WRF Model Study of the Great Plains Low-Level Jet: Effects of Grid Spacing and Boundary Layer Parameterization

ELIZABETH N. SMITH

School of Meteorology, and Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

JEREMY A. GIBBS

Department of Mechanical Engineering, University of Utah, Salt Lake City, Utah

EVGENI FEDOROVICH AND PETRA M. KLEIN

School of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 23 December 2017, in final form 16 August 2018)

ABSTRACT

Previous studies have shown that the Weather Research and Forecasting (WRF) Model often underpredicts the strength of the Great Plains nocturnal low-level jet (NLLJ), which has implications for weather, climate, aviation, air quality, and wind energy in the region. During the Lower Atmospheric Boundary Layer Experiment (LABLE) conducted in 2012, NLLJs were frequently observed at high temporal resolution, allowing for detailed documentation of their development and evolution throughout the night. Ten LABLE cases with observed NLLJs were chosen to systematically evaluate the WRF Model's ability to reproduce the observed NLLJs. Model runs were performed with 4-, 2-, and 1-km horizontal spacing and with the default stretched vertical grid and a nonstretched 40-m vertically spaced grid to investigate which grid configurations are optimal for NLLJ modeling. These tests were conducted using three common boundary layer parameterization schemes: Mellor–Yamada Nakanishi Niino, Yonsei University, and Quasi-Normal Scale Elimination. It was found that refining horizontal spacing does not necessarily improve the modeled NLLJ wind. Increasing the number of vertical levels on a non-stretched grid provides more information about the structure of the NLLJ with some schemes, but the benefit is limited by computational expense and model stability. Simulations of the NLLJ were found to be less sensitive to boundary layer parameterization than to grid configuration. The Quasi-Normal Scale Elimination scheme was chosen for future NLLJ simulation studies.

1. Introduction

The nocturnal low-level jet (NLLJ) is defined as a maximum in the vertical profile of wind speed occurring overnight in the lowest kilometer of the atmosphere (e.g., Stull 2012). Several instances of such wind maxima have been observed around the world including, but not limited to, the Koorin jet (e.g., Brook 1985) and Southerly Buster (e.g., Baines 1980) in Australia and the Somali jet of East Africa (e.g., Ardanuy 1979). The Great Plains NLLJ is a common phenomenon in the United States, is typically southerly, and most often occurs during the warm months of the year. Bonner's (1968) Great Plains NLLJ climatology indicated that the average height of NLLJ wind maxima was approximately 800 m above ground level. This estimate was updated by Whiteman et al. (1997) using data with better height and time resolution to show that half of Great Plains NLLJ wind maxima are located below 500 m. Winds in the NLLJ can reach speeds that are 70% higher than the previous day's geostrophic wind speed (Shapiro and Fedorovich 2010).

Great Plains NLLJs are an important meteorological phenomenon due to their influence on weather and climate over a region larger than their immediate influence (Stensrud 1996). Gulf of Mexico moisture can be transported northward over the central United States by NLLJs. This transport has been related to the observed nocturnal maximum in warm-season rainfall

Corresponding author: Elizabeth N. Smith, elizabeth.n.smith@ou.edu

DOI: 10.1175/JAMC-D-17-0361.1

recorded over the central United States (Markowski and Richardson 2011). Great Plains NLLJs have also been cited as providing support for thermodynamic and dynamic features important to the initiation of deep convection and severe weather (Pitchford and London 1962; Bonner 1966; Maddox 1983; Astling et al. 1985; Trier et al. 2006). More recently, horizontal heterogeneity and nonstationarity in the structure of the Great Plains NLLJ itself have been identified as precursors for nocturnal convection initiation (Gebauer et al. 2018). The strong wind shear associated with NLLJs presents a particular hazard for aviation (National Research Council 1983). Wind energy is another industry impacted by the Great Plains NLLJ, as the increased winds can be a boon to energy production or a hazard due to shear and turbulence impacting wind turbines (e.g., Sisterson and Frenzen 1978; Storm et al. 2009). NLLJs are also of theoretical importance as fluid dynamical phenomena due to the peculiarity of physical mechanisms associated with their formation including inertial oscillations (e.g., Blackadar 1957), thermodynamic effects over sloping terrain (e.g., Holton 1967), buoyancy effects (e.g., Shapiro and Fedorovich 2009), or some combination of various effects (e.g., Shapiro et al. 2016), all of which are difficult to observe or quantify.

Studies of NLLJs with mesoscale models have helped improve the understanding of these phenomena. In particular, Zhong et al. (1996) showed that soil moisture changes impacted jet amplitudes with drier soils leading to stronger NLLJs. Pan et al. (2004) described the role of slope-induced horizontal temperature gradients in the formation of the NLLJ. It was shown that the evolving stable boundary layer (SBL) and embedded NLLJs are generally not well captured by numerical models. Consistent underestimation of the magnitude and depth of the NLLJ and minimal sensitivity to planetary boundary layer (PBL) parameterization schemes was reported by Storm et al. (2009) in a study using the Weather Research and Forecasting (WRF; Skamarock et al. 2008) Model on a 4-km horizontally spaced grid with 10 points below 1 km. Steeneveld et al. (2008) found similar effects in their simulations of the NLLJ using three state-of-the-art mesoscale models. On the other hand, recent studies have shown improved agreement between observed and simulated NLLJ features using the WRF Model (e.g., Vanderwende et al. 2015; Klein et al. 2016). In particular, Mirocha et al. (2016) found the prediction of wind speed near wind turbine height by the WRF Model to be sensitive to PBL scheme choice, but noted that the degree of sensitivity was also related to the choice of other model options such as the grid configuration.

While much progress has been made in recent decades, simulating the SBL remains difficult. Such difficulties include the applicability and validity of Monin–Obukov similarity theory, limited vertical resolution causing issues with thin surface layers and flux-gradient relationships, and interactions with mesoscale motions such as gravity waves and cold air drainage flows (Mahrt 1998). A review of boundary layer representation in modern weather models by Holtslag et al. (2013) points toward the complexity of PBL parameterization approaches being further muddled by differing schools of thought and historical influences. Current PBL and turbulence parameterization schemes are often inadequate, especially when applied to the SBL. These deficiencies are attributed, at least in part, to insufficient understanding of turbulent exchange processes in the SBL (Steeneveld et al. 2008). The literature suggests that simulations of the SBL and NLLJs can improve when model settings are chosen carefully with the relevant application in mind. As such, horizontal grid spacing, vertical grid spacing, and PBL parameterization approaches should be explored in more detail.

While the use of horizontal grid spacing at or below 1 km is increasingly feasible as computational resources improve, such scales often fall into the so-called *terra incognita*, where refined spacing can result in unrealistic flow structures in the daytime boundary layer (Wyngaard 2004; LeMone et al. 2013; Zhou et al. 2014). Since sensitivities to horizontal grid spacing are most likely related to partially resolving the largest eddies in the daytime boundary layer, such sensitivities may not be as pronounced in the overnight period when eddies are smaller. However, the availability of computational resources coupled with the natural assumption that more grid points add more information means that the impact of horizontal grid spacing is still worth testing. Improvements in the vertical grid resolution may be more important for simulating NLLJs since sharp gradients are difficult to resolve with coarser grids (Mirocha et al. 2016). Finally, the optimal grid configurations may be sensitive to PBL scheme choice.

PBL schemes parameterize the turbulent transport of momentum, heat, and scalars in the vertical direction within and beyond the PBL. Different PBL parameterization approaches may produce different characteristic SBLs and thus have an impact on NLLJ simulations. As stated before, known deficiencies in the performance of PBL schemes in the SBL environment can be attributed to our incomplete knowledge of turbulent exchange processes in the SBL (e.g., Salmond and McKendry 2005; Steeneveld et al. 2008; Fernando and Weil 2010; Steeneveld 2014). The PBL schemes used in the WRF Model may be conventionally categorized as local or nonlocal (e.g., Stull 2012, chapter 6). Local closure schemes only allow the vertical levels that are in the immediate vicinity of a particular level to directly impact the variables at that level. Conversely, nonlocal closure schemes allow for interactions between remote

vertical levels. Local schemes are often thought to offer some disadvantage for application to the PBL since local gradients of meteorological variables are not always representative of the processes that extend throughout the entire layer. Nonlocal schemes are generally assumed to better represent the effect of the large turbulent eddies than local schemes, and thus to better account for deep PBL circulations that are typical under convective conditions. However, such characterizations of each closure type may not be relevant for SBLs since in stable conditions all schemes act as local schemes. This study aims to address some of the outstanding challenges described above by performing tests that include a local transport scheme [Mellor–Yamada–Nakanishi–Niino (MYNN); Nakanishi and Niino 2009], a nonlocal transport scheme [Yonsei University (YSU); Hong et al. 2006], and a scale elimination scheme [Quasi-Normal Scale Elimination (QNSE); Sukoriansky et al. 2005], which acts locally.

Our study is motivated by initial experimental findings from the Plains Elevated Convection at Night (PECAN; Geerts et al. 2017) experiment, which indicate that the spatial evolution of the NLLJ throughout the PECAN domain may be important for modulating the turbulence and thermal structure in the boundary layer (e.g., Gebauer et al. 2018). Sudden local changes in stability and turbulent mixing may be primarily triggered by veering and advection of the NLLJ from west to east (Smith et al. 2018). Combining the sparse observations of NLLJs collected during PECAN with numerical model output is critical for investigating these effects in more detail. Our objective was thus to find an optimal WRF Model configuration for simulating NLLJs in the Great Plains. For this model evaluation study, we focused on selected cases from the Lower Atmospheric Boundary Layer Experiment (LABLE) campaign. LABLE took place at the Southern Great Plains (SGP) site of the Department of Energy Atmospheric Radiation Measurement (ARM) Program in Lamont, Oklahoma, during the autumn of 2012 and the summer of 2013 to investigate the dynamic, thermal, and turbulence characteristics in the atmospheric boundary layer (Klein et al. 2015). LABLE datasets provided detailed continuous observations that captured the NLLJ evolution throughout the entire night, while PECAN NLLJ observations were limited to only four nights, and often ended before the early morning transition impacted the NLLJ. To find an optimal configuration in which to run the WRF Model for NLLJ studies, we chose 10 LABLE cases during which NLLJs were observed. These 10 cases were classified following Bonin (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*) as either strongly or weakly turbulent NLLJs. Two example cases (one weakly and one strongly turbulent NLLJ) and model setup are described first in the subsequent sections. These two cases are then used to explore the sensitivity of the model to horizontal grid spacing, vertical grid spacing, and PBL scheme. After detailed results are shown for the two selected cases, the results are verified using the 10-case dataset. Conclusions are presented in the final section.

2. LABLE cases

Ten LABLE cases of observed NLLJs were considered in this study. While more cases would have been desirable, we choose to retain only these 10 cases as they were not directly impacted by boundaries or other mesoscale features and to prevent any complications from including different seasons or locations in the analysis. Figure 1 shows the maximum 10-case mean normalized NLLJ magnitude in the lowest 1 km at three different times from North American Mesoscale Forecast System (NAM) analysis data. For these cases, the LABLE observation site was located within the core region of the NLLJ. Among many other instruments, LABLE deployed a Doppler lidar and an Atmospheric Emitted Radiance Interferometer (AERI) to observe temperature and wind profiles. Conical Doppler lidar scans occurred every 15 min using 18-m range gates and a 70° scan elevation. Velocity azimuthal display analysis permitted retrieval of the horizontal wind from these data. Between other scans, the Doppler lidar collected continuous vertical stare data to measure vertical velocity. The AERI observations allowed for retrieval of vertical thermodynamic profiles with 5-min temporal resolution (Turner and Löhnert 2014). More details on the instrument operations during LABLE can be found in Klein et al. (2015). For reference, the local time at the observation site was central daylight time (CDT; UTC − 5 h). These observations provide high-resolution information about NLLJs throughout the night, which allows evaluation of model representation of processes relevant to NLLJs (NLLJ magnitude, NLLJ evolution, SBL structure, etc.). In analyses completed by Bonin (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*), LABLE NLLJs were classified based on the observed near-surface turbulence, which was quantified in terms of the vertical velocity variance σ_w retrieved from the vertical stare Doppler lidar data. The turbulence regime is defined by the mean measured value of σ_w below 100 m ($\overline{\sigma_w}_{100\text{m}}$). It is important to note that the lowest usable range gate is near 63 m, so the effective mean is from 63 to 100 m. Bonin et al. (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*) identified a threshold of $\overline{\sigma_w}_{100\text{m}} = 0.2\,\text{m s}^{-1}$, above which a weakly stable boundary layer existed and below which a very stable boundary layer existed. Bonin (2015) and Bonin et al. (2018,

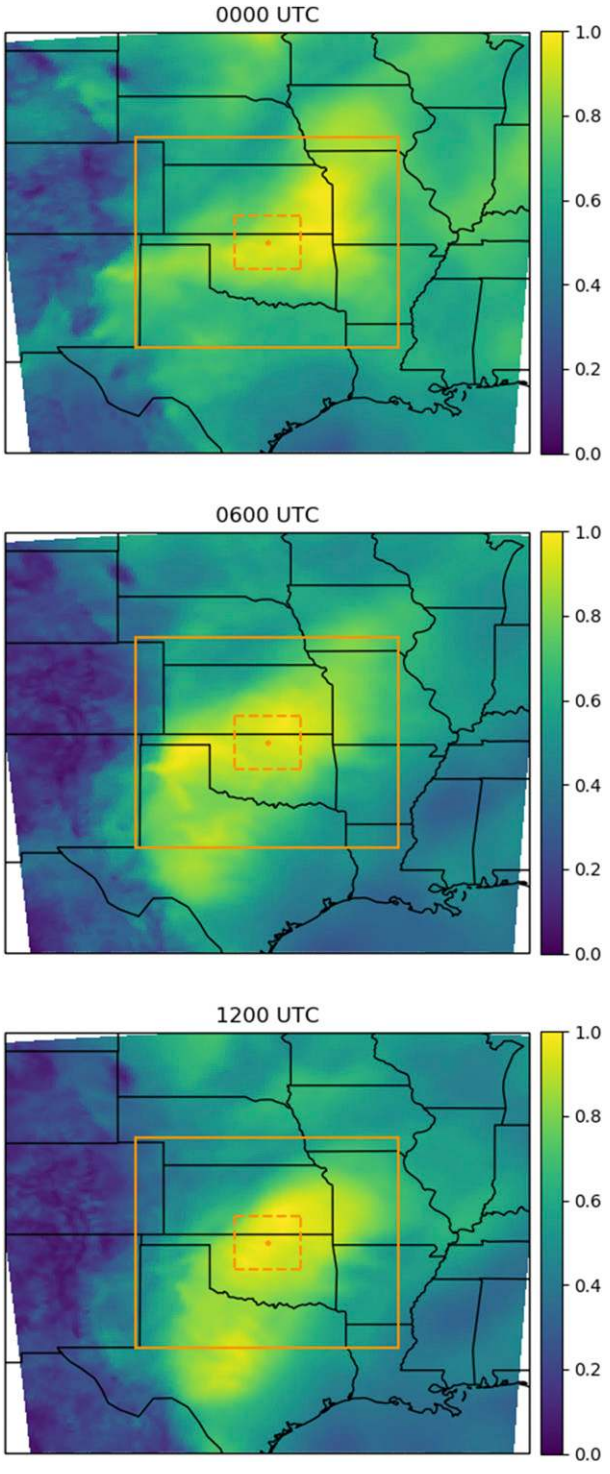

FIG. 1. Color fill shows the maximum mean normalized NAM analysis NLLJ magnitude in the lowest 1 km for all 10 cases at three times, where a value of 1 indicates the maximum wind speed. The solid orange box shows the 1024 km × 1024 km simulation domain, and the dashed orange box shows the 256 km × 256 km simulation domain. The star marks the location of the ARM-SGP site where LABLE observations were collected.

