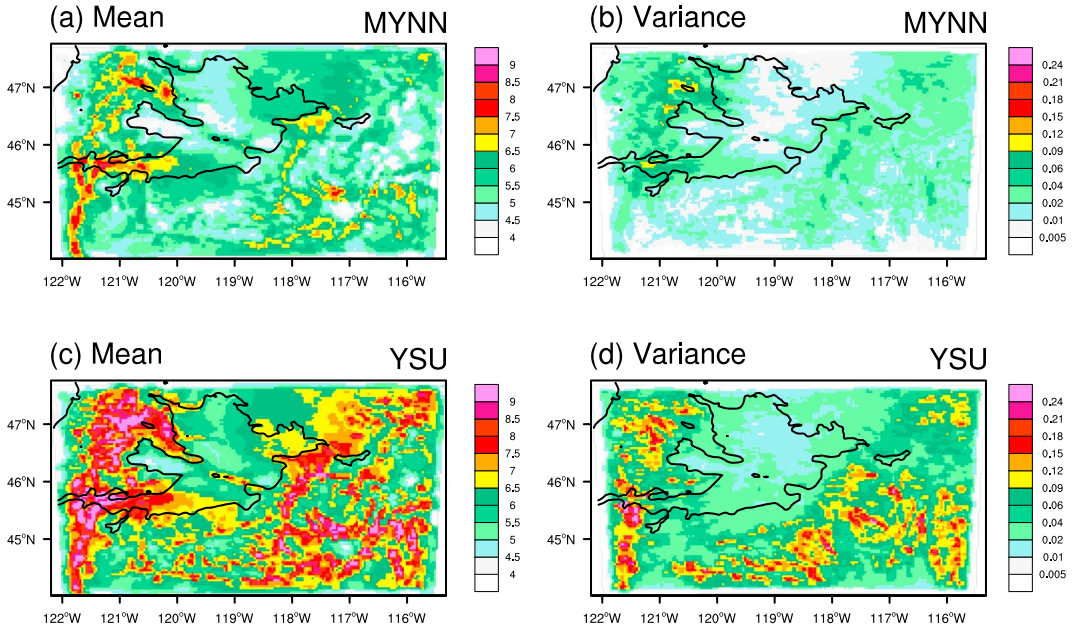

Figure 3. Spatial distributions of (a, c) ensemble mean (m/s) and (b, d) intermember variance (m^2/s^2) of 80-m wind speeds from the 256 experiments with perturbed parameters in (a, b) MYNN and (c, d) YSU during daytime (0600–1800 LT). Terrain height below 500 m is denoted by black lines, which approximately represents the Columbia Basin. MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

(Figure 1). During the simulation period, that is, May 2011, the observed low-level wind speeds show strong day-to-day variations at all sites (black lines in Figure 2; only the results at four sites are shown here). The wind speeds generally range from 0 to 20 m/s. All the experiments with the MYNN (blue belt in left panel, Figure 2) or YSU (red belt in right panel, Figure 2) scheme capture the variations in wind speed, but biases can be seen in some instances where observations fall outside the envelope of the ensemble members.

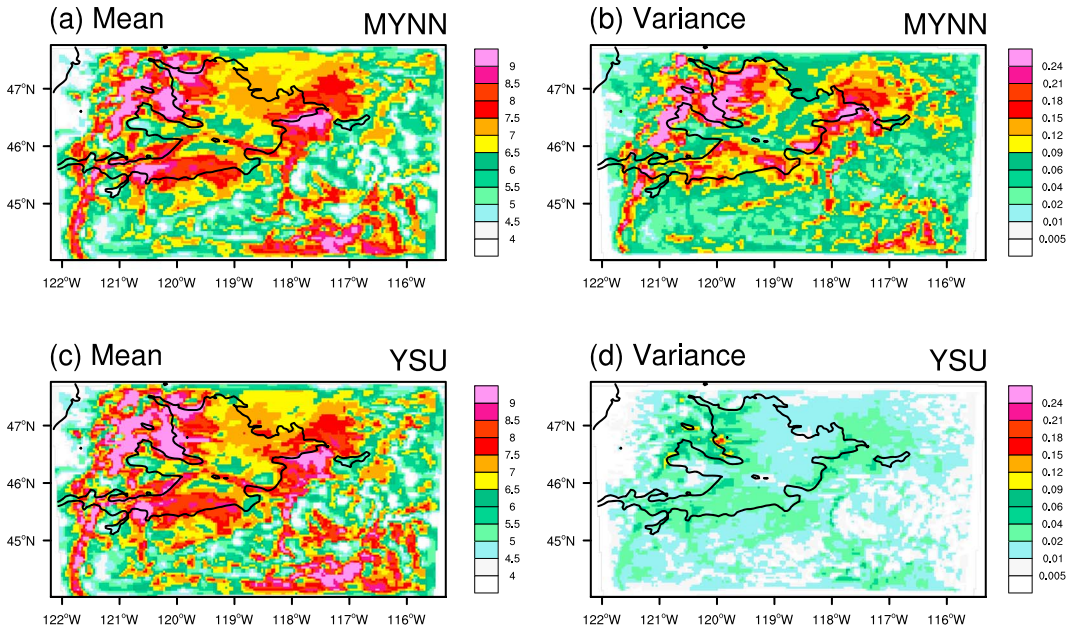

Figure 4. Same as Figure 3 but for nighttime condition (1800–0600 LT). MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

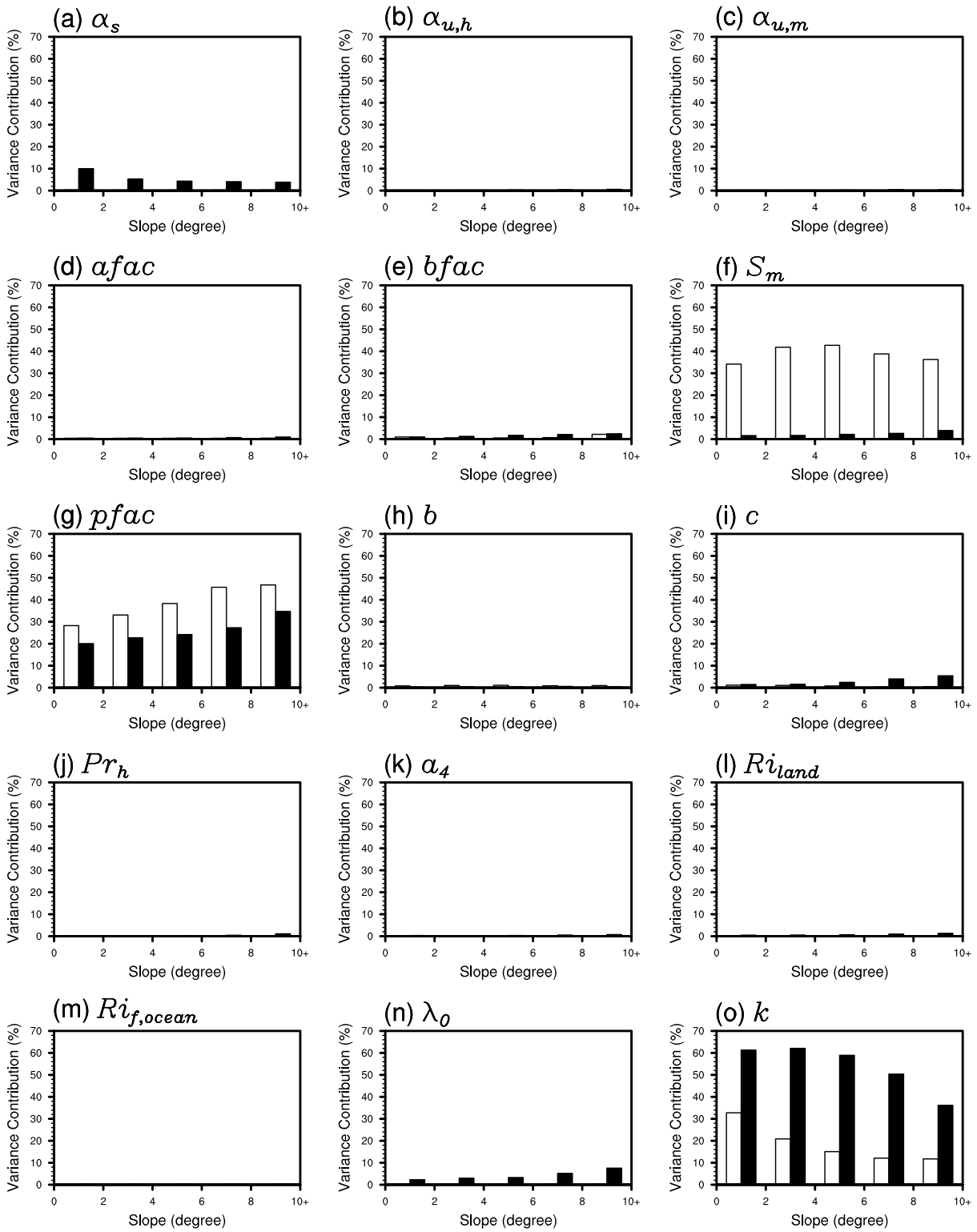

Figure 5. Relative contributions of different Yonsei University parameters to the variance of 80-m wind speeds as a function of terrain slope during daytime (blank bars) and nighttime (solid bars) conditions.

