Lidar-based evaluation of HRRR performance in California's Diablo Range

Gabriel Rios [a], Robert S. Arthur [b], Sonia Wharton [b], Jerome D. Fast [c]

[a] *Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, United States*

[b] *Lawrence Livermore National Laboratory, Livermore, California, United States*

[c] *Pacific Northwest National Laboratory, Richland, Washington, United States*

Corresponding author: Gabriel Rios, gabriel.rios@princeton.edu

ABSTRACT: The performance of the NOAA High Resolution Rapid Refresh (HRRR) model for capturing low-level winds near a wind energy production site during summer 2019 is evaluated. This study catalogues the ability of HRRR to predict boundary layer dynamics relevant to wind energy interests over complex terrain, which has presented challenges for weather and energy forecasting. Performance is evaluated by comparing HRRR output to wind-profiling Doppler lidars at Lawrence Livermore National Laboratory Site 300. HRRR captured the diurnal profile of horizontal winds in the observed 150 m layer, despite strong underpredictions ($\sim 4\,\mathrm{m\,s^{-1}}$) during evening and nighttime hours. These underpredictions may be a result of local speed-up flows observed by the lidars, which were unresolved in HRRR due to their small spatial extent. HRRR bias magnitude relative to observations was found to be minimal during days with synoptic-scale troughs and strong 850 hPa geopotential gradients, while bias magnitude was maximal during days with synoptic ridging and weak 850 hPa geopotential gradients. To translate wind speed predictions to energy forecasting, generic turbine models were used to estimate power generation for turbines characteristic of the nearby Altamont Pass Wind Resource Area. Results show that HRRR-based energy estimates predicted daytime power generation adequately relative to lidar-based estimates with an 18-hour lead time (bias magnitude < 0.4 MW from 09:00-14:00 local time), but overpredicted power during the rest of the diurnal cycle (bias > 1 MW). These results demonstrate conditions under which HRRR performs well for wind energy applications in complex terrain, while highlighting biases that require further investigation to support usage of a high-resolution model for wind energy forecasts.

SIGNIFICANCE STATEMENT: Accurate prediction of surface winds is essential for forecasting atmospheric phenomena, such as boundary layer dynamics and surface-atmosphere energy exchange, to enable the prediction of operational quantities, such as wind energy output. However, prediction is complicated by complex terrain. To assess prediction accuracy, we evaluate performance of NOAA's HRRR model against wind speed data in central California using observational data from vertically-profiling lidars. This study found that winds at turbine height are accurately predicted during the daytime but overpredicted overnight. Additionally, small-scale hill speed-up events at sunset were not captured by the model, leading to consistent underprediction of near-surface winds. These results have implications for wind energy forecasting in the complex terrain of central California, and potentially other areas with similar terrain.

1. Introduction

Complex terrain (e.g., hills, mountains, valleys, ridges, etc.) presents a challenge for numerical weather prediction (NWP). The challenge is particularly significant in the atmospheric boundary layer, as parameterized surface exchange processes and spatiotemporally variable flow patterns may be difficult to capture. Moreover, the horizontal resolution of operational NWP models is often too coarse to fully resolve local-scale topographical features that influence these processes and flow patterns.

This challenge is relevant beyond the NWP community due to the prevalence of wind turbine placement in areas with complex terrain. As wind energy capacity and demand grows (Wiser et al. 2022), the forecasting of energy output becomes increasingly important for the public and private sectors. Prediction of wind energy output is useful for planning and operational purposes alike, and often requires forecasting lead times of a day or more for many stakeholders reliant on wind energy. Additionally, the high sensitivity of wind turbine production to changes in wind speed and direction make accurate and precise predictions critical for energy forecasts. However, such predictions are complicated by the highly variable nature of boundary layer dynamics over complex terrain (Olson et al. 2019).