TABLE 1. Ten LABLE cases classified on the basis of Bonin (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*). The case marked by an asterisk was classified using data from the ARM Doppler lidar rather than the OU Doppler lidar (see the footnote in section 2).

Date	$\overline{\sigma_{w100\mathrm{m}}}$ ($\mathrm{m^2\,s^{-2}}$)	Classification
9 Oct 2012	0.10	Weakly turbulent
13 Oct 2012	0.27	Strongly turbulent
15 Oct 2012	0.01*	Weakly turbulent
16 Oct 2012	0.47	Strongly turbulent
21 Oct 2012	0.07	Weakly turbulent
22 Oct 2012	0.52	Strongly turbulent
23 Oct 2012	0.47	Strongly turbulent
24 Oct 2012	0.54	Strongly turbulent
8 Nov 2012	0.07	Weakly turbulent
9 Nov 2012	0.50	Strongly turbulent

manuscript submitted to *Bound.-Layer Meteor.*) then used the nighttime (0200–1200 UTC) mean $\overline{\sigma_{w100\mathrm{m}}}$ to categorize NLLJs as weakly turbulent ($\overline{\sigma_{w100\mathrm{m}}} < 0.2\,\mathrm{m\,s^{-1}}$) or strongly turbulent ($\overline{\sigma_{w100\mathrm{m}}} > 0.2\,\mathrm{m\,s^{-1}}$). The 10 LABLE cases and their classifications are shown in Table 1. Six cases were classified as strongly turbulent while four cases were classified as weakly turbulent.[1]

In addition to the difference in near-surface turbulence, Bonin (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*) found that weakly and strongly turbulent NLLJs have other defining features. Weakly turbulent NLLJs generally form under more weakly synoptic forced conditions, with shortwave troughs at 850 hPa often positioned a few hundred kilometers west, and do not form until a strong surface-based inversion forms. Strongly turbulent NLLJs often form with large-scale 500-hPa troughs over the west coast of the United States. The strong large-scale synoptic forcing and associated southerly winds are present before sunset preventing strong surface-based inversions from forming. Furthermore, the large-scale troughs propagate more slowly than shortwaves and can thus result in several consecutive nights of strongly turbulent NLLJs. The NLLJ height and maximum wind speed also evolve differently for strongly and weakly turbulent cases. Strongly turbulent NLLJs have a relatively constant height and reach a maximum magnitude by about 0500 UTC. On the other hand, weakly turbulent NLLJs tend to follow isentropic surfaces with the height changing in time and slowly strengthen

[1] The case on 15 Oct 2012 was classified using data from the ARM Doppler lidar rather than the University of Oklahoma (OU) Doppler lidar because the OU Doppler lidar data were too noisy. The ARM Doppler lidar range gate was longer and thus averages out smaller scales of turbulence, resulting in lower $\overline{\sigma_{w100\mathrm{m}}}$ values than those from the OU Doppler lidar. However, this value is so small that misclassification is unlikely.

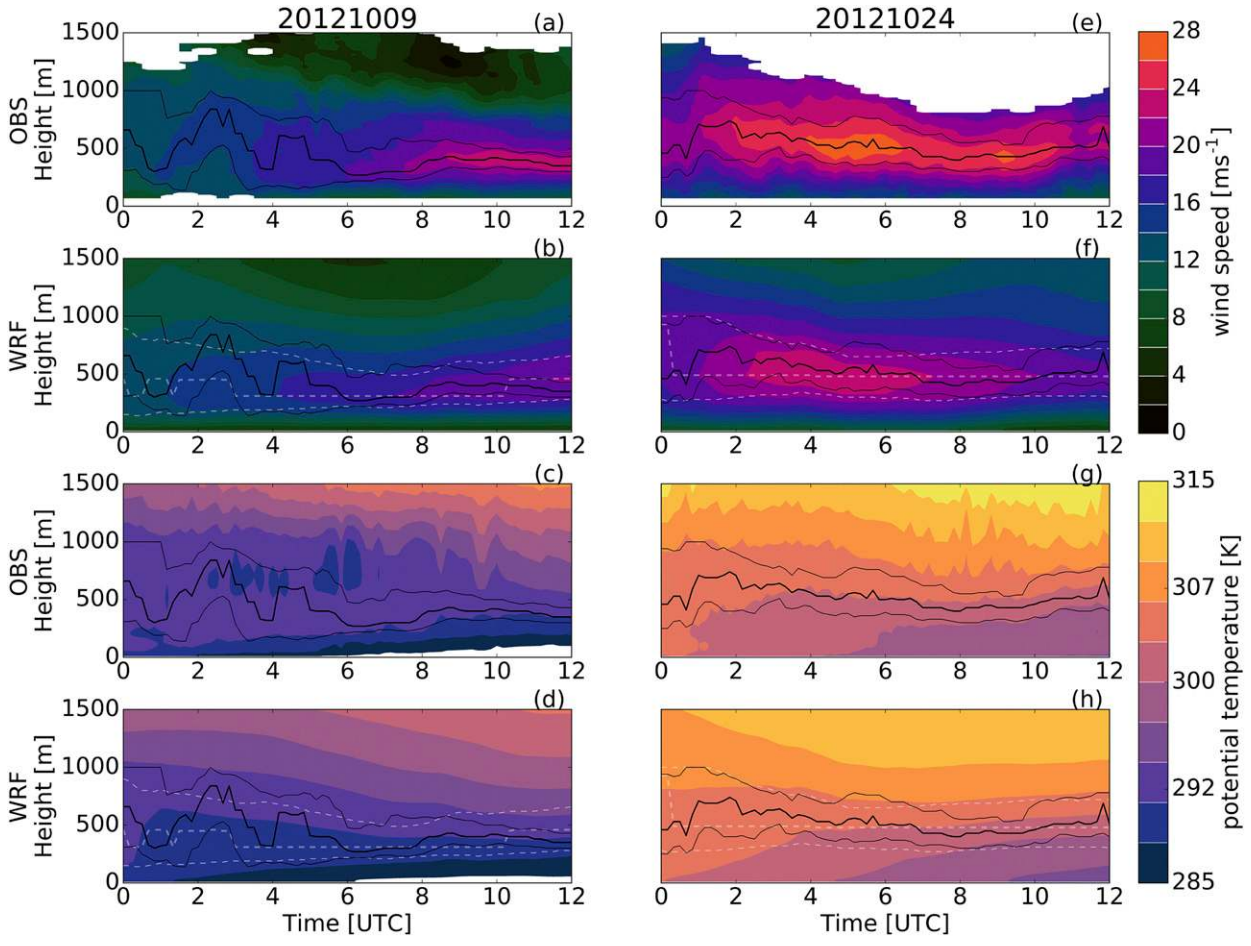

FIG. 2. Time–height cross sections of observed and YSU4kmDef WRF simulated data for (left) 9 and (right) 24 Oct 2012: (a),(e) observed and (b),(f) simulated wind speeds are shown together with (c),(g) observed and (d),(h) simulated potential temperatures. The heavy solid black and dashed white lines show the height of the observed and simulated NLLJ maximum, respectively, and the lighter solid black and dashed white lines encompass the region in which the observed and simulated wind speed, respectively, is at least 90% of the NLLJ maximum wind speed. Local time at the observation site was CDT (UTC − 5 h).

throughout the night reaching a maximum shortly before sunrise. More details about strongly and weakly turbulent NLLJs are available in Bonin (2015) and Bonin et al. (2018, manuscript submitted to *Bound.-Layer Meteor.*).

The NLLJ that formed on 9 October 2012 was classified as weakly turbulent. At 0000 UTC 9 October 2012, a surface low pressure system was positioned in the northern United States over northern Minnesota connected to a 500-hPa trough. An associated cold front draped southwest through the center of Nebraska. Near the observation site, 700- and 500-hPa flows were approximately zonal. A shortwave trough was located at 850 hPa over eastern New Mexico and Colorado. As the shortwave trough propagated east during the overnight hours, a secondary low pressure formed in western Kansas by 1200 UTC. The approaching trough and surface low increased the pressure gradient force and resulted in stronger synoptic-scale forcing overnight at the observation site. The observed wind speed is shown in Fig. 2a and observed potential temperature is shown in Fig. 2c. The NLLJ strengthened throughout the night until about 1000 UTC, while stability increased at the levels where the NLLJ formed. The core region of the NLLJ (the region where wind speed is at least 90% of the maximum value, shown by the thin black lines in Figs. 2a,c) became narrower with time. The height of the NLLJ also tended to follow isentropic surfaces through the night, with the height of the NLLJ increasing as the inversion grew deeper.

On 24 October 2012, a strongly turbulent NLLJ formed. This night was the third night of consecutive strongly turbulent NLLJs associated with a large-scale 500-hPa trough over the western United States. A closed low propagated from northeastern Colorado to western Kansas between 0000 and 1200 UTC as a cold front

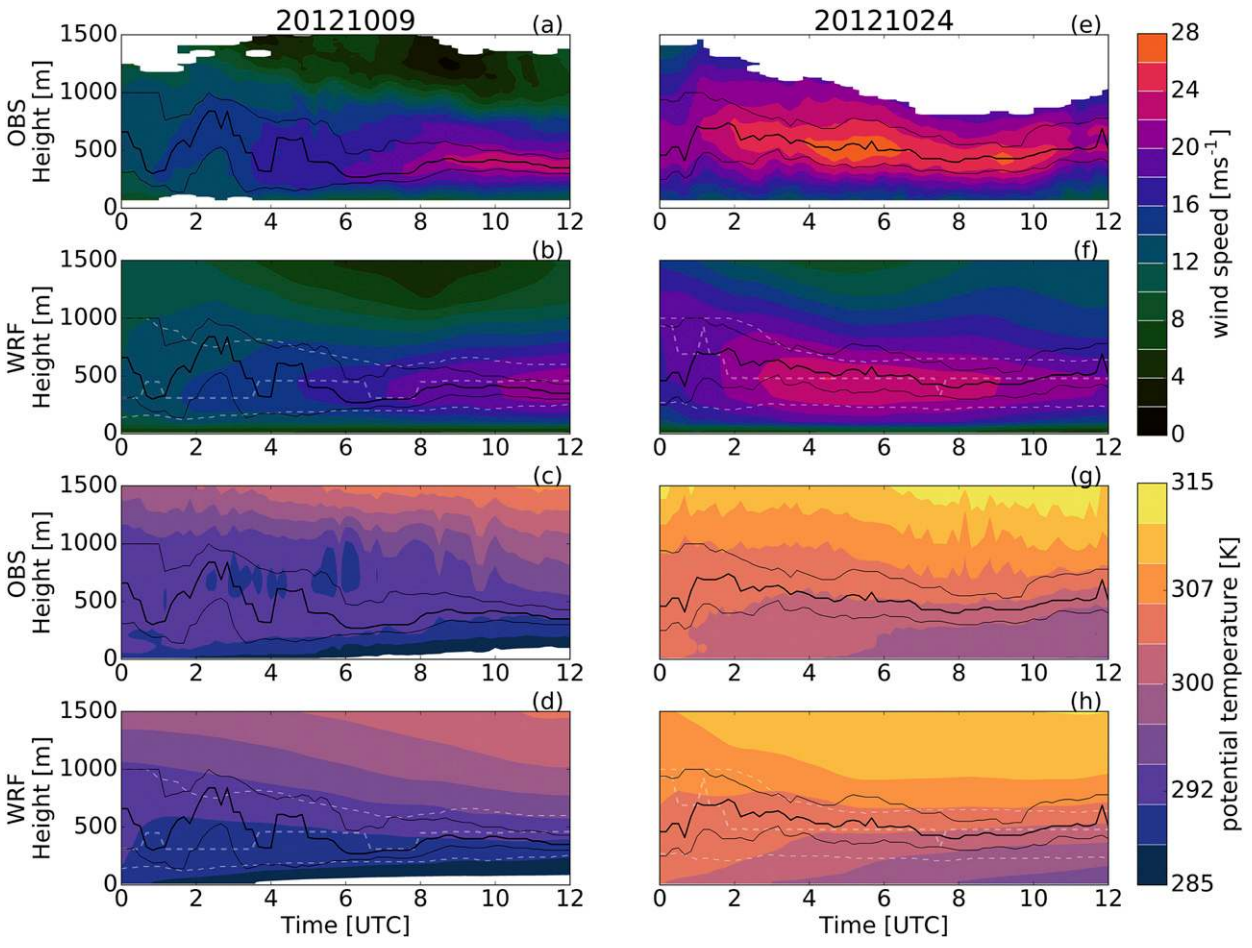

FIG. 3. As in Fig. 2, but the WRF simulation used the MYNN4kmDef configuration.

moved south into northwestern Kansas. The observed wind speed is shown in Fig. 2e, and observed potential temperature is shown in Fig. 2g. The NLLJ formed early in the period, strengthened until it reached a maximum just before 0400 UTC, and stayed at or near that magnitude until about 1000 UTC when the NLLJ began to weaken. The height of the NLLJ remained fairly constant through the night, but the core region (thin black lines in Figs. 2e,g) of the NLLJ did get narrower after the NLLJ reached maximum magnitude.

These two cases were selected as examples of a weakly and a strongly turbulent NLLJ for which model evaluation will be presented in detail. Once the results based on these two cases are described in detail, the model evaluation study is expanded using the results of the full 10-case dataset.

3. Model setup

In this study, the WRF Model (WRF-ARW, version 3.8.1) was applied to study the Great Plains NLLJ. The model domain was centered over the ARM-SGP site in Lamont, Oklahoma. The ARM-SGP site and the 1024 km × 1024 km and 256 km × 256 km model domains are shown in Fig. 1. Numerical experiments were run for the 10 LABLE cases presented in Table 1. Detailed descriptions of the model experiments are presented for the cases observed on 9 October 2012, a weakly turbulent NLLJ, and 24 October 2012, a strongly turbulent NLLJ (Figs. 2–4).