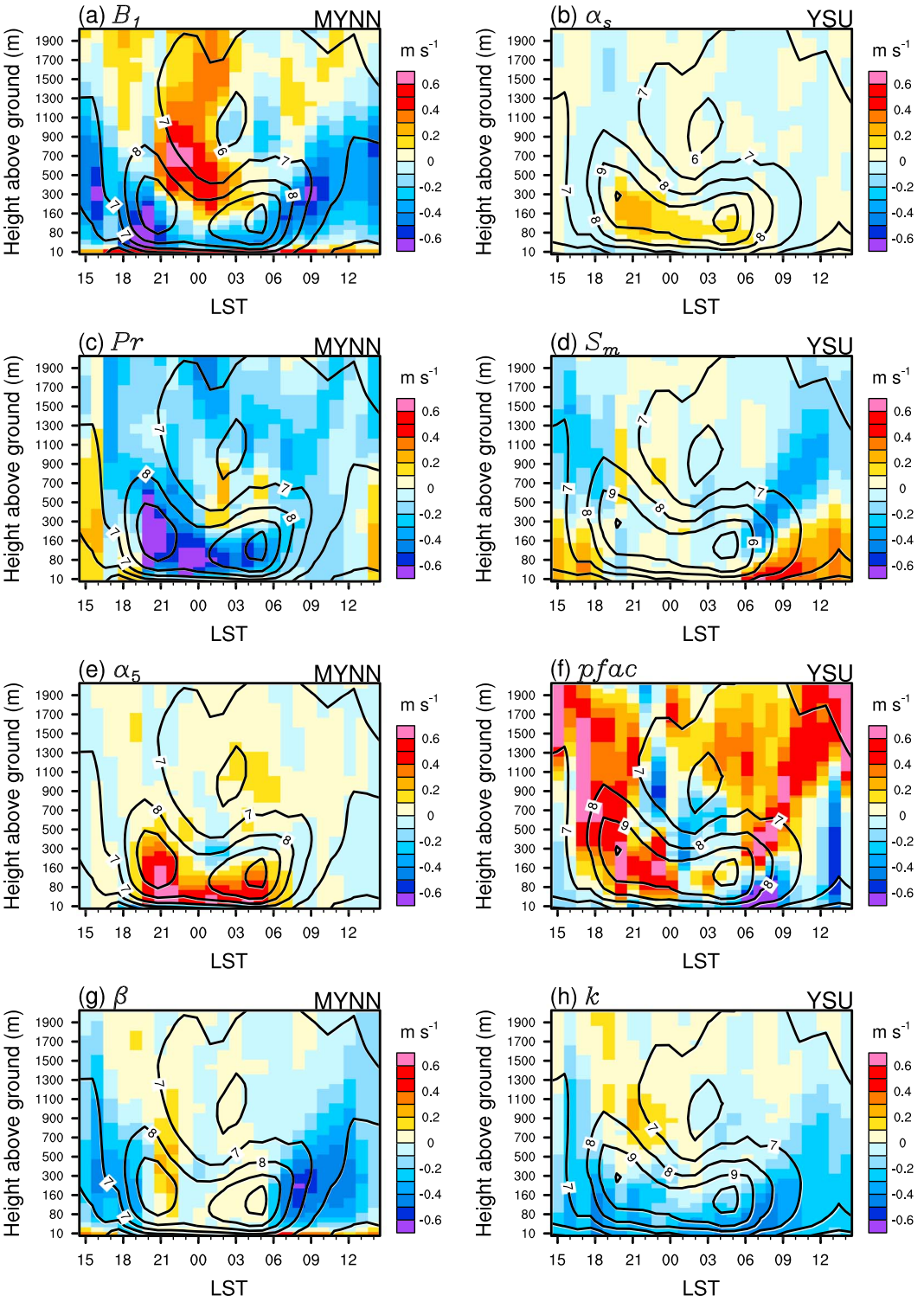

Figure 6. Diurnal vertical distributions of responses (average of simulations in the eighth group minus that in the first group) of wind speed (m/s) at the Butler Grade site (see Figure 1 for location) to the four influential parameters in (a, c, e, g) MYNN and (b, d, f, h) YSU. The ensemble mean wind speeds (m/s) for each scheme are indicated by contours and numbers in the white boxes. MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

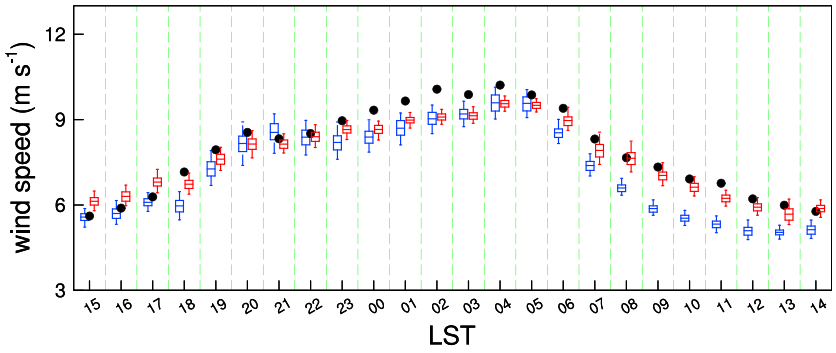

Figure 7. Monthly mean diurnal variations of 62.5-m wind speeds in observation (black dots) and simulations using MYNN (blue boxes) and YSU (red boxes) schemes at the Butler Grade site. At each hour, the mean, 25th to 75th percentiles, and 5th to 95th percentiles in the 256 experiments are represented as box-and-whisker plots. LST = local standard time.

Because of the nonlinear dependence of wind power on wind speed, very small changes or biases in wind speed can cause remarkable variations in the wind power, particularly for cases with wind speed ranging from the cut-in speed to rated speed (i.e., 3–13 m/s) of the wind turbine (e.g., Yang et al., 2017). In the following, we first examine the overall sensitivity of the simulated turbine-height wind speed in each PBL scheme. The responses of the vertical profile of wind speed are also examined to understand the variations of wind speeds at turbine height (normally 80–100 m above the surface). The structural sensitivity to the choice of PBL scheme and incompatible biases among different wind conditions due to structural error are then explored in section 3.2.

3.1. Parametric Sensitivity in the Two PBL Schemes

3.1.1. Spatial Feature of Parametric Sensitivity

The spatial distributions of the ensemble mean and intermember variance (i.e., the variance of the 256 members, which is caused by parameter perturbations) of the simulated 80-m wind speeds are different in the MYNN and YSU PBL ensembles (Figures 3 and 4). During daytime (Figure 3), the average 80-m wind speed (Figures 3a and 3c) within the Columbia Basin (approximately represented by the black contour in Figure 3) is generally less than 7 m/s, which is much smaller than that in the surrounding regions with elevated terrain. The YSU scheme (Figure 3c) produces larger ensemble-mean wind speed than the MYNN scheme (Figure 3a). This is at least partly due to the inclusion of nonlocal momentum mixing in the YSU scheme, which rapidly increases winds in the lower PBL during the morning transition by transporting large momentum from nocturnal low level jets to the surface (e.g., Hourdin, 2015; Xie & Fung, 2014). Correspondingly, compared to the MYNN parameters, the YSU parameters cause larger intermember variance of 80-m wind speeds during daytime (Figures 3b and 3d), but the large variability is mainly seen in areas with elevated terrain and larger wind speeds. The momentum flux in the YSU scheme consists of three terms (i.e., local [or eddy diffusivity], nonlocal, and entrainment fluxes), while the MYNN represents the momentum flux only with the eddy diffusivity term. Given the relatively large contribution of nonlocal transport to the mixing in the convective boundary layers (e.g., LeMone, 1973; Young, 1988), this explains why the YSU shows larger parametric sensitivity in daytime conditions.

Nighttime conditions (Figure 4) can be compared to daytime. The average 80-m wind speeds are larger during nighttime (Figures 4a and 4c) over most of the domain. The increase of wind speed from daytime to nighttime is more apparent in the MYNN ensemble than in the YSU ensemble. Thus, the nighttime wind speeds simulated by the two schemes are comparable with each other. In the YSU ensemble, the nighttime wind speeds exhibit a very weak response to the perturbed parameters (Figure 4d) despite the larger ensemble-mean wind speeds during nighttime than during daytime. In contrast, perturbing the MYNN parameters have a larger impact on the simulated wind speed during nighttime (Figure 4b), particularly over the elevated regions surrounding the Columbia Basin.