The modeling of boundary layer flows over complex terrain for wind energy applications has been extensively studied in the literature. As far back as Sisterson and Frenzen (1978) and Liu and Yocke (1980), the importance of the numerical modeling of boundary layer flows for wind energy forecasting has been recognized by the meteorological community. Numerous studies examined the ability to forecast winds in the boundary layer over a variety of different terrains using models across scales, ranging from mesoscale models (Carvalho et al. 2012; Cheng et al. 2017; Heppelmann et al. 2017) to large-eddy simulations (Bauweraerts and Meyers 2019; Mirocha et al. 2014; Santoni et al. 2018) to wind forecasting models (e.g., statistical, deep-learning, etc.) (Kariniotakis et al. 1996; Li et al. 2022; Sideratos and Hatziargyriou 2007). Despite these and other efforts, sources of forecast accuracy are not fully understood, due in part to subgrid-scale processes and the lack of long-term observational data from the surface through the boundary layer (Pichugina et al. 2019).

A major step forward in diagnosing model errors and guiding model improvements for wind forecasting was ushered in by the Wind Forecast Improvement Project field campaigns, WFIP and WFIP2 (Olson et al. 2019; Shaw et al. 2019; Wilczak et al. 2015, 2019). WFIP presented a significant push by the public and private sectors to improve the accuracy of NWP in forecasting wind energy at short lead times (up to 24 h) through improvements to observational data assimilation and modeled boundary layer dynamics. WFIP2 marked a shift in mission goals and complexity by assessing the ability of NWP models to resolve atmospheric conditions in complex terrain. The WFIP2 campaign was based in the northwestern United States and was composed of an 18-month observational period with comprehensive profiling of surface and boundary layer processes. WFIP2 led to numerous studies on flow dynamics and their representation in NWP models specific to areas with complex terrain, such as cold-air pools, gap flows, and mountain waves (Adler et al. 2021, 2023; Arthur et al. 2022; Bianco et al. 2019; Draxl et al. 2021; Xia et al. 2021). Several of these studies focused on the forecasting of boundary layer properties directly relevant to wind energy forecasting with the intent of diagnosing operational model errors and verifying model modifications relative to observations (Banta et al. 2021; Bianco et al. 2022; Djalalova et al. 2020; Pichugina et al. 2019).

The need to resolve such phenomena has motivated the development of NWP models with increasingly higher spatial and temporal resolutions. One such model is the NOAA High-Resolution Rapid Refresh (HRRR) (Benjamin et al. 2016), which is an operational NWP model used for short-term weather forecasting over the continental United States (CONUS). Due in part to high spatial and temporal resolution relative to other operational NWP models, HRRR is widely used for short-term wind and solar energy forecasting applications (Juliano et al. 2022b; Shaw et al. 2019). A major goal of the WFIP2 project was to support development of HRRR for improved wind predictions over complex terrain (Olson et al. 2019), and various model improvements have since been included in experimental HRRR configurations (Adler et al. 2023; Banta et al. 2023; Bianco et al. 2019; Pichugina et al. 2020).

An additional phenomenon that presents modeling challenges in wind energy forecasting are speed-up flows (Banta et al. 2021; Clifton et al. 2022; Djalalova et al. 2020; Giebel and Kariniotakis 2017; Pichugina et al. 2019; Quon et al. 2019; Safaei Pirooz and Flay 2018). Speed-up flows, which are characterized as near-surface increases in wind speed over hills and ridges relative to neighboring surfaces, are typical features of flows over hills and ridges (Coppin et al. 1994; Lubitz and White 2007; Mickle et al. 1988) and are relevant for wind energy applications, such as wind farm siting (Hyvärinen et al. 2018; Tian et al. 2013, 2021) and energy output forecasting (Castellani et al. 2016; Wagenbrenner et al. 2016; Wharton et al. 2015). Because of their occurrence near the surface (among the lowest modeled vertical levels) and non-logarithmic velocity profiles, as well as their transient nature over the course of a day, forecasting of these phenomena has presented continued challenges for NWP modeling. Given the non-logarithmic shape of speed-up flow wind profiles, in which the wind speed decreases with height through a typical turbine rotor layer, NWP models are likely to overestimate hub-height wind speeds. This could lead to large overestimates of wind energy production. Thus, the goal of this work is to quantify model wind speed bias during observed speed-up events to inform future model improvements, especially for wind energy applications.

The present study aims to evaluate HRRR predictions of boundary layer dynamics in a region with significant wind energy production that features recurring speed-up flows over complex terrain. The analysis focuses on model predictions of local-scale wind profiles, as analysis of localized HRRR performance is useful for model evaluation against lidar observations. However, NWP models exhibit greater predictive skill at larger spatial scales as their spatiotemporal resolutions exceeding those of localized atmospheric phenomena. Therefore, an additional component of this analysis explores the connection between synoptic-scale conditions and model performance to determine synoptic-scale predictors of localized HRRR performance.