Each simulation was run for a 24-h period starting at 1200 UTC on the day preceding the NLLJ. The first 12 h of each simulation were used to accommodate model spinup effects. Simulations with additional spinup time were tested, but differences were found to be negligible beyond 12 h. The WRF Model was initialized using reanalysis data from the NAM with 12-km grid spacing. These NAM data were also used to generate lateral boundary conditions for the model every 6 h. While previous studies included land surface model physics in sensitivity tests (e.g., Mirocha et al. 2016), impacts of degraded land surface representation are likely

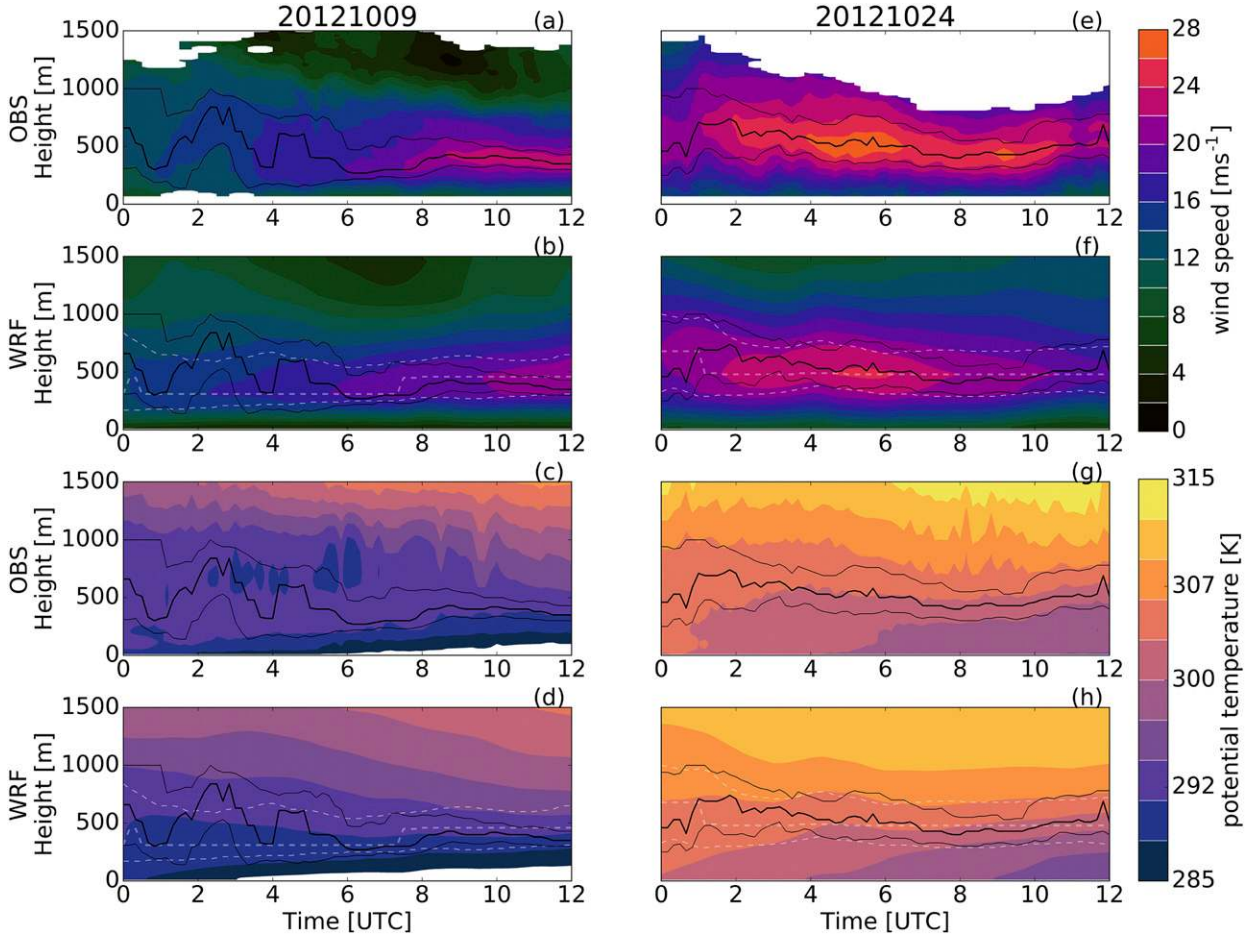

FIG. 4. As in Fig. 2, but the WRF simulation used the QNSE4kmDef configuration.

minimized during nighttime conditions due to decoupling from the surface (Mahrt 1999). It is important to note that the WRF Model is a limited area model, allowing the initial and boundary conditions to reflect the larger-scale conditions. Model settings are summarized in Table 2. Only the PBL parameterization and associated surface layer (SL) scheme, vertical grid spacing, horizontal grid spacing, and horizontal grid expanse were varied in this study.

The default vertical grid in the WRF Model employs 41 terrain-following normalized pressure levels, the so-called eta levels. These levels are stretched vertically in such a manner that the grid spacing increases with distance from the surface. For this study, the default vertical grid was chosen using ''out-of-the-box'' settings making as few changes to the input settings of the model as possible. In the simulations with the default grid configuration, the WRF stretching algorithm produced 10 vertical levels located within the lowest 2 km, as shown in the leftmost panel of Fig. 5. The first model level was located approximately 30 m above the ground and the model top was near 16 km. This default vertical grid is intended to act as a control configuration against which adjustments to the vertical grid spacing can be compared.

TABLE 2. WRF Model physics options used in this study.

Setting	Choice
Land surface model	Noah land surface model (Tewari et al. 2004)
Microphysics	WRF single-moment five-class scheme (Hong et al. 2004)
Shortwave radiation	Fifth-generation Penn. State–NCAR Mesoscale Model (MM5) shortwave radiation scheme (Dudhia 1989)
Longwave radiation	Rapid Radiative Transfer Model (Mlawer et al. 1997)
Horizontal diffusion	Horizontal deformation first-order closure scheme (Smagorinsky 1963)
PBL/SL	YSU PBL/revised MM5 Monin–Obukhov SL; MYNN PBL/MYNN SL; QNSE PBL/QNSE SL

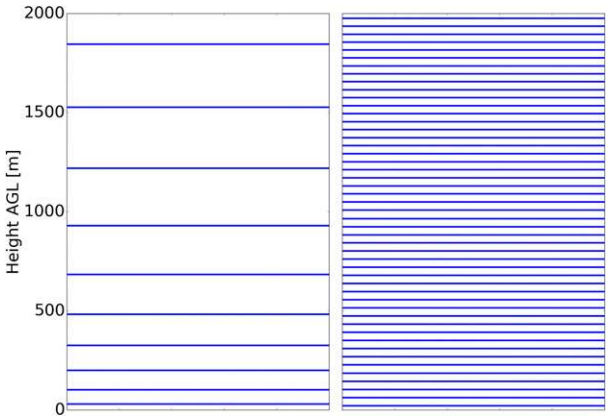

FIG. 5. The vertical spacing below 2 km for (left) default vertical spacing and (right) 40-m vertical spacing.

As discussed in section 1, three PBL schemes were evaluated for their ability to reproduce NLLJs: a local scheme (MYNN), a nonlocal scheme (YSU), and a scale-elimination scheme (QNSE), which can be classified as local. The MYNN level-2.5 parameterization is a scheme with a prognostic turbulence kinetic energy (TKE) equation and parameterized equations for turbulent variances/fluxes. Through the use of various length scales, MYNN induces extra mixing to account for the poor local approximation of fluxes based on eddy diffusivity. The YSU scheme is based on a first-order turbulence model that directly specifies the eddy diffusivity and explicitly treats turbulent entrainment, which is particularly important for convective conditions. Last, QNSE is a local closure with a prognostic TKE equation. The scheme uses the scale elimination approach to approximate exchange by interacting turbulent eddies, and was designed specifically for the SBL. The scale elimination approach uses successive averaging over small intervals of the turbulence spectrum providing expressions for turbulent viscosity and diffusivity as functions of local Richardson number.

WRF Model output was saved in the same time increments as Doppler lidar observations (i.e., every 15 min) to facilitate comparisons between the two. The first two range gates of the lidar were removed along with any other suspicious data after applying a signal-to-noise ratio filter to the data. AERI observations were interpolated to the same 15-min temporal resolution as Doppler lidar and WRF data. WRF Model data were extracted from the four grid cells nearest to the observational site and then spatially averaged in order to facilitate a comparison of the simulated and observed data. This method was chosen since the WRF Model uses a staggered grid. Averaging the four nearest boxes provides a way to include effects from each cell and to account for any small translation errors. Additionally, for the computation of performance metrics, the data were vertically interpolated to match the lidar levels. Performance metric comparisons between observations and simulations were limited to the lowest 800 m. The lidar data are not as trustworthy above this level and are often removed by data quality filters as lidar returns are dependent on aerosol load.

The investigated performance metrics include normalized bias,

$$\text{bias}_N = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(U_{\text{WRF}_i} - U_{\text{lidar}_i}\right)}{0.5\left(\overline{U}_{\text{WRF}} + \overline{U}_{\text{lidar}}\right)}, \tag{1}$$

normalized mean absolute error,

$$\text{MAE}_N = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|U_{\text{WRF}_i} - U_{\text{lidar}_i}\right|}{\left(\overline{U}_{\text{WRF}} + \overline{U}_{\text{lidar}}\right)}, \tag{2}$$

and normalized centered root-mean-square error,

$$\text{CRMSE}_N = \frac{\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\left(U_{\text{WRF}_i} - \overline{U_{\text{WRF}}}\right) - \left(U_{\text{lidar}_i} - \overline{U_{\text{lidar}}}\right)\right]^2\right\}^{1/2}}{\left(\sigma_{\text{WRF}}\sigma_{\text{lidar}}\right)^{1/2}}, \tag{3}$$

where index i loops over N, the number of points in the time/height domain, U_{WRF} and U_{lidar} refer to the WRF-simulated and lidar-observed wind speed, respectively, and overbars denote mean quantities. Standard deviations are represented by σ. "Bias" shows, on average, whether the WRF Model over or underestimates wind speed compared with observations, MAE represents the average difference between the WRF Model solutions and lidar data, and CRMSE describes the random component of the error since the mean bias has been removed. To facilitate comparison between cases, each metric is shown as a normalized quantity. The normalization method employed here is similar to that in Chang and Hanna (2004). To summarize the performance metrics, the mean metric (MM) is defined as

$$\text{MM} = \frac{\text{bias}_N + \text{MAE}_N + \text{CRMSE}_N}{3}. \tag{4}$$

Performance metric data for the 9 and 24 October 2012 cases are visualized in Fig. 6, with each value presented in Tables 3 and 4.

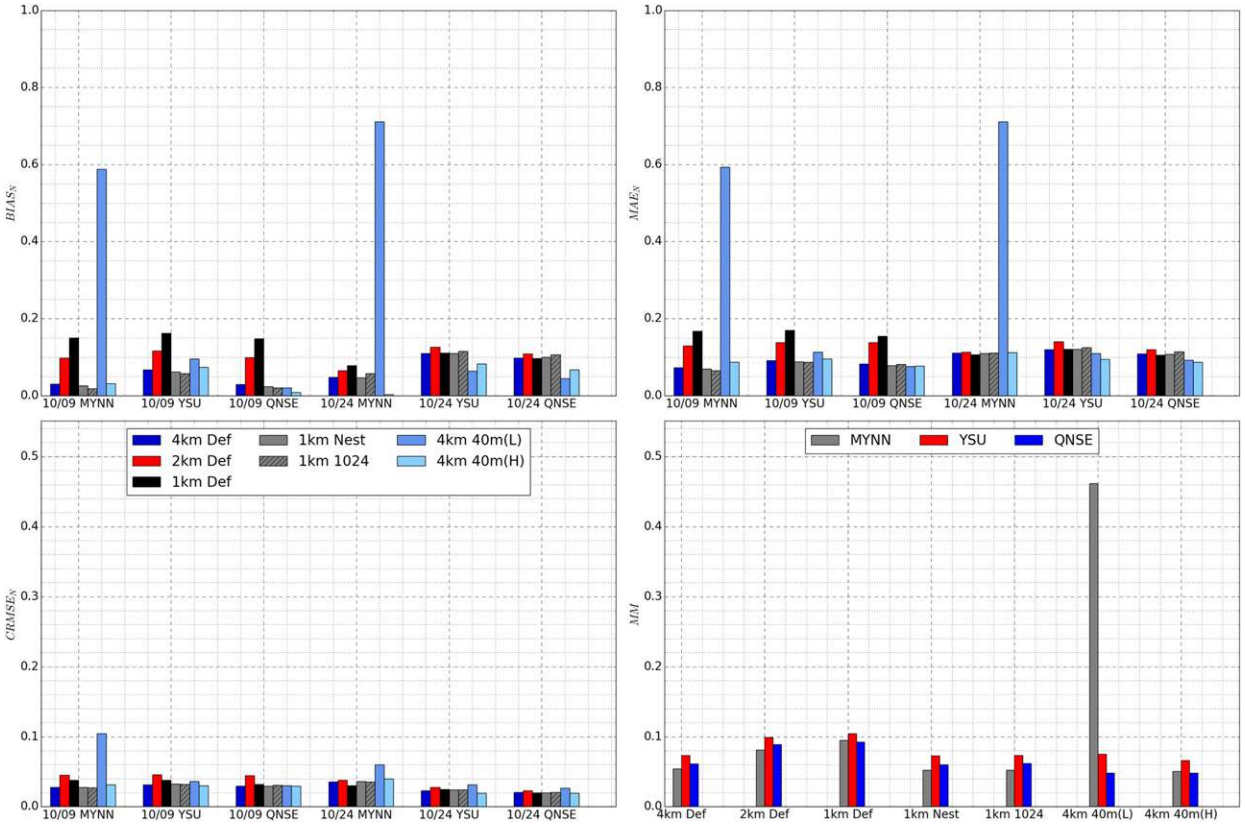

FIG. 6. Bar plots show the values of (top left) bias$_N$ (displayed as an absolute value), (top right) MAE$_N$, and (bottom left) CRMSE$_N$ for 9 and 24 October 2012. Also shown is (bottom right) MM for both cases for each grid configuration. Lower values indicate better performance.

The model configurations are labeled as PBL $\Delta h \Delta z$, where PBL indicates the PBL scheme (MYNN, YSU, or QNSE), Δh indicates the horizontal spacing (1, 2, or 4 km), and Δz indicates the vertical grid spacing [40 m or default (Def)]. For instance, a model run using the YSU PBL scheme with Δh = 1 km and the default stretched vertical grid is labeled YSU1kmDef.