Following Yang et al. (2017), the contributions of each individual parameter to the variances of 80-m wind speeds are calculated at each grid point based on generalized linear model analysis. Then, we average the

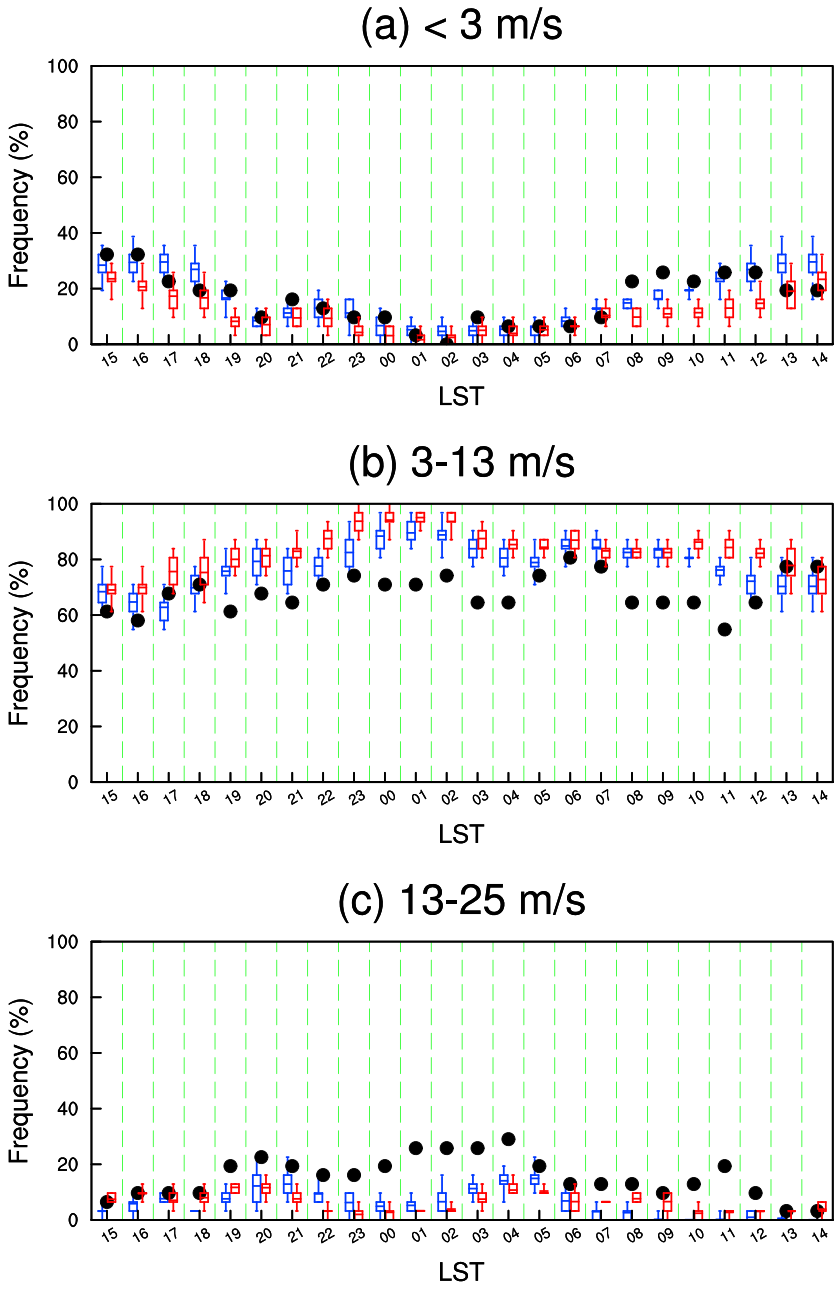

Figure 8. Diurnal variations of frequency (%) of 62.5-m wind speeds of (a) 0–3 m/s, (b) 3–13 m/s, and (c) 13–25 m/s in observation (black dots) and simulations using Mellor-Yamada-Nakanishi-Niino (blue boxes) and Yonsei University (red boxes) schemes at the Butler Grade site. The meaning of the box-and-whisker plot at each hour is the same as that in Figure 7. LST = local standard time.

results for different categories of grid points based on their terrain slope. Yang et al. (2017) and Berg et al. (2019) revealed that most of the variances of 80-m wind speeds in the MYNN ensemble can be explained by a few parameters. We briefly describe the parameter impacts in MYNN to provide a proper context for the following analyses. Detailed results can be referred to Yang et al. (2017). Among all the parameters, B_1 associated with the TKE dissipation rate is the most influential regardless of the terrain slope and time of day. The impacts of the parameter *Pr* are mainly seen at night. The perturbations of the turbulent length scale parameters considerably contribute to the variances in the 80-m wind speed. For example, the parameter α_1 can affect the simulated wind speed during both daytime and nighttime. By contrast, the

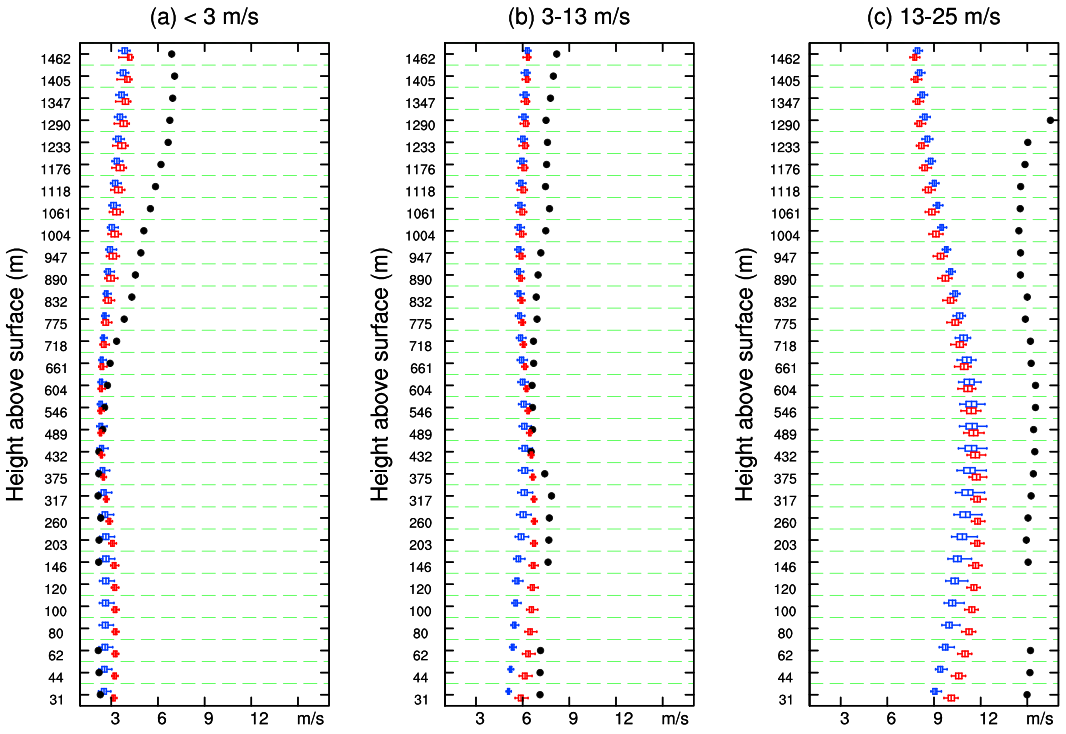

Figure 9. Vertical profiles of wind speeds during (a) weak, (b) moderate, and (c) strong wind conditions in observation (black dots; a combination of data from tower at heights of 31, 44, and 62 m and radar for 146–1,462 m) and simulations using Mellor-Yamada-Nakanishi-Niino (blue boxes) and Yonsei University (red boxes) schemes at the Butler Grade site during 0800–1200 LT. The meaning of the box-and-whisker plot at each height is the same as that in Figure 7. Note that the weak, moderate, and strong conditions are defined as those with the observed 62.5-m wind speeds of 0–3, 3–13, and 13–25 m/s, respectively.

impacts of α_5 and β are only seen during nighttime and daytime, respectively. The impacts of the other parameters in MYNN are much weaker compared to the above ones.

The relative contributions of each parameter in YSU are analyzed here (Figure 5). During daytime (blank bars in Figure 5), S_m (Figure 5f) and *pfac* (Figure 5g), both related to the momentum fluxes (Table 2), are the two parameters contributing the most to the 80-m wind speed variance, with the impact of the latter increasing with slope. The parameter k (i.e., Von Kármán constant; Figure 5o) can also affect the daytime wind speeds, especially over simple terrain with small slope. During nighttime (solid bars in Figure 5), most of the variance of 80-m wind speed can be explained by *pfac* and k. As terrain complexity increases, the relative impacts of the two parameters increase and decrease, respectively. The impacts of the other parameters on the simulated 80-m wind speed are nearly negligible.