The area studied is the Altamont Pass, which is located within the Diablo Range in central California. This location is considered due to its importance for wind energy in California as well as its proximity to a facility operated by Lawrence Livermore National Laboratory (named Site 300), which allows for observations of boundary layer properties in the 0-150 m layer agl in which turbines largely operate. This work follows on the observational analysis performed at this site by Wharton and Foster (2022) as part of the Hill Flow Study, hereafter referred to as HilFlowS (Wharton 2019). As stated in Wharton and Foster (2022), the objective of HilFlowS was to supplement the WFIP2 campaign by providing observations in a region with complex terrain relevant to wind energy generation outside the spatial domain of the WFIP2 campaign.

This study is outlined as follows: Section 2 details the site where observations are recorded, as well as the data (observational data, HRRR model data, and reanalysis data) and analytical methods used for this study. Section 3 provides results from the observational period and an evaluation of HRRR model performance relative to observed conditions. Additionally, this section investigates the association between site-specific HRRR model performance and synoptic- and mesoscale atmospheric conditions (see Section 3d). Afterwards, the utility of HRRR for wind energy forecasting is discussed by exploring wind energy forecast accuracy over an 18 h forecast horizon relative to observations (see Section 3e). Section 4 provides a summary of the findings, a discussion of HRRR performance relevant to boundary layer dynamics and wind energy interests, and suggestions for future work.

2. Site information, data, and methods

a. Site information

The area analyzed in the HilFlowS study is located in north-central California to the east of the San Francisco Bay, between the California Southern Coast Ranges and the San Joaquin Valley (see Figure 1 for a map of the study area). Within this area, relevant sites considered are the Altamont Pass Wind Resource Area (APWRA) and Lawrence Livermore National Laboratory Site 300 (Site 300).

APWRA spans approximately 202 km^2 (≈50,000 acres) along the northern end of the Diablo Range, which runs approximately northwest to southeast, and is a significant wind farm region in California, with nearly 200 operating turbines and a rated capacity of approximately 264 MW during the time of the HilFlowS observation period (Hoen et al. 2018). For the purposes of this study, an important parameter to consider for wind forecasting is median turbine hub height, which is 80 m for newer APWRA turbines (Wharton and Foster 2022).

Site 300 is approximately 10 km southeast of the APWRA along the Diablo Range. The site features variable topography composed of hills, several ridges, and valleys. The site is covered in grassland of roughly uniform height (less than 1 m) and is largely devoid of trees and shrubs. Elevation of terrain within Site 300 ranges from 150 to 500 m above mean sea level (a.m.s.l.), with higher terrain immediately to the south and southwest of the site and gentler downsloping terrain toward the California Central Valley to the east. Topographical variance is high, with typical variations of O(100 m) within 1 km. The slopes of the hills upon which the lidars are mounted reach maximum angles of approximately 20°, although the effective angle is dependent on wind direction.

b. HRRR model dataset

Forecasts from the operational HRRR [HRRRv3, implemented operationally in 2018 and documented in James et al. (2022)] are analyzed and evaluated in this study. The HRRR is nested within the domain set by the NOAA Rapid Refresh (Benjamin et al. 2016), with the HRRR spatial domain ($\Delta x = 3$ km) covering the continental United States. HRRR is rerun hourly, producing 18 h forecasts for most runs, and 48 h forecasts every 6 h (Olson et al. 2019). For results concerning boundary layer dynamics in Sections 3a-d, model data from forecast hour 1 are used to evaluate the ability of HRRR to resolve dynamics observed at Site 300, as HRRR output at forecast hour 1 of was found by (Banta et al. 2021) to have minimum bias. For results relevant to wind energy forecasting in Section 3e, data from forecast hours 0-18 are used to evaluate the ability of HRRR to predict wind energy generation relative to observations. Additional details regarding model setup and data assimilation methods can be found in Benjamin et al. (2016).