4. Results

First, observations of wind speed and potential temperature for the two selected LABLE cases were compared with WRF simulations using 4-km horizontal spacing, the default stretched vertical grid, and varying PBL schemes (Figs. 2, 3, and 4). These WRF simulations represent the expected simulation of the NLLJ and SBL from a typical mesoscale model with a standard grid configuration like those found in previous studies. While a NLLJ is predicted with all three PBL schemes, details of SBL and NLLJ evolution are generally not well captured, and the magnitude and depth of the NLLJ are poorly represented as compared with the observations. Similar results hold across all 10 LABLE cases shown in Fig. 7. Runs using 4-km horizontal spacing and the default vertical grid tended to underestimate the wind speed maxima and evolved much more smoothly, missing sudden increases and decreases in the NLLJ maximum wind speed. With this standard grid configuration, no single PBL scheme can be qualified as the best to use for simulating the NLLJ. All schemes struggle more with the late forming weakly turbulent NLLJ on 9 October 2012 than with the strongly turbulent NLLJ on 24 October 2012. Tests conducted in our study have been designed to examine changes to this baseline configuration that would be beneficial for simulating the NLLJ and to find out if associated improvements are consistent for each PBL scheme. The overarching goal was to identify a suitable WRF Model configuration for NLLJ research applications.

a. *Horizontal grid spacing*

Simulations were conducted on numerical grids of size 256 × 256 × 41 with 1-, 2-, and 4-km horizontal spacing in order to test the sensitivity of model results to horizontal resolution. Runs were conducted for each SL–PBL combination described in section 3. First, our

TABLE 3. Performance evaluation for each WRF Model configuration is presented for 9 Oct 2012, where Δh (km) is horizontal grid spacing, npts is the number of grid points in the horizontal model domain, Δz (m) is vertical grid spacing, bias_N is the normalized mean error, MAE_N is normalized mean absolute error, and CRMSE_N is normalized centered root-mean-square error. For vertical grid spacing, "default" refers to the traditional stretched vertical grid used in the WRF Model.

PBL scheme	Δh (npts)	Δz (npts)	Bias_N	MAE_N	CRMSE_N
YSU	1 (256^2)	Default (41)	−0.16	0.17	0.04
	1 (256^2, nest)	Default (41)	−0.06	0.09	0.03
	1 (1024^2)	Default (41)	−0.06	0.09	0.03
	2 (256^2)	Default (41)	−0.12	0.14	0.05
	4 (256^2)	Default (41)	−0.07	0.09	0.03
	4 (256^2, low model top)	40 (101)	−0.10	0.11	0.04
	4 (256^2, high model top)	40 (126)	−0.07	0.10	0.03
MYNN	1 (256^2)	Default (41)	−0.15	0.17	0.04
	1 (256^2, nest)	Default (41)	−0.03	0.07	0.03
	1 (1024^2)	Default (41)	−0.02	0.07	0.03
	2 (256^2)	Default (41)	−0.10	0.13	0.04
	4 (256^2)	Default (41)	−0.03	0.07	0.03
	4 (256^2, low model top)	40 (101)	−0.59	0.59	0.10
	4 (256^2, high model top)	40 (126)	−0.03	0.09	0.03
QNSE	1 (256^2)	Default (41)	−0.15	0.15	0.03
	1 (256^2, nest)	Default (41)	−0.02	0.08	0.03
	1 (1024^2)	Default (41)	−0.02	0.08	0.03
	2 (256^2)	Default (41)	−0.10	0.14	0.04
	4 (256^2)	Default (41)	−0.03	0.08	0.03
	4 (256^2, low model top)	40 (101)	−0.02	0.08	0.03
	4 (256^2, high model top)	40 (126)	−0.01	0.08	0.03

discussion focuses on results from runs using the YSU scheme.

For the weakly turbulent 9 October 2012 case, the YSU4kmDef (Fig. 8d) configuration underestimated the magnitude of the NLLJ as mentioned previously, but the temporal evolution of the wind field overnight resembled the observed evolution. Errors were most obvious in the NLLJ core region. The YSU2kmDef (Fig. 8c) configuration resulted in a simulated wind field that was less representative of the observations than the

TABLE 4. As in Table 3, but for 24 Oct 2012.

PBL scheme	Δh (npts)	Δz (npts)	Bias_N	MAE_N	CRMSE_N
YSU	1 (256^2)	Default (41)	−0.11	0.12	0.03
	1 (256^2, nest)	Default (41)	−0.11	0.12	0.02
	1 (1024^2)	Default (41)	−0.11	0.12	0.02
	2 (256^2)	Default (41)	−0.13	0.14	0.03
	4 (256^2)	Default (41)	−0.11	0.12	0.02
	4 (256^2, low model top)	40 (101)	−0.06	0.11	0.03
	4 (256^2, high model top)	40 (126)	−0.08	0.09	0.02
MYNN	1 (256^2)	Default (41)	−0.08	0.11	0.03
	1 (256^2, nest)	Default (41)	−0.05	0.11	0.04
	1 (1024^2)	Default (41)	−0.06	0.11	0.04
	2 (256^2)	Default (41)	−0.07	0.11	0.04
	4 (256^2)	Default (41)	−0.05	0.11	0.04
	4 (256^2, low model top)	40 (101)	−0.71	0.71	0.06
	4 (256^2, high model top)	40 (126)	−0.01	0.11	0.04
QNSE	1 (256^2)	Default (41)	−0.10	0.11	0.02
	1 (256^2, nest)	Default (41)	−0.10	0.11	0.02
	1 (1024^2)	Default (41)	−0.11	0.11	0.02
	2 (256^2)	Default (41)	−0.11	0.12	0.02
	4 (256^2)	Default (41)	−0.10	0.11	0.02
	4 (256^2, low model top)	40 (101)	−0.04	0.09	0.03
	4 (256^2, high model top)	40 (126)	−0.07	0.09	0.02

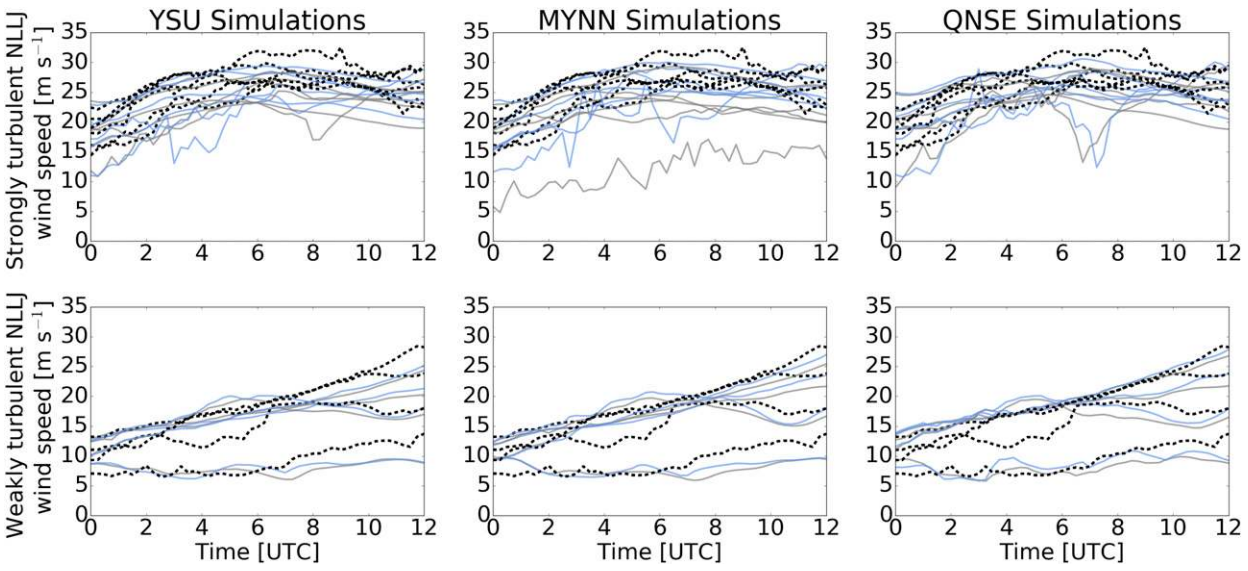

FIG. 7. NLLJ wind speed maxima plotted in time for the (top) strongly turbulent and (bottom) weakly turbulent NLLJ cases listed in Table 1. Black dashed lines show the observed wind speed maxima. Gray lines show NLLJ wind speed maxima from simulations using 4-km horizontal spacing and the default vertically stretched grid. Blue lines show NLLJ wind speed maxima from simulations using 4-km horizontal spacing and the 40-m nonstretched vertical grid (with a high model top). Each column shows simulations using a different PBL scheme.

simulation with 4-km horizontal spacing. The magnitude and structure of the NLLJ were poorly captured, and there was little detail provided about the temporal evolution of the wind field. Last, the YSU1kmDef (Fig. 8b) configuration captured the timing of the NLLJ, but failed to predict a NLLJ with structure and magnitude comparable to the observed NLLJ. The predicted height of the NLLJ maximum also appeared too low compared with the observations. Simulations with 4-km spacing generally predicted the magnitude of the NLLJ better than the simulations with 1- or 2-km spacing, but large errors were still observed in the core region. The bias_N and MAE_N values in Table 3 and Fig. 6 agree with this analysis since each value decreases as horizontal spacing increases.

For the strongly turbulent 24 October 2012 case, the YSU4kmDef (Fig. 8h) configuration predicted a NLLJ that is visually most similar to observations. The NLLJ magnitude was still underestimated, but the jet was generally stronger, and its temporal evolution was better captured than in the YSU1kmDef or YSU2kmDef simulations. The performance metrics were slightly improved compared with metrics of other considered configurations (Table 4 and Fig. 6). The YSU2kmDef (Fig. 8g) run offered some information about temporal evolution by depicting a slight descent of the jet at 0700 UTC and realistic flow structure early in the simulation window. However, it is clear that the jet strength was still underestimated, which is especially true for the region around the jet maximum. Increased values of bias_N, MAE_N, and CRMSE_N (Table 4 and Fig. 6) suggested a degradation of the YSU2kmDef modeled jet as compared with the YSU4kmDef run. The YSU1kmDef (Fig. 8f) configuration generally captured the timing and vertical extent of the jet. However, the model failed to reproduce the slow descent and subsequent ascent of the jet between 0700 and 0900 UTC. In addition, the YSU1kmDef configuration delayed the erosion and lifting of the jet. Metrics in Table 4 and Fig. 6 show that the YSU1kmDef run was slightly improved over the YSU2kmDef, but the YSU4kmDef run still fared the best.

Each simulation discussed in the previous tests of horizontal spacing used numerical meshes with 256 × 256 points in the horizontal. Accordingly, the mesh with 1-km horizontal grid spacing spanned a 256-km region (dashed box shown in Fig. 1), while the mesh with 4-km horizontal grid spacing covered a 1024-km domain (solid box shown in Fig. 1). It is possible to suppose that allowing flow features more space to develop within the model domain may improve the performance of the simulations with 1-km horizontal grid spacing. To investigate whether the mesh size, and thus also the proximity of the interior grid cells to the boundaries, had an impact on the modeled NLLJ, results of model runs in two domains with a horizontal expanse of 1024 km were compared: in the domain of size 256 × 256 using 4-km horizontal grid spacing, and in the domain of size 1024 × 1024 with 1-km horizontal grid spacing.

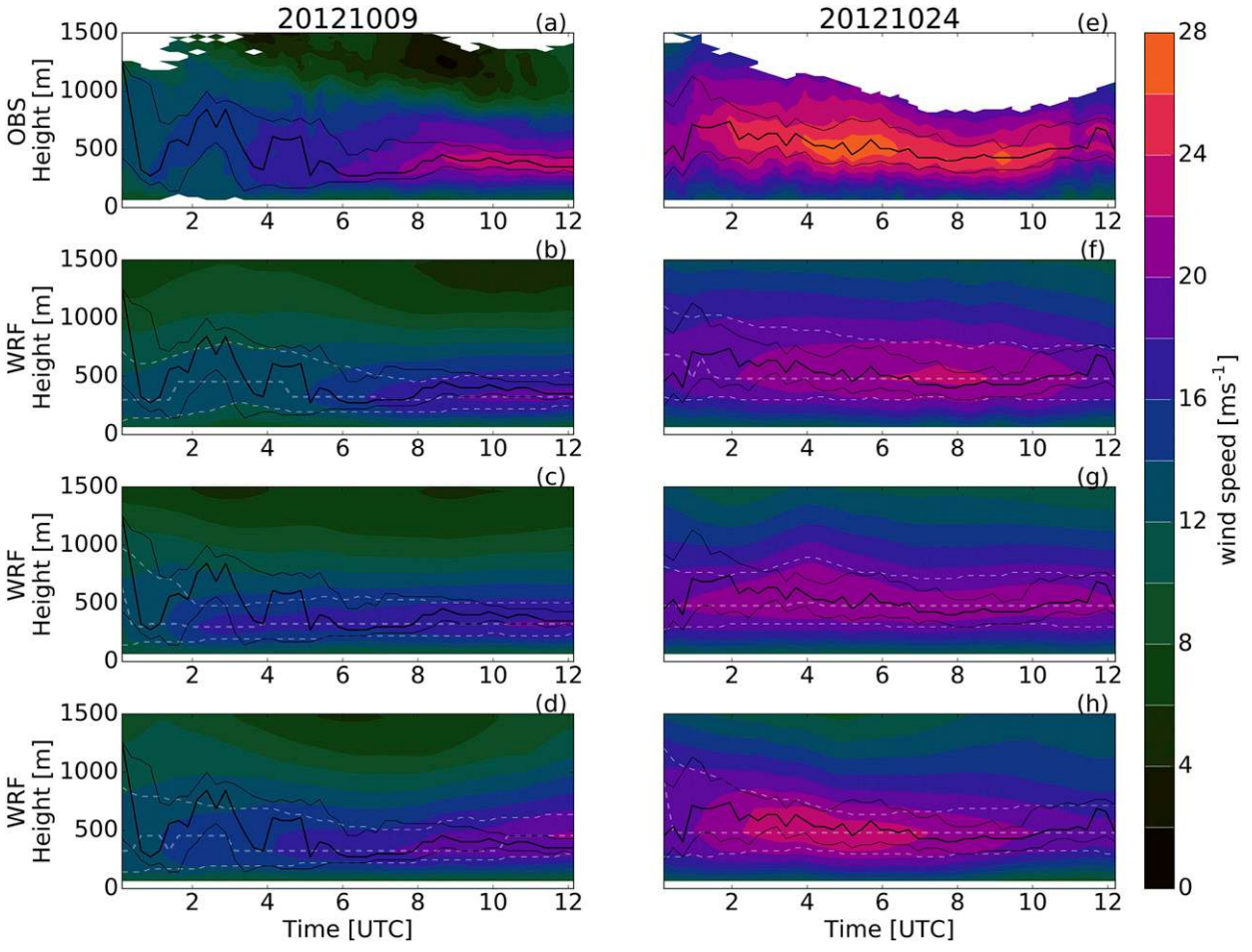

FIG. 8. Time–height cross sections of observed and WRF simulated data from (left) 9 Oct 2012 and (right) 24 Oct 2012 for (a),(e) observed wind speed and simulations with (b),(f) 1-, (c),(g) 2-, and (d),(h) 4-km horizontal spacing, all using the default stretched vertical grid and the YSU PBL scheme. Heavy solid black and dashed white lines show the height of the observed and simulated NLLJ maximum, respectively. Lighter solid black and dashed white lines encompass the region in which the observed or simulated wind speed, respectively, is at least 90% of the NLLJ maximum wind speed.