3.1.2. Sensitivity of Wind Speed Vertical Profiles

The sensitivities of turbine-height wind speeds to parameters are related, at least in part, to changes in the momentum flux at different levels. Therefore, the response of wind speed is generally a function of height above surface. We show in Figure 6 the impacts of the most influential parameters (i.e., B_1, Pr, α_5, and β in MYNN based on Yang et al., 2017, and α_s, S_m, *pfac*, and k in YSU based on Figure 5) on the diurnal variations of wind speed vertical profiles at the Butler Grade site. Other locations, for example, Hanford, show similar parameter impacts with those found for the Butler Grade site (not shown). In the following, all the analyses are based on the results at Butler Grade unless otherwise noted. The physical interpretation about the responses of wind speed vertical profiles to the MYNN parameters are described by Yang et al. (2017). The 256 simulations in each ensemble are categorized into eight groups (each group with 32 members) according to the value of each parameter, with the first (eighth) group corresponding to the smallest (largest) values of the parameter. The difference of the averages between the eighth and first groups associated with each parameter are used to represent the parameter's main effect on wind speed. In the mean distributions (i.e., average of all the 256 experiments; contour lines in Figure 6), there exists a period with large wind speed

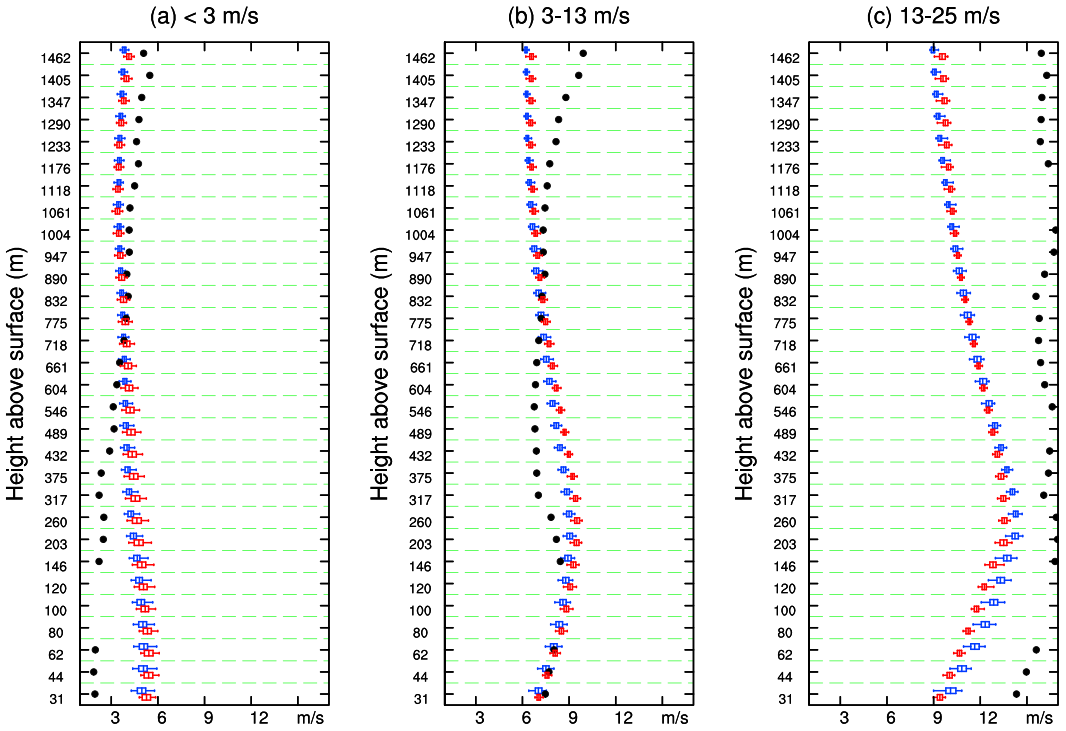

Figure 10. Same as Figure 9 but for the results during 1900–2300 LT.

below 500 m from 1800 LT to 0800 LT (nocturnal low-level jets), corresponding to the stronger nighttime 80-m wind speed than during daytime (Figures 3 and 4). The relative importance of parameters on 80-m wind speed is consistent with that in Figure 5, but with information about the sign of wind speed change included. In the MYNN ensemble, the three parameters, B_1, Pr, and α_5, can each cause a response of wind speed up to 0.6 m/s below 100 m during nighttime hours. On the contrary, the impacts of the YSU parameters on nighttime low-level wind speed are relatively weak. Increasing S_m increases (decreases) wind speeds in the lower (upper) PBL during daytime, by enhancing nonlocal momentum transport. Since the nonlocal transport term is effective only in convective boundary layers, the impact of S_m is relatively negligible during nighttime. With respect to *pfac*, increasing it lowers the height of the maximum diffusivity in the PBL (from 0.5 × PBL height for *pfac* = 1.0 to 0.2 × PBL height for *pfac* = 4.0) and therefore decreases (increases) the momentum diffusivity coefficient above (below) the maximum height. By changing the shape of the momentum diffusivity coefficient, larger *pfac* can lead to weaker and stronger wind at low and upper levels, respectively (Figure 6f). However, its effects cannot persist throughout the entire nighttime at low levels. During daytime, there is less sensitivity to the MYNN parameters than is found for the YSU parameters.

3.2. Structural Sensitivity to PBL Scheme

3.2.1. Diurnal Cycle of Turbine-Height Wind Speeds

Any specific physical component in the model, for example, the PBL process, can be parameterized in various ways, involving different types of physical concepts, theoretical assumptions, or simplifications. For example, one of the biggest structural difference between the YSU and MYNN PBL schemes is the inclusion of nonlocal transport terms; the YSU considers nonlocal transport of heat and momentum, while the MYNN does not. However, the existence of numerous uncertain parameters makes it difficult to attribute the simulation difference among schemes to the scheme structural discrepancy. With the optimized parameter values to fit the observations best, two different schemes could possibly perform similarly. In this section, the simulated wind speeds from two PPEs with the MYNN and YSU schemes are compared to isolate the effects of structural sensitivity from those due to parameter uncertainty. The relationship between the two ensembles can generally fall into three conditions: (1) the two ensembles that do not overlap with each other, (2) the two ensembles that have an intersection, and (3) one ensemble that is a subset of the other. The first

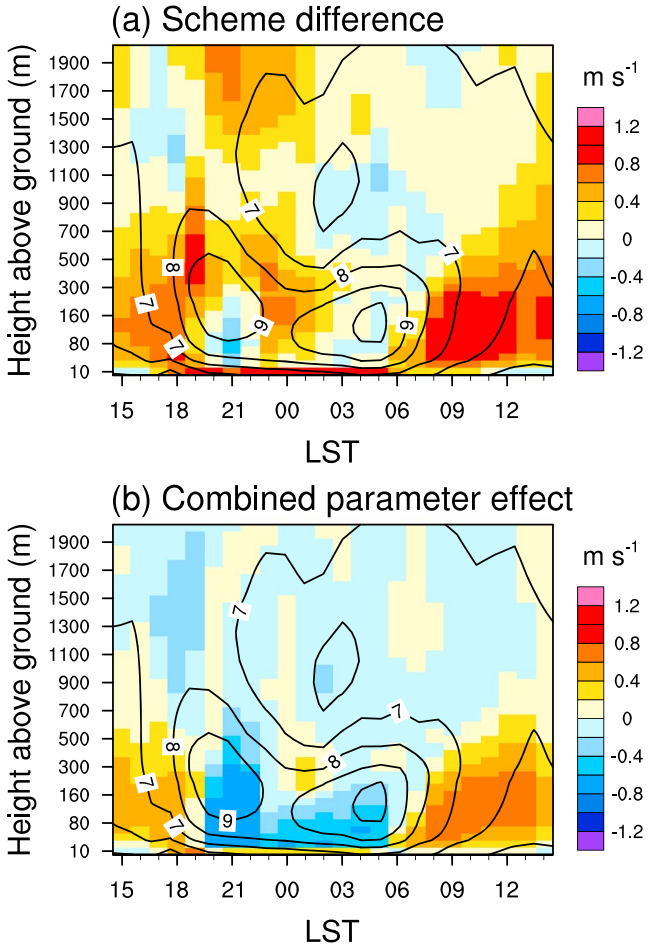

Figure 11. Diurnal vertical distributions of (a) wind speed difference between the two ensemble means using YSU and MYNN (YSU minus MYNN) and (b) combined effect due to parameters S_m in YSU and α_5 and β in MYNN (i.e., values shown in Figure 6d minus that in Figure 6e and in Figure 6g) at the Butler Grade site. The averages of the two ensemble means of wind speeds (m/s) are indicated by contours and numbers in the white boxes. LST = local standard time; MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

condition means there are no chances to tune the two schemes to perform the same. The latter two conditions indicate some members in one ensemble cannot be reproduced by the other. When observations are considered, bias evident in the best member of the ensemble may also imply the existence of scheme structural error (Williamson et al., 2015). Note that structural error in one scheme could become a parametric error in the other scheme, particularly for the third condition, for example, some fixed component in one ensemble is represented by tunable parameters in the other. Such information can have important implications for further parameterization development.

The observed diurnal variation of monthly mean 62.5-m wind speeds at Butler Grade (black dots in Figure 7) shows that the wind speed peaks at around 0400 LT and then rapidly decreases after sunrise due to the intensified land-air coupling and boundary layer mixing. The two schemes (blue and red boxes in Figure 7) are able to produce the overall diurnal features of wind speeds. Differences are evident between the two ensembles during daytime, particularly for the period from 0800 LT to 1200 LT when the two distributions do not overlap with each other. Obviously, the YSU scheme has a greater potential to match the observations. A large bias exists in the MYNN results even with well-constrained parameter values, indicating structural error in the scheme regarding the monthly mean 62.5-m wind speeds during morning to noon hours. The structural error of MYNN is at least in part related to the insufficient momentum transport due to the absence of the nonlocal transport term (e.g., Hourdin, 2015; Xie & Fung, 2014). During nighttime, both schemes have the potential to reproduce the observed 62.5-m wind speed via parameter calibration, except for the period between 0000 and 0002 LT when all the simulations underestimate the wind speed. After sunrise, the MYNN ensemble generally produces stronger decelerations of wind speeds than observation, while the YSU ensemble performs better in this regard.