For the present analysis, HRRR model grid values were bilinearly interpolated to the observation points, following Pichugina et al. (2019). HRRR hybrid-sigma levels were remapped to align with the vertical levels at which lidar data was available. To ensure remapped levels are representative of the lidar-observed levels, a 5% error tolerance was imposed between HRRR hybrid-sigma levels and lidar levels, with any remapping errors exceeding the tolerance being rejected. For the analysis of the lowest 150 m, the lowest 10 to 12 hybrid-sigma levels were used depending on surface pressure. Given the height variability on hybrid-sigma coordinate levels due to atmospheric conditions, the vertical grid spacing of the hybrid-sigma levels ranged from 2 to 5 m within the first 3 hybrid-sigma levels, 5 to 10 m for the following 5 levels, and 10 to 25 m for the remaining levels. In general, vertical resolution was on the order of that for the Doppler lidars used (~10 m). Further details regarding instrumentation are provided in Section 2c.

c. Observational instrumentation and data availability

Data analyzed for this study were collected by a pair of Doppler lidars located on parallel ridgelines within Site 300, with a meteorological tower located on a smaller, third ridge. The instruments are aligned such that they are directly in line with one another when the winds are from the southwest or northeast. The two vertically-profiling Doppler lidars (ZephIR 300, ZXLidars, United Kingdom) were used for observations of boundary layer winds at several vertical levels. The lidars were deployed at two hilltops (western observation point, *WOP*, and eastern observation point, *EOP*) within the Site 300 facility (see Figure 1). Although the distance between WOP and EOP is approximately 1 km, the observation sites correspond to neighboring HRRR grid cells as shown in the figure. The lidars were operated in a velocity azimuth display scanning mode, with a measurement frequency of 50 Hz and a scan frequency of 1 Hz (50 measurements per scan). The lidars use 55 beams which are emitted from a rotating scanning head at an elevation angle of 30° from the vertical, and are rotated a full 360° to make the conical scan. Each conical scan requires approximately 15 s, as each vertical level is measured individually at 1 Hz. The lidars were oriented using GPS to align the instruments with true north and subsequently cross-validated to ensure agreement in measurements of wind speed and direction. Processed scan output thus resulted in an observational temporal resolution of 15 s. This mode allowed for measurement of the zonal, meridional, and vertical components of wind speed at vertical levels ranging from 10 to 150 m agl. Processed scan output is then averaged over 10 min intervals, allowing mean wind profiles of the surface and lower mixed sublayers of the atmospheric boundary layer to be captured. Note that this observed layer encompasses the vertical extent of the wind turbine rotor disks installed in APWRA. Quality control filtering was performed by (i) eliminating observations recorded during precipitation events, (ii) rejecting lidar data with signal-to-noise (SNR) ratios lower than −22 dB, and (iii) removing outliers exceeding 4 standard deviations from a 30 min window mean centered on the sample time. Installation of lidars in complex terrain introduces the potential for measurement error; for example, Bingöl et al. (2010) concluded that measurements of horizontal wind speed using conically-scanning lidars are on the order of ±10%. The error is introduced by heterogeneity in flow patterns over complex terrain and is considered in throughout this analysis. A 52 m-tall meteorological station (referred to as the meteorological tower) was located on a third parallel ridgeline east of the EOP Doppler lidar and was used to evaluate surface layer properties not captured by the lidars [see Wharton and Foster (2022) for more information].

Data was collected from 7 July to 23 September 2019 for a total observation period of 1872 h in 10 min intervals after internal quality control. For evaluation of HRRR, instrument data is averaged hourly to match the temporal frequency of HRRR output, with averaging windows centered on each hour. After processing and data rejection due to quality control, the WOP lidar retained 1828 h of compliant observational data, the EOP retained 1562 h, and the meteorological tower retained 1316 h. Note that EOP lidar has lower data availability than WOP because the EOP lidar had more downtime due to its electrical source (EOP lidar ran on solar and battery power, WOP ran on grid power) and because of aforementioned filtering of outliers from the time-window means (filtering step iii). Additionally, note that wake effects from APWRA, which lies to northwest of Site 300, are not considered to have effects on observational quality due to the distance between APWRA and the observation site (approximately 5 km for the closest turbines), and the prevailing winds largely coming from the west and west-southwest. Although it has been shown that wake effects downstream from a wind farm are possible at this distance (Christiansen and Hasager 2005; Fitch et al. 2013; Platis et al. 2018), these studies have been perfomed over homogeneous surfaces (flat surfaces in numerical studies, sea surface in observational studies), have accounted for taller turbines than those on the lee side of the APWRA wind farm, and have noted the mitigating effects of rough terrain on wake distance. The location of data collection is considered topographically similar to APWRA given their siting along the Diablo Range and a similar degree of terrain variability at Site 300 and APWRA.