The use of one-way nesting was also investigated. As stated previously, data from the 12-km NAM provided boundary conditions to the WRF Model every 6 h. This means that the 12-km NAM reanalysis data used for setting the initial and boundary conditions were downscaled differently to match the 4-km grid than they were to match the 1-km grid. To evaluate the impact that this downscaling difference had on the resulting wind fields, a grid with 1-km horizontal spacing was nested within a grid with 4-km spacing. Thus, the boundaries of the 256 × 256 grid with 1-km horizontal spacing (dashed box shown in Fig. 1) were driven by 4-km spaced data rather than 12-km spaced data.

Results for the 9 October 2012 case (see Table 3 and Fig. 6) indicate that the 1-km 1024-point simulation performed comparably with the model runs using the 4-km 256-point domain and offered some improvement over the 1-km, 256-point domain. For example, bars depicting bias_N in Fig. 6 show a modest difference between the 1-km 1024-point simulation and the 4-km 256-point simulation, but a much larger difference exists between the 1-km 256-point domain and the 1-km 1024-point domain. This implies that choosing an appropriate horizontal grid size and considering a location's proximity to the simulation domain boundaries can be just as important as selecting an appropriate horizontal grid spacing when reproducing the NLLJ over U.S. Great Plains. However, results for the 24 October 2012 case (see Table 4 and Fig. 6) did not necessarily agree with the above conclusion. In this case, the 1-km 1024-point, 4-km 256-point, and 1-km 256-point simulations all performed similarly. For both cases, the 1-km 256-point nested domain performed similarly to the 1-km 1024-point and 4-km 256-point domains. In summary, any

possible advantages of using 1-km spacing on the 256-point domain without nesting was likely washed out by the relatively coarse representation of the NLLJ from the 12-km NAM lateral boundary conditions. In this study, using 1-km spacing in the NLLJ only (i.e., the 1-km 256-point simulations) and using 1-km spacing in the NLLJ and in the surrounding region (i.e., the 1-km 1024-point simulations) to represent additional features that could influence the jet had similarly small impacts on the considered performance metrics. The differences between the simulations using 1-km and 4-km horizontal grid spacing did not justify the 16-fold increase in computational burden associated with the finer-resolution simulation, and did not motivate further study of larger nested domains. While we could envision additional tests of grid extent and nesting applications, we felt that investigating vertical grid spacing would be generally more important for our study's objectives.

Comparison with observations clearly indicates that the model underestimated the magnitude of the NLLJ using the default vertical grid with any horizontal spacing. This underestimation was confirmed by the negative bias_N values presented in Tables 3 and 4. In both cases, the underestimation was particularly large within the core of the jet. The model produced a NLLJ that is too diffuse in the vertical direction compared to the observed jet. Tests showed that the increase in computational costs associated with the 1-km spacing apparently offers no direct benefit over the 4-km configurations. While specific details of model solutions differed among configurations with differing PBL schemes, the results were overall similar (see Tables 3 and 4 and Fig. 6). This finding is consistent with results from previous studies, especially those conducted over the terrain representative of the Great Plains. Gibbs et al. (2011) found that model runs employing 4-km horizontal grid spacing produced near-surface turbulent flow parameters that were as accurate, if not more so, than those conducted using a grid with 1-km spacing. Kain et al. (2008), Schwartz et al. (2009), and Clark et al. (2012) all found that the finer details in convection forecasts produced by a WRF Model run with 2-km horizontal grid spacing did not add enough value over the forecasts provided by the 4-km configuration to justify the increased computational burden. Similar results were found in evaluations using object-based evaluation methods (Johnson et al. 2013).

b. *Vertical grid spacing*

Given the strong vertical gradients shown in LABLE observations and their absence in simulations shown in Fig. 8, it is reasonable to hypothesize that adding additional vertical levels might improve the modeled NLLJ. This hypothesis was tested by comparing results obtained with the WRF Model default vertical grid to those obtained using quasi-constant vertical spacing of 40 m. The 40-m constantly spaced vertical grid was constructed using the hypsometric equation. Initially, the 40-m grid in our case had 101 vertical levels with the first model level at 20 m and the model top near 4 km. Employing such a grid constrained the vertical extent of the model domain, but the NLLJ exists within this portion of the atmosphere, and these cases did not include persistent strong vertical motions, so the initial and boundary conditions were relied upon to represent the larger scale conditions. These simulations will be called low model-top simulations. For all grids, a Rayleigh damping layer was applied in the top 600 m of the domain. The construction of the quasi-constant vertical grid and the damping option are similar to the approach described in Mirocha et al. (2014). Levels for both vertical grids in the lowest 2 km of the model domain are illustrated in Fig. 5. We initially focus on the runs with the YSU PBL parameterization scheme (Fig. 9).

As mentioned in section 4a, a consistent underestimation of NLLJ magnitude was observed when applying the default vertically stretched grid for both considered nights. This underestimation was most prevalent in the core of the jet. Simulations with the default WRF grid can therefore miss important features of the temporal evolution and structure of the NLLJ. For both cases, the noted underestimation in YSU4kmDef was still present in the YSU4km40m simulations with a low model top (Figs. 9c, g), but the structure and temporal evolution of the jet were both improved. However, the YSU4km40m runs with a low model top overpredicted wind speed below the height of the observed NLLJ maximum, which is likely caused by spatial errors in the predicted and observed NLLJ height and depth. The appearance of the wind field was notably improved by better representation of the jet core and changes to its structure in time. For 9 October 2012, the WRF Model predicted lower values for the height of the NLLJ maximum early on and after 0700 UTC. Additionally, the model struggled to reproduce the narrow jet core, but simulations with finer vertical spacing and a low model top still showed some improvement in this regard. For 24 October 2012, the depth of the NLLJ was well predicted, but the NLLJ maximum was again positioned too high early on and after 0900 UTC, when the runs with finer vertical spacing and a low model top failed to reproduce the lifting of the jet. In both cases, the finer vertical grid spacing allowed for the temporal evolution of the height and magnitude of the NLLJ to be more representative of the observed evolution. Indeed, the overall performance for the strongly turbulent NLLJ on 24 October 2012 was

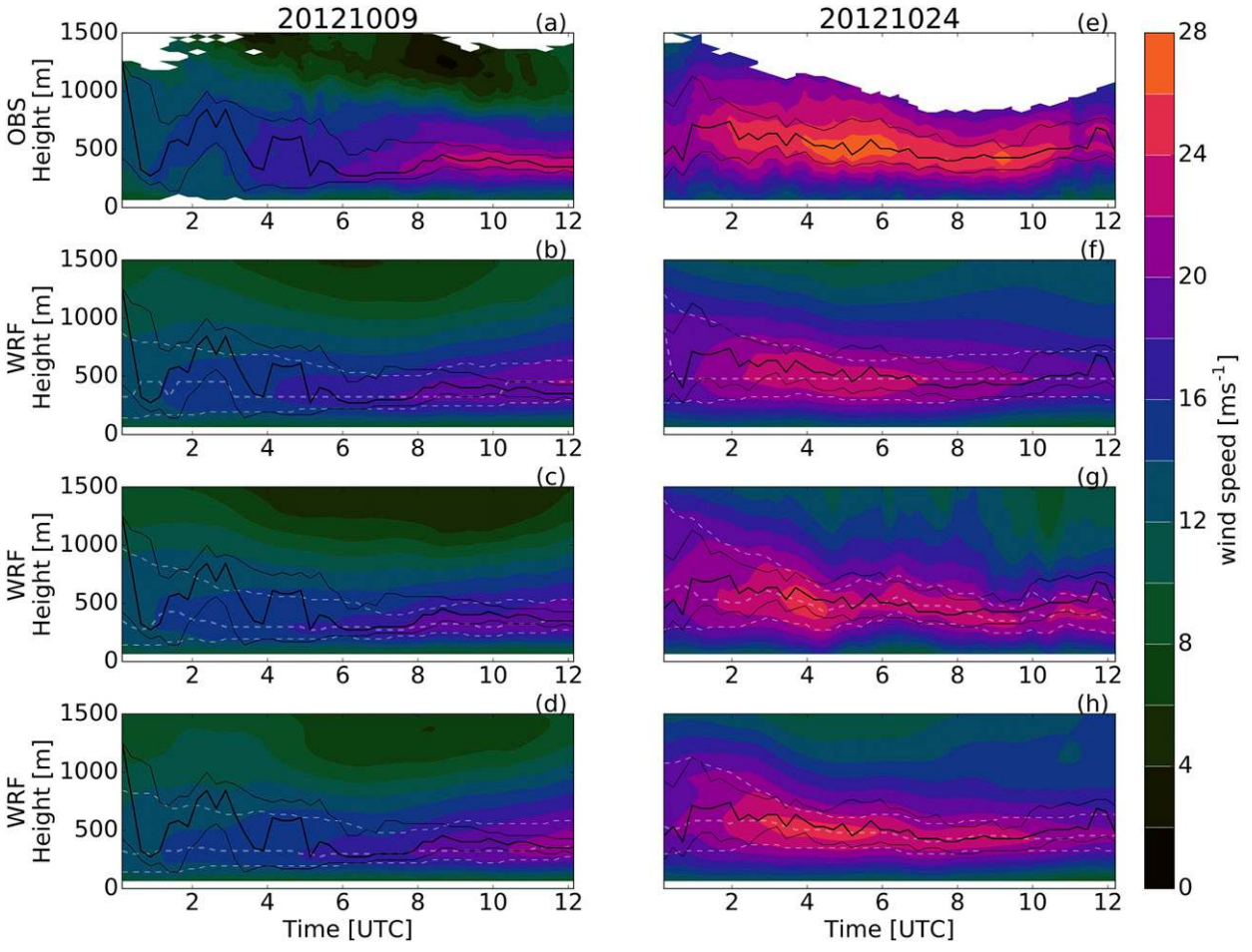

FIG. 9. As in Fig. 8, but WRF simulations are shown with (b),(f) the default stretched vertical grid, (c),(g) 40-m vertical spacing with a low model top, and (d),(h) 40-m vertical spacing with a high model top. All simulations shown use 4-km horizontal spacing and the YSU PBL scheme.

generally improved with bias_N and MAE_N values decreasing for the YSU4km40m runs with a low model top as compared with the YSU4kmDef run (see Table 4 and Fig. 6). CRMSE_N increased slightly. The same cannot be said for the weakly turbulent case on 9 October 2012, with metrics either increasing or not changing for the YSU4km40m runs with a low model top as compared with the YSU4kmDef run. However, MM values including both cases for the YSU4km40m runs with a low model top were similar to those of the YSU4kmDef runs. The 40-m spaced vertical grid offered some improvement of jet prediction over the default stretched grid based on both the visual comparison of the simulated and observed NLLJ wind fields and the more objective performance metric evaluation.

Mesoscale simulations of the NLLJ with a typical default vertical grid configuration showed weak sensitivity to the choice of PBL scheme. Since it has been shown that a 4-km horizontally spaced grid with finer constant vertical spacing offers improvement to the simulated NLLJ, sensitivity to the three PBL parameterization schemes discussed in section 3 was revisited with this grid configuration (i.e., 4-km horizontal spacing and 40-m vertical spacing with a low model top). Simulated wind speeds are shown in Fig. 10, and simulated potential temperature data are shown in Fig. 11.

For both dates, it is immediately clear that the MYNN4km40m runs with a low model top (Figs. 10c,g) did not develop a valid NLLJ, with the modeled jet being diffused vertically through the entire boundary layer. Corresponding potential temperature fields from the MYNN4km40m simulations with a low model top (Figs. 11c,d) were similarly mixed throughout the boundary layer. We ran the MYNN scheme within three previous versions of the WRF Model to confirm this behavior for grid configurations with a low model top. In each version and for both cases, the MYNN scheme

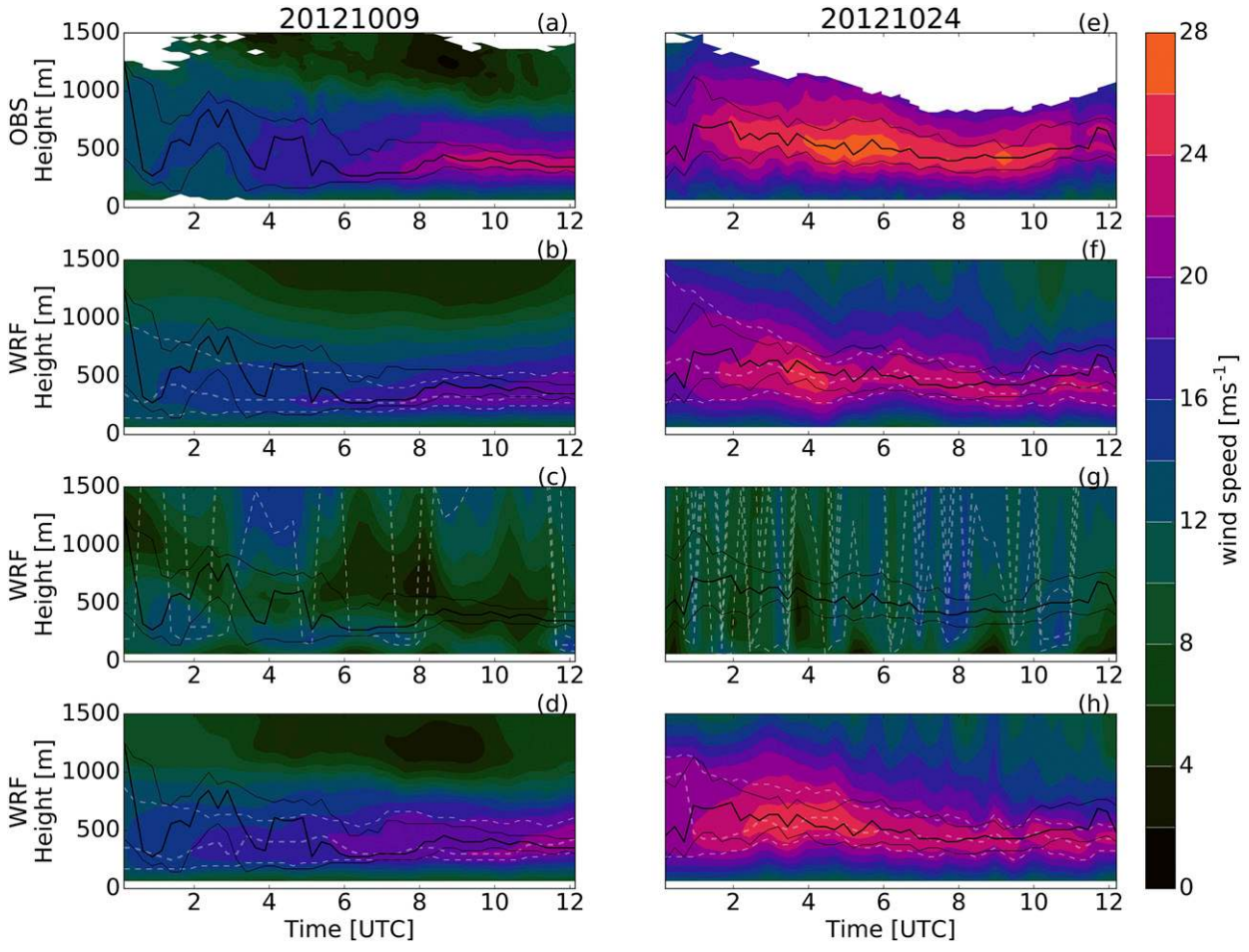

FIG. 10. As in Fig. 8, but WRF simulations are shown with the (b),(f) YSU, (c),(g) MYNN, and (d),(h) QNSE PBL schemes. All simulations shown use 4-km horizontal spacing and 40-m vertical spacing with a low model top.

produced results similar to those shown in Figs. 10c,g and 11c,g for the 40-m vertically spaced grid with a low model top (see Tables 3 and 4 and Fig. 6). Similar behavior of the MYNN scheme was noted in Olson and Brown (2012) and Jahn et al. (2014). In the latter paper, the authors suggested that the LES-derived closure constants in the scheme may be unsuitable for some conditions. We hypothesize that the MYNN scheme may be especially sensitive to the grid configuration when determining the mixing lengths prescribed in the scheme.