The above analysis indicates that parameter-tuned YSU scheme could perform better than the MYNN scheme in regard to the monthly mean diurnal cycle of wind speed. However, opposite conclusions can be drawn when focusing on the diurnal features for cases with different classes of wind speeds (Figure 8). In the observations, the frequency of occurrences for wind speeds less than 3 m/s is approximately 30% during daytime, which is much larger than during nighttime (black dots in Figure 8a). The MYNN scheme generally simulates larger frequencies of the weak wind cases than the YSU scheme. The spreads in both the two ensembles overlap the observed value for most times of a day, except for the period between 0800 and 1200 LT when MYNN shows a potential to produce smaller biases than YSU. Similar differences between the two schemes can be found for conditions with wind speeds ranging between 3–13 m/s (Figure 8b) and 13–25 m/s (Figure 8c), in which the best member from the MYNN ensemble agrees more with observations than that from the YSU ensemble during nighttime. All the members in the two ensembles underestimate the frequency of occurrence of strong wind cases (i.e., wind speed ranging between 13 and 25 m/s, Figure 8c) during 2300–0400 and 1000–1200 LT.

3.2.2. Vertical Profiles of Wind Speeds

We have shown that the YSU scheme generally simulates stronger turbine-height wind speeds than the MYNN scheme during daytime (Figures 7 and 8). The simulated vertical profiles from the two ensembles during 0800–1200 LT are further examined (Figure 9) to help understand the differences. Hourly data during 0800–1200 LT are classified into three categories based on the observed wind speeds at 62.5-m height. The low-level winds (from near the surface to 100 m) simulated by the MYNN scheme are systematically weaker than those simulated by the YSU scheme for all ensemble members across all conditions, especially for cases

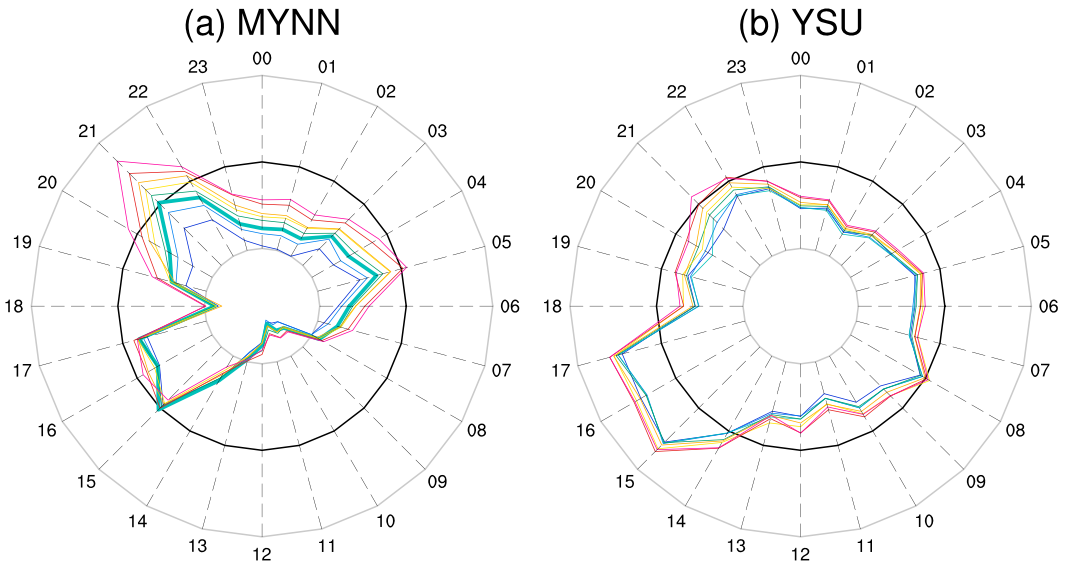

Figure 12. Relationship of model biases (normalized by observational values) of 62.5-m wind speeds across different time of day in the ensemble experiments using (a) MYNN and (b) YSU schemes at the Butler Grade site. For each scheme, the 256 experiments are grouped into eight bins (with different colors) based on the wind speeds at 2100 LT. The black and gray lines denote the relative biases of 0 and ±15%, respectively. The colored curves from blue to red represent the groups with relative bias approximately from −7% to +10% at 2100LT. MYNN = Mellor-Yamada-Nakanishi-Niino; YSU = Yonsei University.

with moderate and strong winds (Figures 9b and 9c). In strong wind cases, both schemes tend to underestimate the wind speeds in the entire PBL even with well-tuned parameters, indicating the existence of structural errors or errors in other model components.

For nighttime conditions (Figure 10), the differences between the two schemes are smaller than during daytime. The MYNN scheme has a larger potential to match observations in both weak and strong wind conditions. However, the best members for one wind speed category may be the worst ones for the other wind speed category (discussed in section 3.2.3).

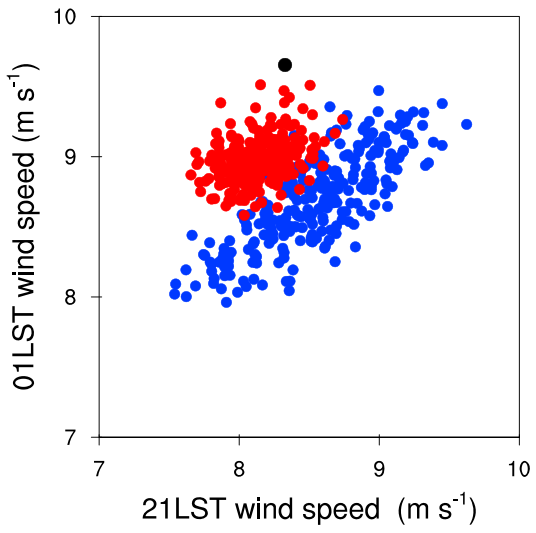

Figure 13. Scatter plots of 62.5-m wind speeds at 2100 LT (*x* axis) and 0100 LT (*y* axis) in observation (black dot) and simulations using Mellor-Yamada-Nakanishi-Niino (blue dots) and Yonsei University (red dots) schemes at the Butler Grade site. LST = local standard time.

The systematically weaker daytime winds in MYNN compared to YSU (Figure 9) are likely due, at least in part, to the absence of nonlocal transport in MYNN. Figure 11a indicates that the mean differences between the two ensembles are most evident at heights ranging between 50 and 300 m during daytime. Our sensitivity analyses show that nonlocal transport in YSU leads to an increase in wind speeds in the lower PBL and decrease in wind speeds in the upper PBL (Figure 6d). We can see that in the upper PBL (above 500 m) and near the surface, the differences between the two ensembles (Figure 11a) are not consistent with the impacts associated with nonlocal transport (Figure 6d). In addition to the absence of nonlocal transport, the parameterization for the momentum exchange between near the surface and aloft in MYNN is also important, which is directly affected by the surface layer length scale parameter β during daytime (Figure 6g). We combine the impacts due to nonlocal transport in YSU (Figure 6d) and the surface layer length scale in MYNN (Figure 6g), which produces a pattern (Figure 11b) resembling the difference between the two ensembles during daytime (Figure 11a). The results indicate that compared to MYNN, the nonlocal term in YSU transports more momentum from the upper PBL to the lower PBL. Meanwhile, stronger mixing between near the surface and aloft in MYNN transports more momentum downward to near the surface, amplifying the differences between the two ensembles for heights ranging

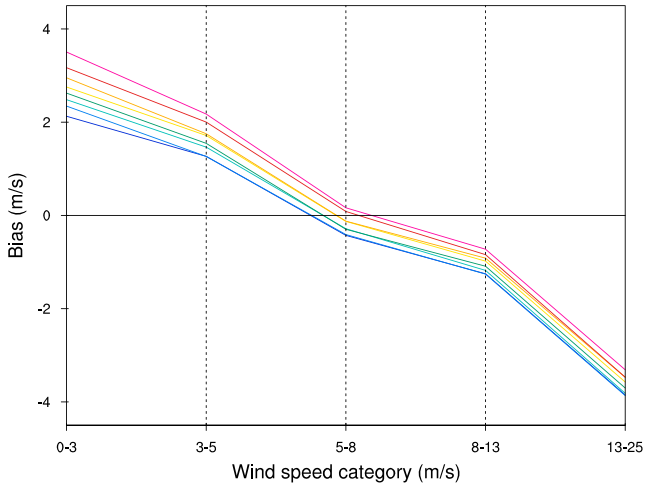

Figure 14. Model biases of 62.5-m wind speeds across various wind strengths of 0–3, 3–5, 5–8, 8–13, and 13–25 m/s in the Mellor-Yamada-Nakanishi-Niino ensemble at the Butler Grade site during nighttime. Hourly data are grouped into different categories based on the observed wind speeds. The 256 experiments are grouped into eight bins (with different colors) based on the biases for wind speed of 0–3 m/s.

between 50 and 300 m but decreasing the differences below and above. We also include the impact due to α_5 (Figure 6e) in Figure 11b, which partly explains the difference between the two ensembles at night. Hence, simultaneously tuning the parameters related to nonlocal transport in YSU and the surface layer length scale in MYNN can reduce the wind speed differences produced by the two schemes. However, to reduce the differences as much as possible, we have to use parameter values out of their uncertainty ranges or modify the formulations used in the parameterizations. It should be noted that in YSU the impacts of nonlocal transport can be partly compensated by the changes in the height of maximum diffusivity (Figure 6f). Similarly, the mixing strength in MYNN can also be affected by other parameters such as that related to the TKE dissipation rate (Figure 6a). Nevertheless, our results suggest that the daytime differences of using the two schemes are mainly caused by two factors: YSU produces a relatively stronger mixing between the lower part and upper part of PBL, and MYNN simulates a relatively stronger mixing between near the surface and aloft.