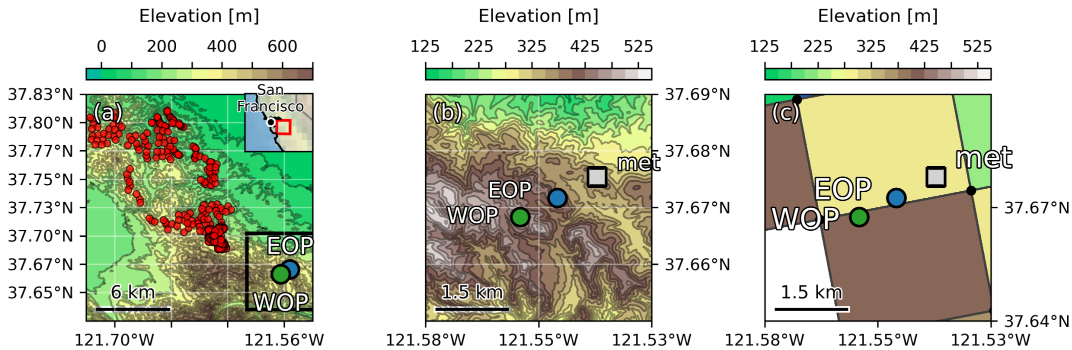

Fig. 1: Topographical map of Lawrence Livermore National Laboratory Site 300 in central California (north is at the top of the map). (a) Regional map showing the location of the APWRA wind turbine complex (red dots indicate individual wind turbines) relative to the Site 300 observation locations (indicated by black box encasing blue and green dots, shown in detail in right inset plot). (b) Inset view showing local map of Site 300 observation locations, with western (WOP) and eastern (EOP) observation point lidars denoted by green and blue dots, respectively, while the meteorological tower is denoted by a grey square. The distance between WOP and EOP is approximately 1 km. Terrain data were obtained from the United States Geological Survey GMTED 2010 survey (Danielson and Gesch 2011) and wind turbine locations were obtained from data provided in Hoen et al. (2018). (c) Inset view showing local map of Site 300 as in panel (b) with soil height data (colored cells) and grid points (black dots) used in HRRR (as a proxy for terrain data) to highlight the spatial resolution of topography within the model.

d. Derived quantities

Several quantities used to analyze HRRR model performance relative to observations are defined in this section.

1) Bias calculation methods

Model bias is defined as:

$$\text{bias} = \psi_{\text{model}} - \psi_{\text{obs}} \tag{1}$$

where ψ is the meteorological variable. For the purposes of this study, *model* refers to HRRR data and *obs* refers to observational data recorded by the lidars at WOP and EOP. A positive bias corresponds to model overprediction and a negative bias corresponds to model undeprediction relative to observations. For bias calculations of horizontal wind properties, a minimum wind speed threshold was established at the 10th-percentile of horizontal wind speeds at the median turbine hub height (80 m agl), as defined in Section 2a.

For several variables analyzed, it is useful to provide the relative (also known as fractional) bias between model and observed values. The relative bias is defined as:

$$\text{relative bias} = 100\left(\frac{\psi_{\text{model}} - \psi_{\text{obs}}}{\psi_{\text{obs}}}\right) \tag{2}$$

2) Rotor-equivalent wind speed

The rotor-equivalent wind speed is a metric used to account for the kinetic energy passing throughout the vertical extent of a swept rotor area (i.e., the span of the wind turbine blades) corresponding to a wind turbine (Wagner et al. 2014). This metric is useful for wind energy forecasting, as it accounts for variations in the vertical wind profile spanning a turbine rotor. The cross-rotor wind variations are often several meters per second (Wagner et al. 2009; Wharton and Lundquist 2012), and can be higher in areas with high vertical wind shear. Accounting for these variations has been shown to improve estimates of wind speeds across turbine rotors (Liu et al. 2021; Sasser et al. 2022), especially in areas with complex terrain and variable boundary layer flows (Van Sark et al. 2019), which has implications for the accuracy of wind energy forecasting.