The unrealistic behavior of the MYNN4km40m simulations with a low model top motivated us to test the importance of the model-top height to the resulting wind field. A new vertical grid was constructed by blending the 40-m grid with 101 vertical levels with the default grid above 4 km. The resulting grid consisted of 126 vertical levels and brought the model top to about 16 km to match the height of the default grid. These simulations will be called high model-top simulations. The wind fields for the YSU4km40m configuration with a high model top are shown in Figs. 9d and 9h. In comparison with the YSU4km40m configuration with a low model top (Figs. 9c,g), the increased model-top height had a weak impact on the wind field. Figures 12 and 13 show the wind speed and potential temperature fields from simulations using the higher model top and all three PBL schemes. Impacts on simulations using the MYNN scheme were large. The performance metrics shown in Fig. 6 confirm that the impact of the higher model top is small for runs using YSU or QNSE but large for simulations using the MYNN scheme. The improvement of simulations using the MYNN scheme when increasing the model top suggests a sensitivity in the scheme that is not present in the other evaluated schemes. The sensitivity of the MYNN scheme to model-top height cannot be readily explained in physical terms at this point.

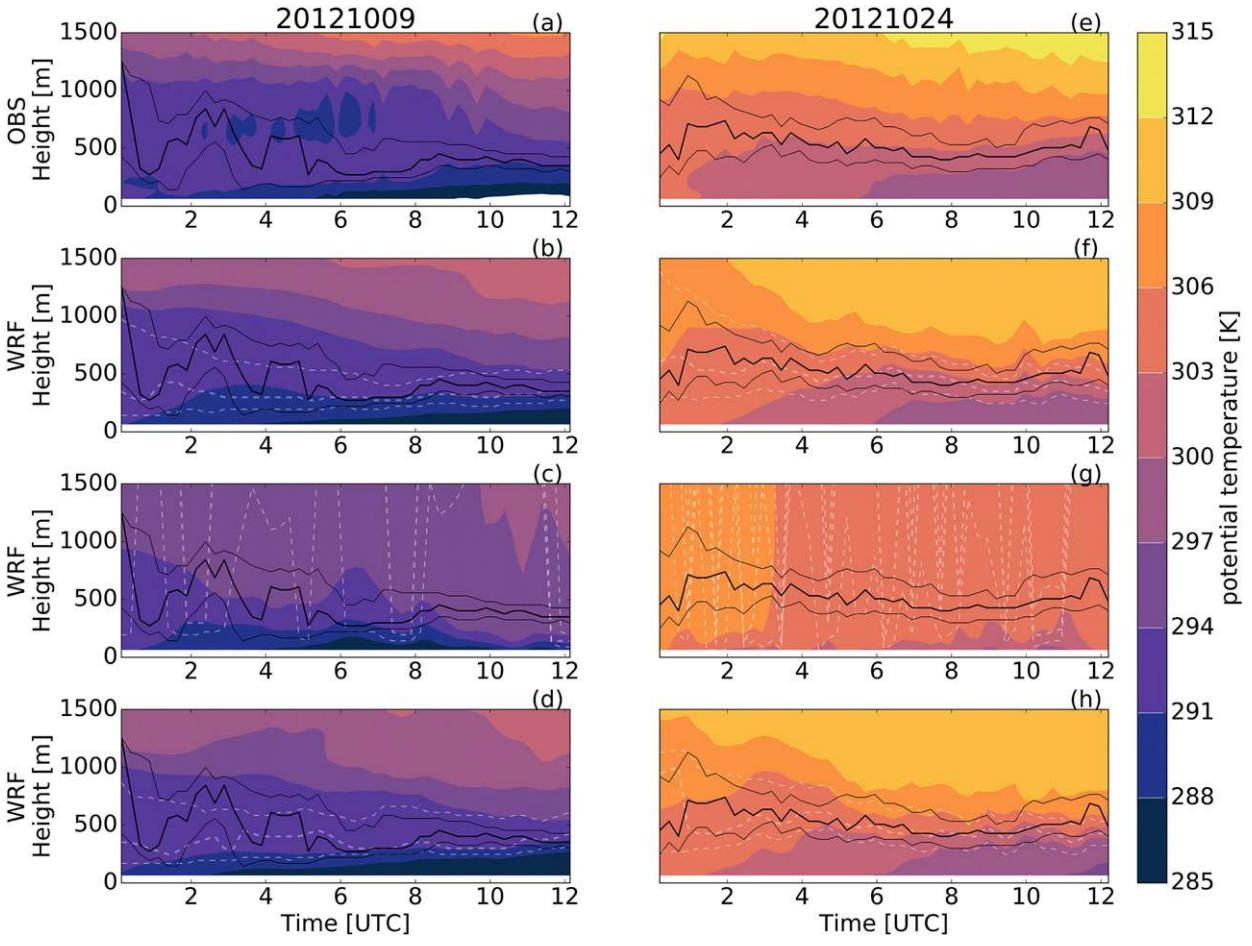

FIG. 11. As in Fig. 10, but showing observed and simulated potential temperature.

The previous analysis showed that the YSU4km40m configurations with either model top (Figs. 10b,f and 12b,f) yielded a NLLJ with realistic temporal evolution and vertical structure for both cases, despite an apparent underestimation of the NLLJ magnitude. The MYNN4km40m simulations produced nonphysical results with a low model top (Figs. 10c,g), but performed comparably to simulations using other schemes when used with a high model top (Figs. 12c,g). The QNSE4km40m simulations (Figs. 10d,h and 12d,h) performed comparably to YSU4km40m simulations with either model top, but offered improvements in the appearance and behavior of the jet, especially in the 9 October 2012 case. Specifically, the QNSE4km40m configurations more closely predicted the magnitude and structure of the NLLJ. Areas where the wind speed was under or overestimated were limited to the periods of jet development and decay, leaving a more accurate representation of the jet core as compared to other configurations. The temporal evolution of the jet was generally consistent with observations, as was its vertical structure. This was true even for the period before the NLLJ fully developed prior to 0600 UTC 9 October 2012. The metrics summarized in Fig. 6 are consistent with this interpretation, where QNSE4km40m runs performed slightly better than YSU4km40m runs and the MYNN4km40m run using a high model top. In addition, the QNSE4km40m configurations offered subtle improvements to the predicted evolution of the potential temperature fields (Figs. 11d,e and 13d,e).

Similar conclusions were drawn by Vanderwende et al. (2015), in which the QNSE scheme performed the best among the considered PBL schemes, especially in moderate-to-strong NLLJ cases. Mirocha et al. (2016) also noted the superior performance of the QNSE scheme in predicting boundary layer wind at night as compared with MYNN and YSU. They specifically noted how the QNSE performance demonstrates the benefits of improving model physics for particular environments for which more traditional approaches may

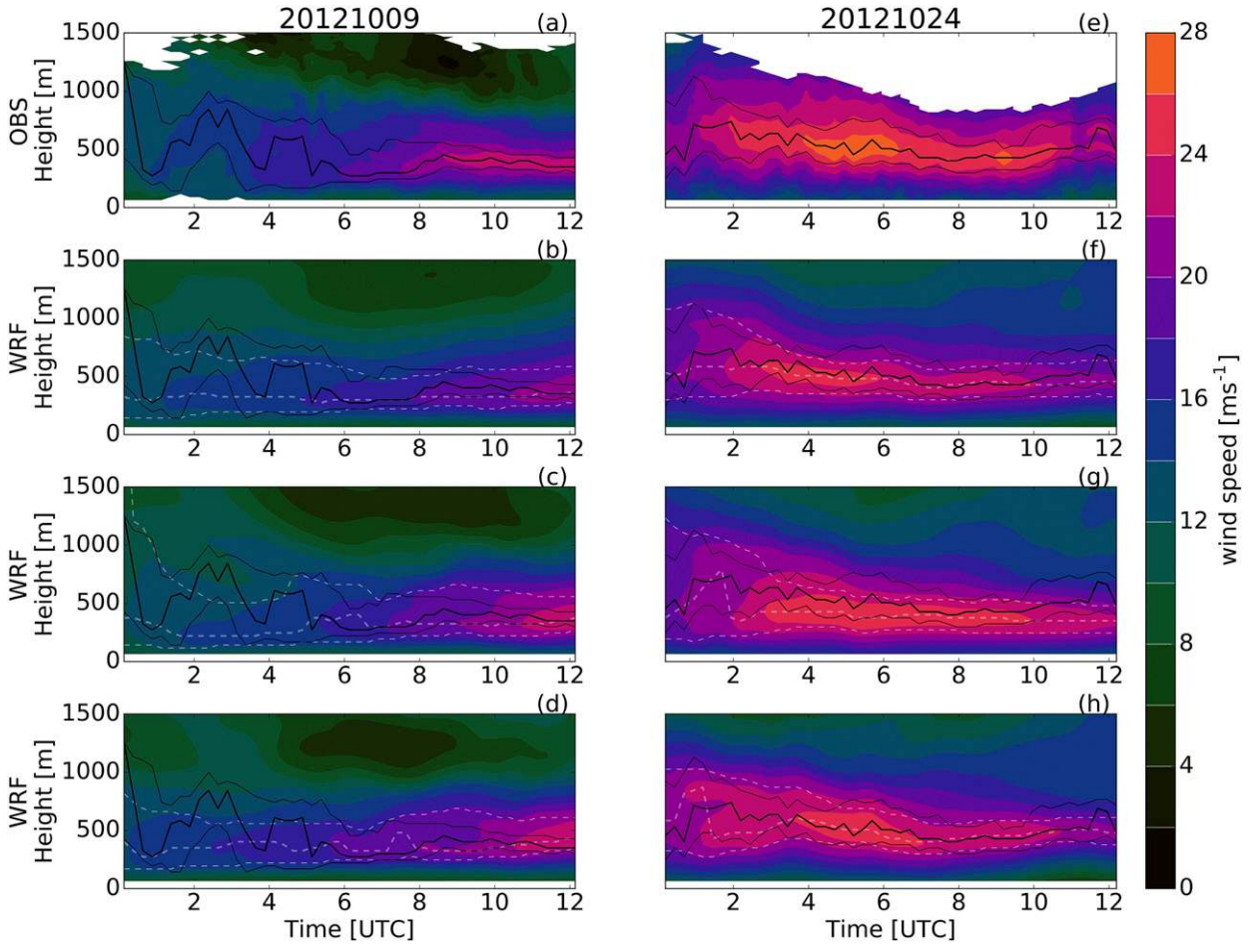

FIG. 12. As in Fig. 8, but WRF simulations are shown with the (b),(f) YSU, (c),(g) MYNN PBL, and (d),(h) QNSE PBL schemes. All simulations shown use 4-km horizontal spacing and 40-m vertical spacing with a high model top.

be inadequate. The SBL is a prime example of such an environment. The present study confirms these results using different grid configurations and for a different set of NLLJ cases than previously considered.

c. *Expanded case dataset*

The same normalized metrics discussed in section 3 were computed for the 10 LABLE cases listed in Table 1. The metrics are visualized in Fig. 14. The conducted analysis included tests of all three PBL schemes using the 4-km 40-m (high model top) and 4-km default grid configurations. Time series of NLLJ maximum wind speed are shown in Fig. 7. The MYNN PBL scheme often struggled with strongly turbulent cases when applied with the default vertically stretched grid. This result was not apparent in either of the cases discussed in detail above. As mentioned in section 4b, we hypothesize that the decline in performance may result from sensitivities related to the mixing length specification in the MYNN scheme. Further investigation is needed to understand the origins of this issue in the scheme, but is beyond the scope of the presented work. Otherwise, the varying PBL schemes and grid configurations produced different outcomes for the 10 cases. There was no consistent trend for improved performance metrics for finer grid spacing or PBL scheme. The bias_N, MAE_N, and CRMSE_N averages over all 10 cases are shown as the last set of bars in each respective panel of Fig. 14. MYNN runs using the default grid produced the highest bias_N values. MYNN runs with the 40-m vertical grid and YSU and QNSE runs on both vertical grids fared much better than the MYNN runs with the default grid, and the 40-m vertical grid offered minor improvement to the bias_N for YSU and QNSE runs. A similar result is seen in the MAE_N values averaged over all 10 cases. MYNN runs with the default grid produced the highest MAE_N, but other configurations all perform similarly with a slight favoring of the 40-m grid. CRMSE_N shows that all