3.2.3. Incompatible Biases Due to Structural Error

Our results show that using parameter calibration has a low probability of eliminating model biases related to structural problem, but errors associated with parameter uncertainties can be reduced to a large degree. In this section, we further explore whether the improvements by tuning for a specific site, time of day, or wind speed category would degrade model performances in others.

We can take the simulated diurnal cycle of wind speed at Butler Grade as an example. The spreads in both ensembles cover the observational wind speeds during 1900–2300 LT (Figure 7). For each ensemble, we split the 256 experiments into eight groups (each with 32 members) based on the simulated magnitudes of wind speeds at 2100 LT, and then we compare the simulated diurnal cycles averaged in each group against the observation (Figure 12). When using the MYNN scheme (Figure 12a), the third group (thick green line) shows the best agreement with observation at 2100 LT but large biases during other periods. To reduce the biases at other times, particularly at 2300–0700 LT, we need to increase the wind speeds (from green line to pink line) by perturbing certain parameters. However, such tuning process inevitably degrades the simulated wind speed at 2100 LT. The incompatible biases can also be seen in the YSU ensemble, but it is less evident when the experiments are grouped based on the wind speed at 2100 LT.

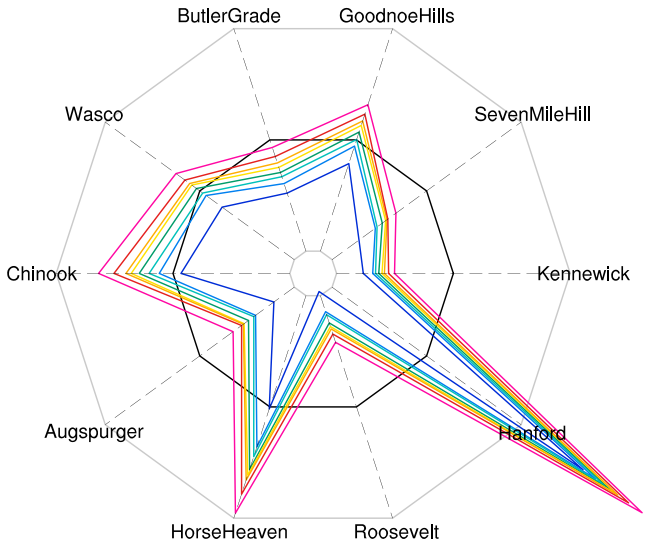

Figure 15. Relationship of nighttime model biases in wind speeds (normalized by observational values) across different sites (see Figure 1 for locations) in the ensemble experiments using Mellor-Yamada-Nakanishi-Niino scheme. The 256 experiments are grouped into eight bins (with different colors) based on the biases at the Butler Grade site. The black and gray lines denote the relative biases of 0 and ±25%, respectively.

The biases of wind speeds at different times are incompatible likely because they are linked with fixed yet unrealistic relationships that are insensitive to the investigated parameters. The results based on averaged values within different individual groups (Figure 12) explain a large portion of the connection among wind responses at different nighttime hours, which can be further attributed to certain parameters (Figure 6). However, some information could still be missing due to the variability within each group. The correlation between wind speeds at 2100 LT and 0100 LT based on all the 256 experiments (Figure 13) confirms the incompatible biases between the two times in the MYNN ensemble (blue dots vs. black dot). By contrast, the YSU scheme (red dots) has a larger chance to produce the observed wind speeds at both times, as shown by the proximity of some of the YSU ensemble members to the observations. For the MYNN ensemble, the two biases could be simultaneously reduced if the wind speeds at the two times are interrelated in different ways (e.g., different intercept, slope, or nonlinearity) or less correlated (i.e., more coverage in the two-dimension space), which can be addressed by additional

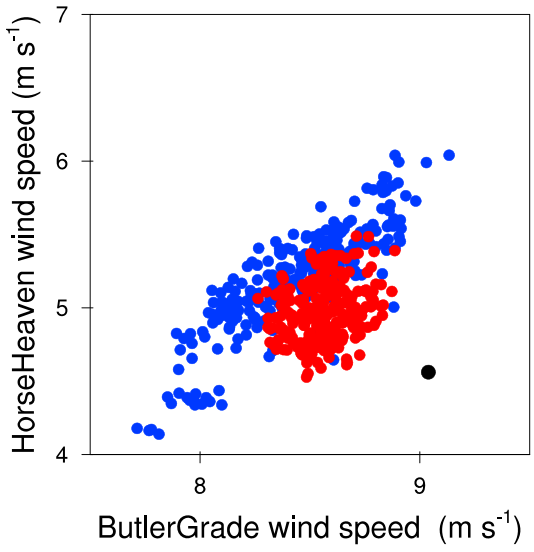

Figure 16. Scatter plots of nighttime wind speeds at the sites of Butler Grade (*x* axis) and Horse Heaven (*y* axis) in observation (black dot) and simulations using Mellor-Yamada-Nakanishi-Niino (blue dots) and Yonsei University (red dots) schemes.

developments in the parameterizations, but it could still possibly be a parametric issue beyond our current knowledge.

The incompatible biases also exist among conditions with different wind speeds (Figure 14). We only show the MYNN results during nighttime when the parameters can exert strong impacts. Nighttime hourly data are classified into five categories based on the observed wind speeds ranging from 0 to 25 m/s. Then, the 256 simulations are divided into eight groups based on the biases for wind speed of 0–3 m/s. The MYNN scheme always overestimates the wind speeds between 0 and 5 m/s but underestimates them between 8 and 25 m/s. The wind speeds are sensitive to the MYNN parameters in cases with both weak and strong winds. However, their biases are anticorrelated when perturbing the parameters. To mitigate these incompatible biases, we can assign the parameter value as a function of wind strength, but whether this treatment is physically justifiable needs further examination.

The biases of wind speeds (averaged during nighttime) in the MYNN ensemble are also explored at different sites (Figure 15). The results indicate a systematic response of wind speed to parameters across different locations. Biases (absolute value) in different sites can be positively or negatively correlated with one another. There is an evident contrast between the results at Butler Grade and Horse Heaven. Detailed relationship between the two sites can be seen based on the scatter plot of all the 256 experiments (Figure 16). Both the wind speeds at these two sites simulated by MYNN (blue dots in Figure 16) can be tuned to match observation individually, but the best parameter sets are mutually exclusive, for example, the optimal parameter set for wind speed at Butler Grade is among the worst ones for wind speed at Horse Heaven.

4. Summary and Discussion

We investigated the sensitivities of turbine-height wind simulations to the structural difference and uncertain parameters in the PBL parameterizations by comparing experiments from two PPEs using the MYNN and YSU PBL schemes for a domain in the Pacific Northwest of the United States. We explored the high-dimensional parameter space by applying an efficient sampling algorithm. The parametric sensitivities of the simulated wind speeds in the two schemes were compared, including using the surrogate-based variance decomposition. The structural sensitivity to the choice of PBL scheme and incompatible biases due to structural error were then explored.

During daytime, the ensemble mean and inter-member variance (i.e., sensitivity to parameters) of the simulated turbine-height wind speeds across the entire domain are larger in the YSU ensemble than in the MYNN ensemble. During nighttime, by contrast, the MYNN parameters have larger impact on the wind speed than the YSU parameters, although the ensemble-mean wind speeds of the two ensembles are comparable. A majority of the intermember variances can be explained by a few parameters in each scheme, such as those associated with the TKE dissipation, Prandtl number, and turbulent length scales in the MYNN scheme, and the momentum flux-related coefficients and Von Kármán constant in the YSU scheme.

The simulated winds from the two ensembles were evaluated against the measurements at the Butler Grade site to explore the structural errors and structural differences in the schemes. In general, the two schemes are able to reproduce the observed diurnal features of the monthly mean turbine-height wind speed. During daytime, differences evident in the turbine-height wind speeds can be found between the two ensembles. The YSU scheme has a potential to match observation during daytime, while a large bias still exists in the MYNN results even with the well-tuned parameter values, suggesting the structural error in the MYNN scheme. Further analyses indicate that the discrepancies between the two schemes are evident across conditions with different magnitudes of wind speed during daytime; compared to the YSU scheme, the MYNN scheme performs better in weak wind cases but worse in strong wind cases. Both schemes underestimate the occurrence frequency of strong winds during both daytime and nighttime. Compared to YSU, MYNN is able to better match observations in both weak and strong wind conditions at night.