Rotor-equivalent wind speed is calculated as in Equation 3:

$$U_{\text{eq}} = \left[\sum_{i=1}^{N} U_i^3 \frac{A_i}{A}\right]^{1/3} \tag{3}$$

where i denotes a vertical level, N denotes the number of vertical levels spanning the swept rotor area, U_i is the horizontal wind speed at vertical level i, A_i is the swept rotor area between vertical levels i and $i-1$, and A is the total swept rotor area. Results using this metric are provided in Section 3e for evaluating model bias of horizontal winds in a context relevant to wind energy applications.

e. North American Regional Reanalysis dataset

To provide insight into nonlocal phenomena influencing HRRR performance at Site 300, the North American Regional Reanalysis (NARR; Mesinger et al. 2006) dataset was used to provide daily synoptic-scale meteorological conditions. These conditions were then associated with time windows of maximal and minimal HRRR bias magnitude relative to lidar-observed horizontal wind speeds at hub height (80 m agl). This analysis is intended to identify synoptic phenomena that are associated with maximal and minimal HRRR bias magnitudes, with the goal of determining connections between synoptic-scale phenomena (which are generally forecast with high accuracy) and local conditions (which present a more difficult forecasting problem). NARR data for geopotential height at daily frequency was chosen as an observationally-constrained dataset that is independent from HRRR and is commonly used for mesoscale and synoptic-scale analysis. Note that for this analysis, HRRR bias at each site is averaged over a 3 h period to filter out transient events and allow for a more consistent comparison to NARR. The analysis proceeds as follows:

1. The HRRR bias at each site was averaged over 3 h windows for the entire study period. Window bias magnitudes exceeding one standard deviation (1σ) above the mean over the period were flagged for maximal bias magnitude, while windows with bias magnitudes less than 1σ below the mean were flagged for minimal bias magnitude.

2. Days with multiple 3-hourly windows of maximal or minimal HRRR bias magnitude were identified at each site.

3. To connect patterns in local observations with synoptic-scale wind patterns, days with multiple 3-hourly windows in common at both sites were considered, as these are indicative of days with synoptic-scale forcing contributing to elevated or suppressed HRRR bias magnitude, rather than shorter-lived local phenomena.

4. These days were then identified within the NARR dataset and used to create respective composite mean fields corresponding to conditions during days with maximal and minimal HRRR bias magnitude.

NARR geopotential height data (ϕ) was used at 500 hPa (termed ϕ_{500}) and 850 hPa (termed ϕ_{850}) for synoptic-scale and mesoscale analyses, respectively. The intent of using ϕ_{500} was to identify synoptic patterns that related with local HRRR performance, while using ϕ_{850} allows for the association of regional wind patterns with local HRRR performance. In total, 30 days that met the maximal HRRR bias magnitude threshold and 10 days that met the minimal HRRR bias magnitude threshold were identified during the observation period (total of 40 days among both groups of days).

3. Analysis and results

a. Study area meteorological conditions

A composite of horizontal wind speeds is shown in Figure 2a and b for the WOP and EOP, respectively. Mean horizontal wind speed minima occurred in the morning, with 10 m agl wind speeds measuring an average of approximately $3\,\mathrm{m\,s^{-1}}$ at 10:00 local time (LT) at both sites during the development of the convective boundary layer. Mean horizontal wind speed maxima occurred in the evenings at approximately 20:00 LT, with 10 m agl wind speeds reaching an average of $14\,\mathrm{m\,s^{-1}}$ at EOP and $11\,\mathrm{m\,s^{-1}}$ at WOP. Because the lidars are placed atop hills, the near-surface wind speed maximum is evidence of a speed-up event over the local topography, which is a regular occurrence just before sunset and has been observed at other locations with similar meteorology (Banta et al. 2021; Djalalova et al. 2020; Pichugina et al. 2019).

Figures 2c and d show a diurnal cumulative frequency plot of wind directions for WOP and EOP, respectively. Similar to horizontal wind speeds, wind directions follow a diurnal profile, with winds at all levels being predominantly west- and west-southwesterly ($225 < \phi < 270°$) during evening and overnight hours, with a northerly shift during the morning hours. This diurnal profile reveals the role played by mesoscale winds during the evening and overnight hours, with westerlies driven by onshore flows due to marine air intrusions, largely induced by land-sea temperature gradients (McClung and Mass 2020). The northwesterly shift in winds during the daytime is less attributable to a given phenomenon, but may be a result of flow channeling through the San Pablo Bay and the Sacramento River Delta to the north.