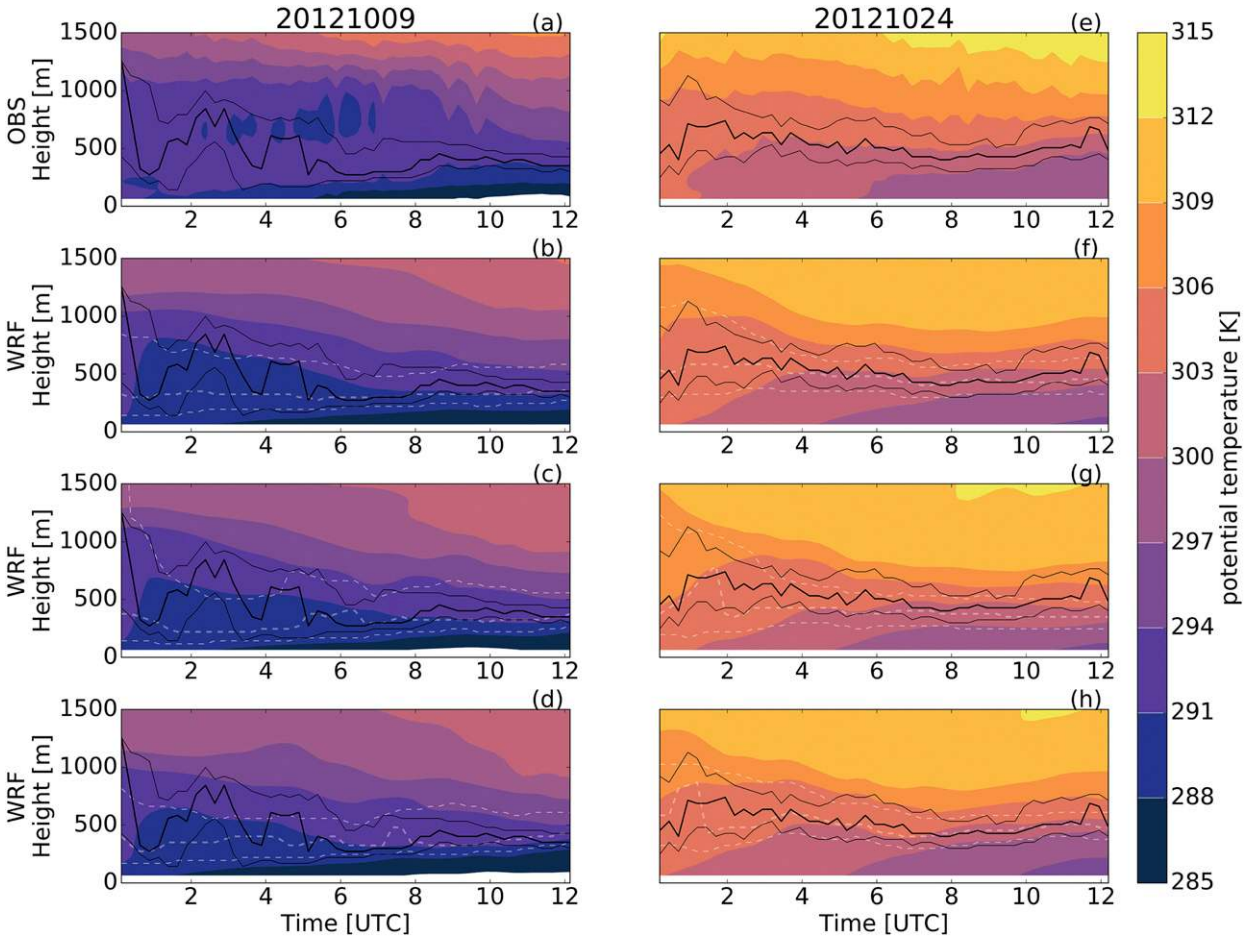

FIG. 13. As in Fig. 12, but showing observed and simulated potential temperature.

configurations performed similarly, but again there was a slight favor toward the 40-m vertical grid.

The final panel of Fig. 14 shows the MM for all cases and for strongly and weakly turbulent cases separately. For all cases, MYNN runs with the default vertical grid performed the worst in terms of MM. MYNN runs with the 40-m grid, YSU runs, and QNSE runs performed similarly with QNSE showing a slight advantage with lowest MM values. For YSU and QNSE, there was very little difference in the runs using the default and vertical grids. Runs using the MYNN scheme benefited substantially from the increased number of vertical levels when the model top was high enough. Strongly turbulent cases followed the same pattern as all cases with similar performance for all configurations except for the MYNN runs with the default vertical grid. Weakly turbulent cases were much better captured in the MYNN runs using the default vertical grid, but these cases were best represented by the QNSE runs using the 40-m grid. In a general sense, runs using the MYNN scheme on the default vertical grid struggled the most to adequately represent the NLLJ. Runs using the MYNN scheme on the 40-m grid and YSU and QNSE on both vertical grids were all quite similar with slight differences between the default and 40-m vertical grid. The MM values are shown for both wind direction and potential temperature in Fig. 15. Similar to the data shown in Fig. 14, The MM values for wind direction and potential temperature indicate that, except for runs using MYNN on the default vertical grid, all configurations were comparable.

The average bias, MAE, and CRMSE were computed for wind speed across all 10 cases. Using these averages, the 95% confidence intervals for each performance metric were constructed using an empirical bootstrap technique with 10 000 replacement iterations, similar to methods described in Hamill (1999). The resulting averages and confidence intervals are shown in Fig. 16. Again, except for runs using MYNN on the default vertical grid, all configurations were fairly comparable. While differences are small the confidence intervals do show that, for all metrics except CRMSE, all configurations performed better on the 40-m vertical grid than the default vertical grid.

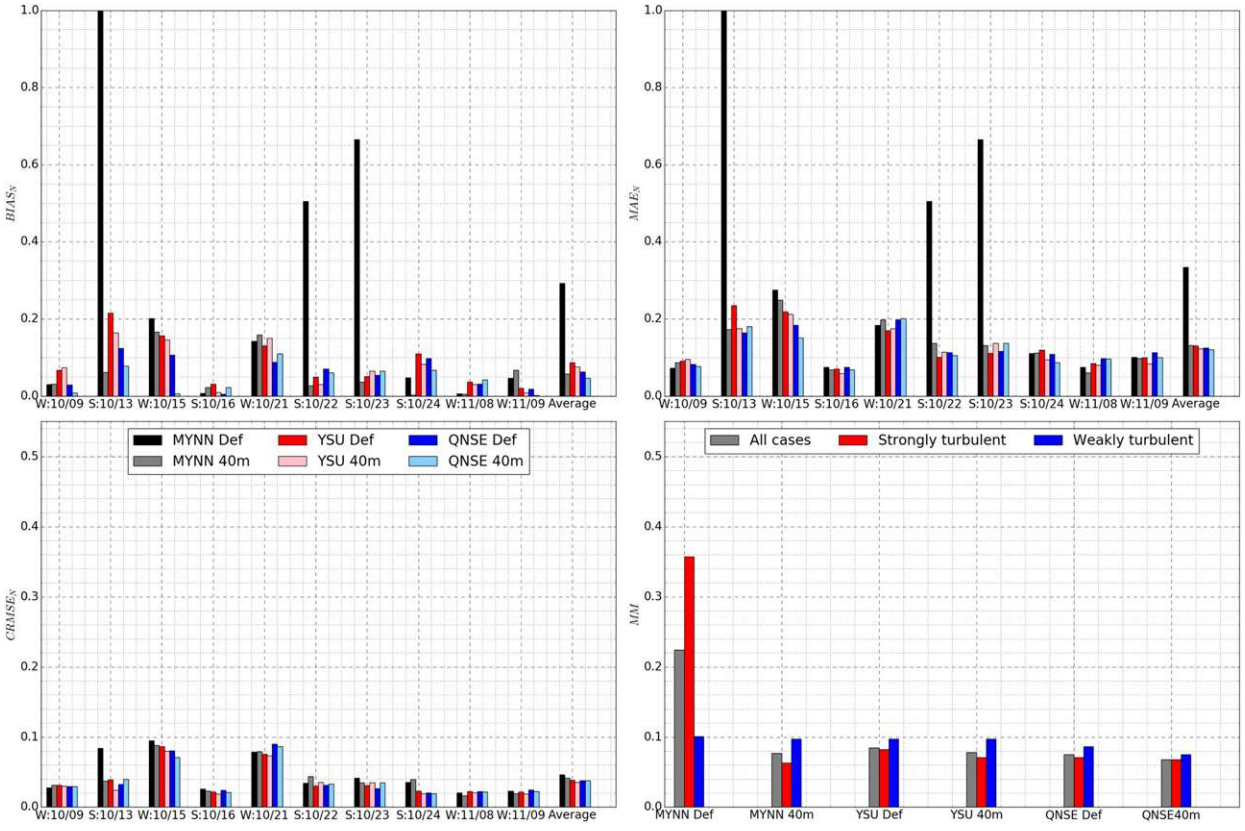

FIG. 14. Bar plots show the values of (top left) bias_N (displayed as an absolute value), (top right) MAE_N, and (bottom left) CRMSE_N for 10 LABLE cases. The case dates and classifications (W for weak; S for strong) are labeled on the x axis. Also shown is (bottom right) MM for both cases for each classification.

It is important to note that performance metrics can possibly degrade if simulations manage to predict a stronger NLLJ but fail to correctly predict the height and depth of the NLLJ core. Similar verification issues have been discussed previously in the literature highlighting the utility of object- or feature-based verification methods (e.g., Johnson et al. 2013; Mittermaier 2008). Bearing this in mind and reflecting on the visual analysis of the wind fields shown such as those in Fig. 9, NLLJ development and persistence was improved with finer vertical spacing. Evaluation of the potential temperature fields shown in Figs. 4 and 13 further supported the use of finer vertical grid spacing. While simulations using the MYNN (with 40-m spacing on a grid with a high model top), QNSE, and YSU schemes performed similarly, the QNSE scheme was ultimately chosen for our future studies because of slightly better performance in these cases and its focus on SBL physics. QNSE uses scale elimination to approximate exchange by interacting turbulent eddies, which can perhaps better capture the processes associated with the sharp vertical gradients of meteorological fields within NLLJs. These test results point to the configuration with a 4-km horizontally spaced and 40-m vertically spaced grid and the QNSE scheme as the most optimal considered configuration of the WRF Model for research simulations of the NLLJ over the Great Plains.

5. Conclusions

Numerical experiments were run for 10 NLLJs observed during LABLE field campaign. Two of these cases were chosen for detailed evaluation: a weakly turbulent NLLJ on 9 October 2012 and a strongly turbulent NLLJ on 24 October 2012. Overall results concerning grid spacing and PBL parameterization were consistent between the two cases even though the NLLJ developed quite differently in each case (see e.g., Fig. 2). The NLLJs simulated with the WRF Model were found to have varying degrees of sensitivity to horizontal grid spacing, vertical grid spacing, and choice of PBL parameterization scheme.

Counter to the seemingly natural assumption that reducing grid spacing should improve predictions of

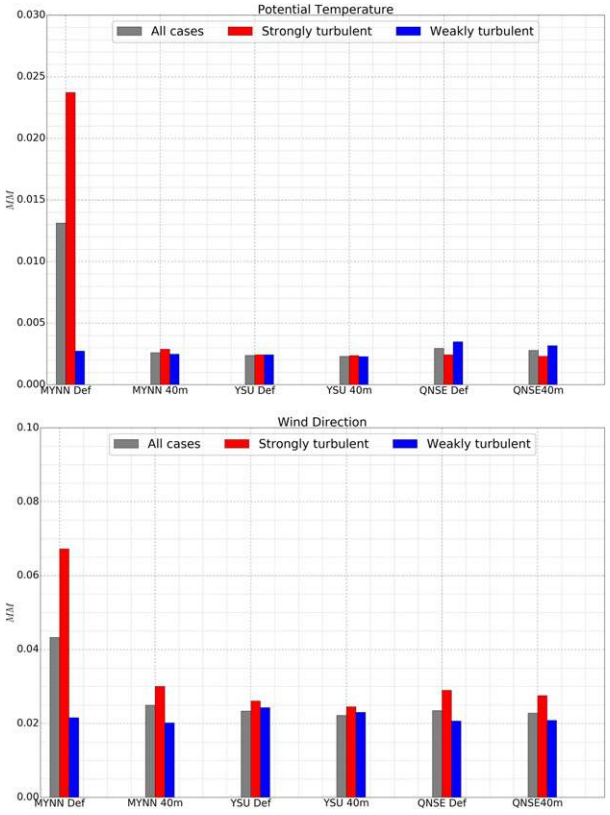

FIG. 15. The MM for all cases and for each classification for (top) potential temperature and (bottom) wind direction.

atmospheric flow variables, findings of our study suggested that using 4-km horizontal spacing is no worse than using finer horizontal spacing (i.e., 1 or 2 km) for representing the NLLJ. In fact, our results showed that finer horizontal spacing could result in degraded wind field simulations. Accordingly, there is little justification for the increased computational burden associated with the use of a 1-km grid. These results are consistent with those reported in the literature and indicate that 4-km horizontal grid spacing is the optimal choice for WRF simulation of the considered NLLJs and associated nocturnal SBLs.

Vertical grid spacing is another key factor in simulating the NLLJ using the WRF Model. While the default stretched vertical grid configuration of the WRF Model performed generally satisfactorily, the use of a vertical grid with quasi-constant and finer spacing proved to be a worthwhile investment of computational resources. The simulations with a finer vertical grid provided improvements to the morphology of the NLLJ, especially to the representation of the temporal evolution and vertical structure of the jet. The MYNN PBL scheme severely underestimated the magnitude of the NLLJ and overmixed the boundary layer flow when applied on a nonstretched vertical grid with a low model top. Similar trends were also observed with the default vertically stretched grid (in some strongly turbulent cases), confirming previously reported behavior (e.g., Olson and Brown 2012; Jahn et al. 2014). This suggests that some grid configurations are not optimal with the MYNN scheme. When applied with a higher model top, the MYNN scheme produced acceptable results. Such strong sensitivity to model-top height was not noted for the YSU or QNSE schemes. The YSU scheme produced a realistic jet in terms of its structure and evolution, but underestimated the jet magnitude. On the other hand, the QNSE scheme more accurately depicted the magnitude, temporal evolution, and structure of the NLLJ, with improvements also noted in the simulated potential temperature fields. The 40-m vertical grid was found to be an optimal alternative to the default stretched grid when simulating the NLLJ in boundary layer studies.

Results presented for the full 10-case dataset in Fig. 14 showed that runs using any scheme with a high model top were fairly similar in terms of the performance metrics. The runs using the QNSE scheme showed a slight advantage over others. While some cases with default vertical spacing scored similarly well or even better than the finer vertical grids in various metrics, these diffuse jets have limited value for detailed NLLJ and SBL analyses. Combining the performance metric evaluation results with the visual comparison of the simulated and observed wind fields made clear the utility of the constantly spaced fine vertical grid over the vertically stretched grid. While all schemes performed similarly in the 4-km horizontally spaced and 40-m vertically spaced (with high model top) configuration, the relevant depiction of stable boundary layer physics in the QNSE makes it particularly attractive for NLLJ modeling.

From the analyses that we presented, we found that the optimal grid configuration for simulation of the Great Plains NLLJ using the WRF Model was a combination of 4-km horizontal spacing and constant 40-m vertical spacing. In conjunction with this grid, the QNSE scheme was identified as the optimal PBL parameterization. WRF Model outputs using this configuration offer detailed information about the structure of the boundary layer and the NLLJ. These simulations also shed light on NLLJ temporal evolution, suggesting that the spatial evolution of the NLLJ is also well captured. The identified configuration will be used in future studies of PECAN cases in which the three-dimensional structure and evolution of the NLLJ over the Great Plains can be evaluated against high-resolution observations at several locations.