Model biases related to parameter uncertainties can be reduced to large degree by tuning, but the improvements in one particular target may compromise model performances in others. The relationships between different targets are not accurately represented or missing in the parameterization. For example, the turbine-height wind speeds are markedly underestimated by the MYNN scheme during late night. These biases can be largely reduced by perturbing certain parameters, but such tuning process inevitably degrades the simulated wind speed at 2100 LT in our example, suggesting the incompatible biases of wind speed among different times due to the scheme structural problem. The incompatible biases can also be detected when focusing on the wind speeds across different wind strengths or at different locations.

Comparing the experiments from different PPEs rather than single default cases can help to identify the distinctive structural characteristics of the simulations in a given scheme from others. For single-objective optimizations, the best members can be used to determine whether there exists a structural error and how it can be characterized, which can provide guidance for further parameterization development focusing on biases that cannot be explained by parameter uncertainty. For multiobjective optimizations, the incompatible biases among different targets and using different metrics will also expose the limitation of parameter tuning and necessity of making structural improvements to the parameterizations. In some cases, a structural problem in one scheme disappears or becomes a parametric issue in other schemes. This means we can better understand the structural errors in the schemes by referring to the response features, including those related to parameter perturbations, in other schemes. In our example, the different parameterizations both for local turbulent flux and for nonlocal transport might be important for the different wind speeds simulated by the two schemes, but additional detailed and comprehensive exploration of the physical reason responsible for those structural sensitivities is necessary. It also deserves future investigation to determine whether the conclusions revealed here can still hold when using a different model configuration (i.e., different options for other physics) or for a different application.

Acknowledgments

This work has been supported by the U. S. Department of Energy's (DOE) Wind Power Program. The work of B. Y. at Nanjing University is supported by the National Natural Science Foundation of China (41675101). The Pacific Northwest National Laboratory (PNNL) is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC07-76RL01830. The PNNL Institutional Computing (PIC) provided computational resources. Meteorological data were proved by the Bonneville Power Administration (PBA towers) and MSA (Hanford). Radar wind profiler data was provided by the DOE Atmospheric Radiation Measurement (ARM) Climate Research Facility. The model outputs and observational data used in this study are publicly available in Github repository (https://github.com/brady1234/UQ_2PBL).

References

Andren, A., & Moeng, C. H. (1993). Single-point closures in a neutrally stratified boundary-layer. *Journal of the Atmospheric Sciences*, *50*(20), 3366–3379. https://doi.org/10.1175/1520-0469(1993)050<3366:Spcian>2.0.Co;2

Banta, R. M., Pichugina, Y. L., Brewer, W. A., James, E. P., Olson, J. B., Benjamin, S. G., et al. (2018). Evaluating and improving NWP forecast models for the future: How the needs of offshore wind energy can point the way. *Bulletin of the American Meteorological Society*, *99*(6), 1155–1176. https://doi.org/10.1175/bams-d-16-0310.1

Banta, R. M., Pichugina, Y. L., Kelley, N. D., Hardesty, R. M., & Brewer, W. A. (2013). WIND ENERGY METEOROLOGY: Insight into wind properties in the turbine-rotor layer of the atmosphere from high-resolution Doppler lidar. *Bulletin of the American Meteorological Society*, *94*(6), 883–902. https://doi.org/10.1175/Bams-D-11-00057.1

Berg, L. K., Liu, Y., Yang, B., Qian, Y., Olson, J., Ma, P.-L., & Hou, Z. (2019). Sensitivity of turbine-height wind speeds to parameters in the planetary boundary-layer parametrization used in the Weather Research and Forecasting model: Extension to wintertime conditions. *Boundary-Layer Meteorology*, *170*(3), 507–518. https://doi.org/10.1007/s10546-018-0406-y

Berg, L. K., Pekour, M., & Nelson, D. (2012). *Description of the Columbia Basin Wind Energy Study (CBWES), Technical Report, PNNL-22036* (p. 14). Richland, Washington: Pacific Northwest National Laboratory.

Boyle, J. S., Klein, S. A., Lucas, D. D., Ma, H. Y., Tannahill, J., & Xie, S. (2015). The parametric sensitivity of CAM5's MJO. *Journal of Geophysical Research: Atmospheres*, *120*, 1424–1444. https://doi.org/10.1002/2014jd022507

Caflisch, R. E. (1998). Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, *7*, 1–49. https://doi.org/10.1017/S0962492900002804

Carvalho, D., Rocha, A., Gomez-Gesteira, M., & Santos, C. (2012). A sensitivity study of the WRF model in wind simulation for an area of high wind energy. *Environmental Modelling and Software*, *33*, 23–34. https://doi.org/10.1016/j.envsoft.2012.01.019

Cheng, Y. G., & Brutsaert, W. (2005). Flux-profile relationships for wind speed and temperature in the stable atmospheric boundary layer. *Boundary-Layer Meteorology*, *114*(3), 519–538. https://doi.org/10.1007/s10546-004-1425-4

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., et al. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*, W00B02. https://doi.org/10.1029/2007WR006735

Dai, H., Ye, M., Walker, A. P., & Chen, X. (2017). A new process sensitivity index to identify important system processes under process model and parametric uncertainty. *Water Resources Research*, *53*, 3476–3490. https://doi.org/10.1002/2016WR019715

Draxl, C., Hahmann, A. N., Pena, A., & Giebel, G. (2014). Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy*, *17*(1), 39–55. https://doi.org/10.1002/we.1555

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean global atmosphere coupled ocean atmosphere response experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Giorgi, F., & Mearns, L. O. (2002). Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method. *Journal of Climate*, *15*(10), 1141–1158. https://doi.org/10.1175/1520-0442(2002)015<1141:Coaura>2.0.Co;2

Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., & Persson, P. O. G. (2007). SHEBA flux-profile relationships in the stable atmospheric boundary layer. *Boundary-Layer Meteorology*, *124*(3), 315–333. https://doi.org/10.1007/s10546-007-9177-6

Grell, G. A., Dudhia, J., & Stauffer, D. R. (1994). A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5), NCAR Technical Note (NCAR/TN-398+STR, 117 pp.).

Guo, Z., Wang, M. H., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., et al. (2014). A sensitivity analysis of cloud properties to CLUBB parameters in the single-column Community Atmosphere Model (SCAM5). *Journal of Advances in Modeling Earth Systems*, *6*, 829–858. https://doi.org/10.1002/2014ms000315

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, *48*, W08301. https://doi.org/10.1029/2011WR011044

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. https://doi.org/10.1175/2009bams2607.1

Hong, S. Y. (2010). A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. *Quarterly Journal of the Royal Meteorological Society*, *136*(651), 1481–1496. https://doi.org/10.1002/qj.665

Hong, S. Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https://doi.org/10.1175/Mwr3199.1

Hong, S. Y., & Pan, H. L. (1996). Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Monthly Weather Review*, *124*(10), 2322–2339. https://doi.org/10.1175/1520-0493(1996)124<2322:Nblvdi>2.0.Co;2

Hou, Z. S., Huang, M. Y., Leung, L. R., Lin, G., & Ricciuto, D. M. (2012). Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model. *Journal of Geophysical Research*, *117*, D15108. https://doi.org/10.1029/2012JD017521

Hourdin, F., & Coauthors (2015). Parameterization of convective transport in the boundary layer and its impact on the representation of the diurnal cycle of wind and dust emissions. *Atmospheric Chemistry and Physics*, *15*(12), 6775–6788. https://doi.org/10.5194/acp-15-6775-2015

Huang, D. Q., Zhu, J., Zhang, Y. C., & Huang, A. N. (2013). Uncertainties on the simulated summer precipitation over eastern China from the CMIP5 models. *Journal of Geophysical Research: Atmospheres*, *118*, 9035–9047. https://doi.org/10.1002/jgrd.50695

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., & García-Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, *140*(3), 898–918. https://doi.org/10.1175/Mwr-D-11-00056.1

Kim, J. W., & Mahrt, L. (1992). Simple formulation of turbulent mixing in the stable free atmosphere and nocturnal boundary-layer. *Tellus A*, *44*(5), 381–394.

Laboratory, N. R. E. (2008). *20% Wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply* (p. 228). Washington, DC: U.S. Department of Energy.