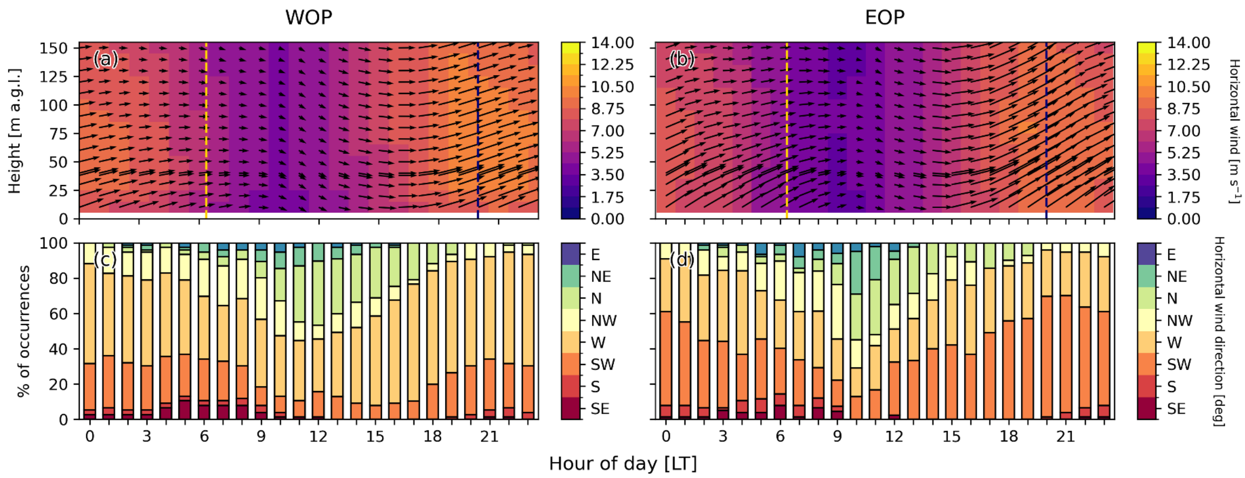

Fig. 2: Diurnal profile of time-averaged observed values for WOP (a, c) and EOP (b, d) for horizontal wind speed (a-b) and horizontal wind direction at 80 m agl (c-d). For panels (a, b), the arrows indicate wind direction, with upward-pointing arrows corresponding to southerly flow and rightward-pointing arrows corresponding to westerly flow. Note that panels (c, d) for wind direction are cumulative frequencies of each wind direction for their given hour. Vertical yellow and blue dashed lines denote approximate sunrise and sunset times at the study area, respectively.