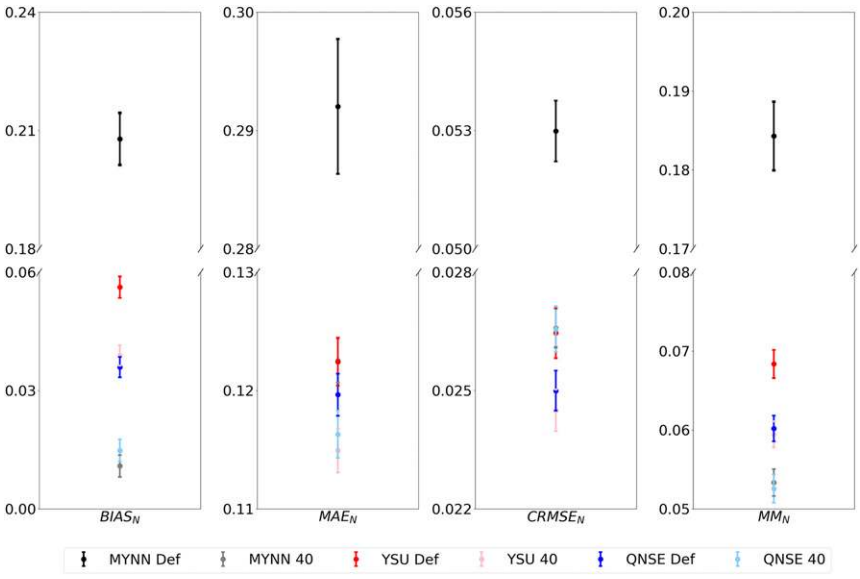

FIG. 16. Averages for each configuration over all 10 cases are shown with 95% confidence intervals for each considered performance metric. Note that the *y* axes vary across panels and contain breaks.

Acknowledgments. This research was completed with support by the National Science Foundation under Grant AGS-1359698. The first author was also supported in part by the Lockheed Martin Corporation–American Meteorological Society Graduate Fellowship. The authors acknowledge the participants in the LABLE field campaign for the observation data used in this work, with specific acknowledgement of Timothy Bonin for his detailed explanation and assistance with LABLE data. Comments from anonymous peer reviewers greatly improved this work. The authors additionally acknowledge high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) that was provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

REFERENCES

Ardanuy, P., 1979: On the observed diurnal oscillation of the Somali jet. *Mon. Wea. Rev.*, **107**, 1694–1700, https://doi.org/10.1175/1520-0493(1979)107<1694:OTODOO>2.0.CO;2.

Astling, E. G., J. Paegle, E. Miller, and C. J. O'Brien, 1985: Boundary layer control of nocturnal convection associated with a synoptic scale system. *Mon. Wea. Rev.*, **113**, 540–552, https://doi.org/10.1175/1520-0493(1985)113<0540:BLCONC>2.0.CO;2.

Baines, P., 1980: The dynamics of the Southerly Buster. *Aust. Meteor. Mag.*, **28**, 175–200.

Blackadar, A. K., 1957: Boundary layer wind maxima and their significance for the growth of nocturnal inversions. *Bull. Amer. Meteor. Soc.*, **38**, 283–290, https://doi.org/10.1175/1520-0477-38.5.283.

Bonin, T. A., 2015: Nocturnal boundary layer and low-level jet characteristics under different turbulent regimes. Ph.D. thesis, The University of Oklahoma, 190 pp., https://shareok.org/handle/11244/14623.

——, W. G. Blumberg, P. M. Klein, and P. B. Chilson, 2015: Thermodynamic and turbulence characteristics of the southern Great Plains nocturnal boundary layer under differing turbulent regimes. *Bound.-Layer Meteor.*, **157**, 401–420, https://doi.org/10.1007/s10546-015-0072-2.

Bonner, W. D., 1966: Case study of thunderstorm activity in relation to the low-level jet. *Mon. Wea. Rev.*, **94**, 167–178, https://doi.org/10.1175/1520-0493(1966)094<0167:CSOTAI>2.3.CO;2.

——, 1968: Climatology of the low level jet. *Mon. Wea. Rev.*, **96**, 833–850, https://doi.org/10.1175/1520-0493(1968)096<0833:COTLLJ>2.0.CO;2.

Brook, R. R., 1985: The Koorin nocturnal low-level jet. *Bound.-Layer Meteor.*, **32**, 133–154, https://doi.org/10.1007/BF00120932.

Chang, J., and S. Hanna, 2004: Air quality model performance evaluation. *Meteor. Atmos. Phys.*, **87**, 167–196, https://doi.org/10.1007/s00703-003-0070-7.

Clark, A. J., and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed experimental forecast program spring experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, https://doi.org/10.1175/BAMS-D-11-00040.1.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Fernando, H. J. S., and J. C. Weil, 2010: Whither the stable boundary layer? *Bull. Amer. Meteor. Soc.*, **91**, 1475–1484, https://doi.org/10.1175/2010BAMS2770.1.

Gebauer, J. G., A. Shapiro, E. Fedorovich, and P. K. Klein, 2018: Convection initiation caused by heterogeneous low-level jets over the Great Plains. *Mon. Wea. Rev.*, **146**, 2615–2637, https://doi.org/10.1175/MWR-D-18-0002.1.

Geerts, B., and Coauthors, 2017: The 2015 Plains Elevated Convection at Night field project. *Bull. Amer. Meteor. Soc.*, **98**, 767–786, https://doi.org/10.1175/BAMS-D-15-00257.1.

Gibbs, J. A., E. Fedorovich, and A. M. Van Eijk, 2011: Evaluating Weather Research and Forecasting (WRF) Model predictions of turbulent flow parameters in a dry convective boundary layer. *J. Appl. Meteor. Climatol.*, **50**, 2429–2444, https://doi.org/10.1175/2011JAMC2661.1.

Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, https://doi.org/10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2.

Holton, J. R., 1967: The diurnal boundary layer wind oscillation above sloping terrain. *Tellus*, **19**, 199–205, https://doi.org/10.1111/j.2153-3490.1967.tb01473.x.

Holtslag, A. A. M., and Coauthors, 2013: Stable atmospheric boundary layers and diurnal cycles: Challenges for weather and climate models. *Bull. Amer. Meteor. Soc.*, **94**, 1691–1706, https://doi.org/10.1175/BAMS-D-11-00187.1.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120, https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Jahn, D. E., W. Gallus, and E. S. Takle, 2014: Evaluation of the MYNN PBL scheme closure constants for low-level jets in a stable boundary layer. *Proc. 21st Symp. on Boundary Layers and Turbulence,* Leeds, United Kingdom, Amer. Meteor. Soc., 16B.5, https://ams.confex.com/ams/21BLT/webprogram/Paper247664.html.

Johnson, A., X. Wang, F. Kong, and M. Xue, 2013: Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Mon. Wea. Rev.*, **141**, 3413–3425, https://doi.org/10.1175/MWR-D-13-00027.1.

Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, https://doi.org/10.1175/WAF2007106.1.

Klein, P., and Coauthors, 2015: LABLE: A multi-institutional, student-led, atmospheric boundary layer experiment. *Bull. Amer. Meteor. Soc.*, **96**, 1743–1764, https://doi.org/10.1175/BAMS-D-13-00267.1.

——, X.-M. Hu, A. Shapiro, and M. Xue, 2016: Linkages between boundary-layer structure and the development of nocturnal low-level jets in central Oklahoma. *Bound.-Layer Meteor.*, **158**, 383–408, https://doi.org/10.1007/s10546-015-0097-6.

LeMone, M. A., M. Tewari, F. Chen, and J. Dudhia, 2013: Objectively determined fair-weather CBL depths in the ARW-WRF Model and their comparison to CASES-97 observations. *Mon. Wea. Rev.*, **141**, 30–54, https://doi.org/10.1175/MWR-D-12-00106.1.

Maddox, R. A., 1983: Large-scale meteorological conditions associated with midlatitude, mesoscale convective complexes. *Mon. Wea. Rev.*, **111**, 1475–1493, https://doi.org/10.1175/1520-0493(1983)111<1475:LSMCAW>2.0.CO;2.

Mahrt, L., 1998: Stratified atmospheric boundary layers and breakdown of models. *Theor. Comput. Fluid Dyn.*, **11**, 263–279, https://doi.org/10.1007/s001620050093.

——, 1999: Stratified atmospheric boundary layers. *Bound.-Layer Meteor.*, **90**, 375–396, https://doi.org/10.1023/A:1001765727956.

Markowski, P., and Y. Richardson, 2011: *Mesoscale Meteorology in Midlatitudes*. John Wiley and Sons, 430 pp.

Mirocha, J., B. Kosovic, and G. Kirkil, 2014: Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **142**, 806–831, https://doi.org/10.1175/MWR-D-13-00064.1.

——, M. D. Simpson, J. D. Fast, L. K. Berg, and R. Baskett, 2016: Investigation of boundary-layer wind predictions during nocturnal low-level jet events using the Weather Research and Forecasting Model. *Wind Energy*, **19**, 739–762, https://doi.org/10.1002/we.1862.

Mittermaier, M. P., 2008: The potential impact of using persistence as a reference forecast on perceived forecast skill. *Wea. Forecasting*, **23**, 1022–1031, https://doi.org/10.1175/2008WAF2007037.1.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-*k* model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, https://doi.org/10.1029/97JD00237.

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

National Research Council, 1983: *Low-Altitude Wind Shear and Its Hazard to Aviation.* National Academies Press, 112 pp.

Olson, J., and J. M. Brown, 2012: Modifications in the MYNN PBL/surface layer scheme for WRF-ARW. *Proc. WRF Users Workshop 2012*, Boulder, CO, NCAR, http://www2.mmm.ucar.edu/wrf/users/workshops/WS2012/ppts/3.3.pdf.

Pan, Z., M. Segal, and R. W. Arritt, 2004: Role of topography in forcing low-level jets in the central United States during the 1993 flood-altered terrain simulations. *Mon. Wea. Rev.*, **132**, 396–403, https://doi.org/10.1175/1520-0493(2004)132<0396:ROTIFL>2.0.CO;2.

Pitchford, K. L., and J. London, 1962: The low-level jet as related to nocturnal thunderstorms over midwest United States. *J. Appl. Meteor.*, **1**, 43–47, https://doi.org/10.1175/1520-0450(1962)001<0043:TLLJAR>2.0.CO;2.

Salmond, J. A., and I. G. McKendry, 2005: A review of turbulence in the very stable nocturnal boundary layer and its implications for air quality. *Prog. Phys. Geogr.*, **29**, 171–188, https://doi.org/10.1191/0309133305pp442ra.

Schwartz, C. S., and Coauthors, 2009: Next-day convection-allowing WRF Model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372, https://doi.org/10.1175/2009MWR2924.1.

Shapiro, A., and E. Fedorovich, 2009: Nocturnal low-level jet over a shallow slope. *Acta Geophys.*, **57**, 950–980, https://doi.org/10.2478/s11600-009-0026-5.

——, and ——, 2010: Analytical description of a nocturnal low-level jet. *Quart. J. Roy. Meteor. Soc.*, **136**, 1255–1262, https://doi.org/10.1002/qj.628.

——, ——, and S. Rahimi, 2016: A unified theory for the Great Plains nocturnal low-level jet. *J. Atmos. Sci.*, **73**, 3037–3057, https://doi.org/10.1175/JAS-D-15-0307.1.

Sisterson, D. L., and P. Frenzen, 1978: Nocturnal boundary-layer wind maxima and the problem of wind power assessment. *Environ. Sci. Technol.*, **12**, 218–221, https://doi.org/10.1021/es60138a014.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:10.5065/D68S4MVH.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Wea. Rev.*, **91**, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Smith, E. N., P. M. Klein, E. Fedorovich, J. A. Gibbs, and J. G. Gebauer, 2018: The Great Plains low-level jet during PECAN:

Observed and simulated characteristics. *Proc. Special Symp. on Plains Elevated Convection at Night (PECAN),* Austin, TX, Amer. Meteor. Soc., 2.6, https://ams.confex.com/ams/98Annual/webprogram/Paper331142.html.

Steeneveld, G.-J., 2014: Current challenges in understanding and forecasting stable boundary layers over land and ice. *Front. Environ. Sci.*, **2**, 1–6, https://doi.org/10.3389/fenvs.2014.00041.

——, T. Mauritsen, E. I. F. de Bruijn, J. Vilà-Guerau de Arellano, G. Svensson, and A. Holtslag, 2008: Evaluation of limited-area models for the representation of the diurnal cycle and contrasting nights in CASES-99. *J. Appl. Meteor. Climatol.*, **47**, 869–887, https://doi.org/10.1175/2007JAMC1702.1.

Stensrud, D. J., 1996: Importance of low-level jets to climate: A review. *J. Climate*, **9**, 1698–1711, https://doi.org/10.1175/1520-0442(1996)009<1698:IOLLJT>2.0.CO;2.

Storm, B., J. Dudhia, S. Basu, A. Swift, and I. Giammanco, 2009: Evaluation of the Weather Research and Forecasting Model on forecasting low-level jets: Implications for wind energy. *Wind Energy*, **12**, 81–90, https://doi.org/10.1002/we.288.

Stull, R. B., 2012: *An Introduction to Boundary Layer Meteorology*. Springer, 670 pp.

Sukoriansky, S., B. Galperin, and V. Perov, 2005: Application of a new spectral theory of stably stratified turbulence to the atmospheric boundary layer over sea ice. *Bound.-Layer Meteor.*, **117**, 231–257, https://doi.org/10.1007/s10546-004-6848-4.

Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF Model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 11–15, https://ams.confex.com/ams/84Annual/webprogram/Paper69061.html.

Trier, S. B., C. A. Davis, D. A. Ahijevych, M. L. Weisman, and G. H. Bryan, 2006: Mechanisms supporting long-lived episodes of propagating nocturnal convection within a 7-day WRF Model simulation. *J. Atmos. Sci.*, **63**, 2437–2461, https://doi.org/10.1175/JAS3768.1.

Turner, D., and U. Löhnert, 2014: Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor. Climatol.*, **53**, 752–771, https://doi.org/10.1175/JAMC-D-13-0126.1.

Vanderwende, B. J., J. K. Lundquist, M. E. Rhodes, E. S. Takle, and S. L. Irvin, 2015: Observing and simulating the summertime low-level jet in central Iowa. *Mon. Wea. Rev.*, **143**, 2319–2336, https://doi.org/10.1175/MWR-D-14-00325.1.

Whiteman, C. D., X. Bian, and S. Zhong, 1997: Low-level jet climatology from enhanced rawinsonde observations at a site in the southern Great Plains. *J. Appl. Meteor. Climatol.*, **36**, 1363–1376, https://doi.org/10.1175/1520-0450(1997)036<1363:LLJCFE>2.0.CO;2.

Wyngaard, J. C., 2004: Toward numerical modeling in the ''terra incognita.'' *J. Atmos. Sci.*, **61**, 1816–1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2.

Zhong, S., J. D. Fast, and X. Bian, 1996: A case study of the Great Plains low-level jet using wind profiler network data and a high-resolution mesoscale model. *Mon. Wea. Rev.*, **124**, 785–806, https://doi.org/10.1175/1520-0493(1996)124<0785:ACSOTG>2.0.CO;2.

Zhou, B., J. S. Simon, and F. K. Chow, 2014: The convective boundary layer in the terra incognita. *J. Atmos. Sci.*, **71**, 2545–2563, https://doi.org/10.1175/JAS-D-13-0356.1.