LeMone, M. A. (1973). The structure and dynamics of horizontal roll vortices in the planetary boundary layer. *Journal of the Atmospheric Sciences*, *30*(6), 1077–1091. https://doi.org/10.1175/1520-0469(1973)030<1077:TSADOH>2.0.CO;2

Leung, L. R., Ghan, S. J., Zhao, Z. C., Luo, Y., Wang, W. C., & Wei, H. L. (1999). Intercomparison of regional climate simulations of the 1991 summer monsoon in eastern Asia. *Journal of Geophysical Research-Atmospheres*, *104*(D6), 6425–6454. https://doi.org/10.1029/1998JD200016

Lu, X., McElroy, M. B., & Kiviluoma, J. (2009). Global potential for wind-generated electricity. *Proceedings of the National Academy of Sciences*, *106*(27), 10,933–10,938. https://doi.org/10.1073/pnas.0904101106

Mahoney, W. P., Parks, K., Wiener, G., Liu, Y. B., Myers, W. L., Sun, J. Z., et al. (2012). A wind power forecasting system to optimize grid integration. *IEEE Transactions on Sustainable Energy*, *3*(4), 670–682. https://doi.org/10.1109/Tste.2012.2201758

Marjanovic, N., Wharton, S., & Chow, F. K. (2014). Investigation of model parameters for high-resolution wind energy forecasting: Case studies over simple and complex terrain. *Journal of Wind Engineering and Industrial Aerodynamics*, *134*, 10–24. https://doi.org/10.1016/j.jweia.2014.08.007

Marquis, M., Wilczak, J., Ahlstrom, M., Sharp, J., Stern, A., Smith, J. C., & Calvert, S. (2011). Forecasting the wind to reach significant penetration levels of wind energy. *Bulletin of the American Meteorological Society*, *92*(9), 1159–1171. https://doi.org/10.1175/2011bams3033.1

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (532 pp.) London, UK: Chapman and Hall. https://doi.org/10.1007/978-1-4899-3242-6

McNeall, D., Williams, J., Booth, B., Betts, R., Challenor, P., Wiltshire, A., & Sexton, D. (2016). The impact of structural error on parameter constraint in a climate model. *Earth System Dynamics*, *7*(4), 917–935. https://doi.org/10.5194/esd-7-917-2016

Mellor, G. L., & Yamada, T. (1974). A hierarchy of turbulence closure models for planetary boundary layers. *Journal of the Atmospheric Sciences*, *31*(7), 1791–1806. https://doi.org/10.1175/1520-0469(1974)031<1791:ahotcm>2.0.co;2

Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure-model for geophysical fluid problems. *Reviews of Geophysics*, *20*(4), 851–875. https://doi.org/10.1029/RG020i004p00851

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360.

Musgrove, P. (2010). *Wind power* (p. 338). Cambridge, UK: Cambridge University Press.

Nakanishi, M. (2001). Improvement of the Mellor-Yamada turbulence closure model based on large-eddy simulation data. *Boundary-Layer Meteorology*, *99*(3), 349–378. https://doi.org/10.1023/A:1018915827400

Nakanishi, M., & Niino, H. (2004). An improved Mellor-Yamada level-3 model with condensation physics: Its design and verification. *Boundary-Layer Meteorology*, *112*(1), 1–31. https://doi.org/10.1023/B:Boun.0000020164.04146.98

Nakanishi, M., & Niino, H. (2006). An improved Mellor-Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, *119*(2), 397–407. https://doi.org/10.1007/s10546-005-9030-8

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan*, *87*(5), 895–912. https://doi.org/10.2151/jmsj.87.895

Nielsen-Gammon, J. W., Hu, X. M., Zhang, F. Q., & Pleim, J. E. (2010). Evaluation of planetary boundary layer scheme sensitivities for the purpose of parameter estimation. *Monthly Weather Review*, *138*(9), 3400–3417. https://doi.org/10.1175/2010mwr3292.1

Noh, Y., Cheon, W. G., Hong, S. Y., & Raasch, S. (2003). Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Boundary-Layer Meteorology*, *107*(2), 401–427. https://doi.org/10.1023/a:1022146015946

Pichugina, Y. L., Banta, R. M., Olson, J. B., Carley, J. R., Marquis, M. C., Brewer, W. A., et al. (2017). Assessment of NWP forecast models in simulating offshore winds through the lower boundary layer by measurements from a ship-based scanning Doppler lidar. *Monthly Weather Review*, *145*(10), 4277–4301. https://doi.org/10.1175/Mwr-D-16-0442.1

Pimentel, D., Hertz, M., Glickstein, M., Zimmerman, M., Allen, R., Becker, K., et al. (2002). Renewable energy: Current and potential issues. *Bioscience*, *52*(12), 1111–1120. https://doi.org/10.1641/0006-3568(2002)052[1111:Recapi]2.0.Co;2

Pleim, J. E. (2007). A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, *46*(9), 1383–1395. https://doi.org/10.1175/Jam2539.1

Posselt, D. J., Fryxell, B., Molod, A., & Williams, B. (2016). Quantitative sensitivity analysis of physical parameterizations for cases of deep convection in the NASA GEOS-5. *Journal of Climate*, *29*(2), 455–479. https://doi.org/10.1175/Jcli-D-15-0250.1

Qian, Y., Jackson, C, Giorgi, F., Booth, B., Duan, Q., Forest, C., et al. (2016). Uncertainty quantification in climate modeling and projection. *Bulletin of the American Meteorological Society*, *97*(5), 821–824. https://doi.org/10.1175/BAMS-D-15-00297.1

Qian, Y., Yan, H., Hou, Z., Johannesson, G., Klein, S., Lucas, D., et al. (2015). Parametric sensitivity analysis of precipitation at global and local scales in the Community Atmosphere Model CAM5. *Journal of Advances in Modeling Earth Systems*, *7*, 382–411. https://doi.org/10.1002/2014ms000354

Schumann, U., & Gerz, T. (1995). Turbulent mixing in stably stratified shear flows. *Journal of Applied Meteorology*, *34*(1), 33–48. https://doi.org/10.1175/1520-0450-34.1.33

Siuta, D., West, G., & Stull, R. (2017). WRF hub-height wind forecast sensitivity to PBL scheme, grid length, and initial condition choice in complex terrain. *Weather and Forecasting*, *32*(2), 493–509. https://doi.org/10.1175/Waf-D-16-0120.1

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). A description of the advanced research WRF version 3.. (NCAR Tech. Note, NCAR/TN-475+STR, 113 pp.).

Stull, R. (1988). *An introduction to boundary layer meteorology* (666 pp.). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-009-3027-8

Troen, I., & Mahrt, L. (1986). A simple-model of the atmospheric boundary-layer—Sensitivity to surface evaporation. *Boundary-Layer Meteorology*, *37*(1–2), 129–148. https://doi.org/10.1007/Bf00122760

Vickers, D., & Mahrt, L. (2004). Evaluating formulations of stable boundary layer height. *Journal of Applied Meteorology*, *43*(11), 1736–1749. https://doi.org/10.1175/Jam2160.1

Williamson, D., Blaker, A. T., Hampton, C., & Salter, J. (2015). Identifying and removing structural biases in climate models with history matching. *Climate Dynamics*, *45*(5–6), 1299–1324. https://doi.org/10.1007/s00382-014-2378-z

Xie, B., & Fung, J. C. H. (2014). A comparison of momentum mixing models for the planetary boundary layer. *Journal of Geophysical Research: Atmospheres*, *119*, 2079–2091. https://doi.org/10.1002/2013JD020273

Yan, H. P., Qian, Y., Zhao, C., Wang, H. L., Wang, M. H., Yang, B., et al. (2015). A new approach to modeling aerosol effects on east Asian climate: Parametric uncertainties associated with emissions, cloud microphysics, and their interactions. *Journal of Geophysical Research: Atmospheres*, *120*, 8905–8924. https://doi.org/10.1002/2015JD023442

Yang, B., Qian, Y., Berg, L. K., Ma, P. L., Wharton, S., Bulaevskaya, V., et al. (2017). Sensitivity of turbine-height wind speeds to parameters in planetary boundary-layer and surface-layer schemes in the Weather Research and Forecasting model. *Boundary-Layer Meteorology*, *162*(1), 117–142. https://doi.org/10.1007/s10546-016-0185-2

Yang, B., Qian, Y., Lin, G., Leung, R., & Zhang, Y. (2012). Some issues in uncertainty quantification and parameter tuning: A case study of convective parameterization scheme in the WRF regional climate model. *Atmospheric Chemistry and Physics*, *12*(5), 2409–2427. https://doi.org/10.5194/acp-12-2409-2012

Yang, B., Zhang, Y. C., Qian, Y., Wu, T. W., Huang, A. N., & Fang, Y. J. (2015). Parametric sensitivity analysis for the Asian summer monsoon precipitation simulation in the Beijing climate center AGCM, version 2.1. *Journal of Climate*, *28*(14), 5622–5644. https://doi.org/10.1175/Jcli-D-14-00655.1

Yang, Q., Berg, L. K., Pekour, M., Fast, J. D., Newsom, R. K., Stoelinga, M., & Finley, C. (2013). Evaluation of WRF-predicted near-hub-height winds and ramp events over a Pacific northwest site with complex terrain. *Journal of Applied Meteorology and Climatology*, *52*(8), 1753–1763. https://doi.org/10.1175/Jamc-D-12-0267.1

Young, G. S. (1988). Turbulence structure of the convective boundary layer. Part II: Phoenix 78 aircraft observations of thermals and their environment. *Journal of the Atmospheric Sciences*, *45*(4), 727–735. https://doi.org/10.1175/1520-0469(1988)045<0727:TSOTCB>2.0.CO;2

Zhao, C., Liu, X., Qian, Y., Yoon, J., Hou, Z., Lin, G., et al. (2013). A sensitivity study of radiative fluxes at the top of atmosphere to cloud-microphysics and aerosol parameters in the community atmosphere model CAM5. *Atmospheric Chemistry and Physics*, *13*(21), 10,969–10,987. https://doi.org/10.5194/acp-13-10969-2013