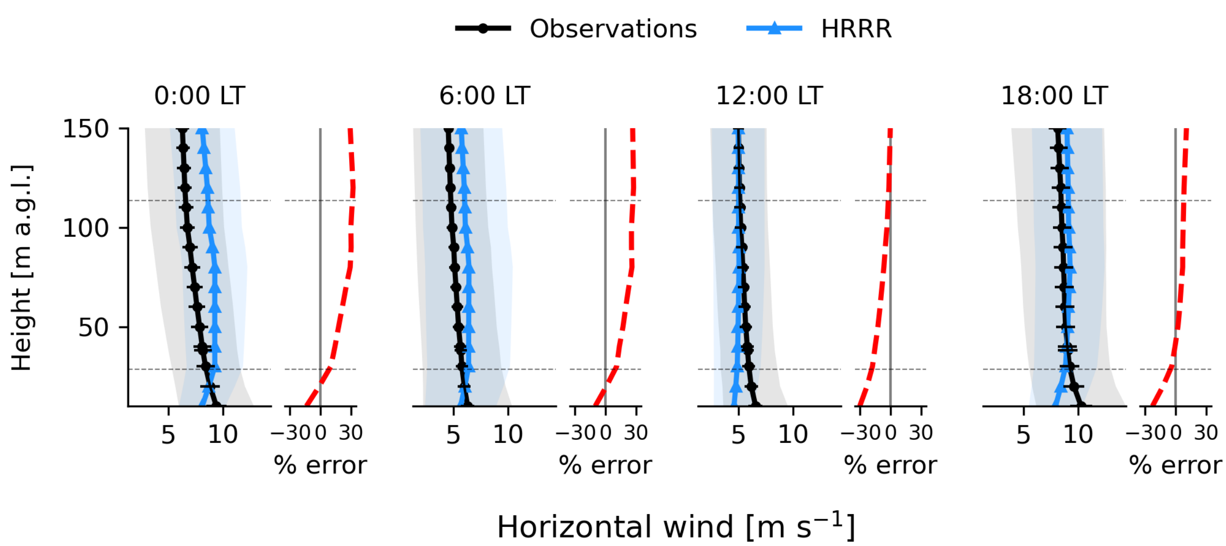

Fig. 3: Site-averaged vertical profiles of observed (black dotted line) and model (blue line with triangular markers) horizontal wind speed hourly averages at 0:00 (midnight), 6:00 (early morning), 12:00 (midday), and 18:00 (early evening) LT, respectively. Mean relative bias (in percent, see Equation 2) for each set of profiles are shown to the right of each plot in red. Grey and blue shading denote one standard deviation from the observed and model means, respectively. Horizontal error bars at the marker points denote a +/- 10% error from the composite mean observed horizontal wind speed to account for instrument error, following Bingöl et al. (2010). Horizontal dashed lines denote the mean minimum and mean maximum rotor extents of turbines installed at APWRA (Hoen et al. 2018).

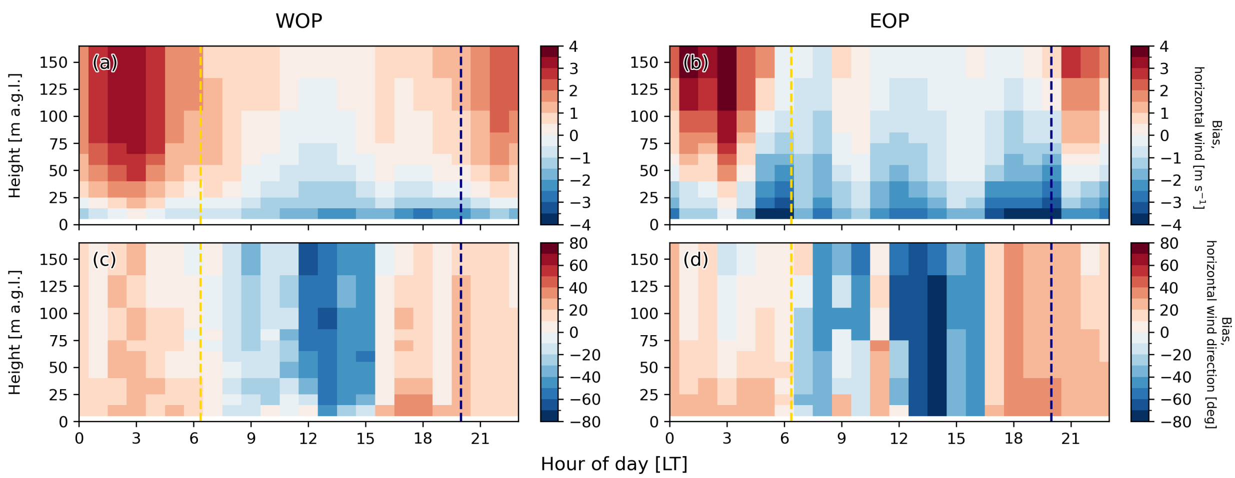

Fig. 4: As in Figure 2, except that model biases are plotted for wind speed (a–b) and wind direction (c-d). Model data used is HRRR output at forecast hour 1.

b. Model performance evaluation of boundary layer dynamics

1) Horizontal winds

The mean diurnal profiles of observed and model horizontal wind speeds among both sites are shown in Figure 3 using hourly averages, shown at 6-hour intervals. Overall, HRRR horizontal wind speed bias was lowest in the afternoon (12:00 to 18:00 LT) and highest during early morning (00:00 to 06:00 LT) when averaged over the observed 150 m. HRRR overpredicted horiztonal wind speeds during nighttime hours and underpredicted during daytime hours, within the 0-150 m layer agl. HRRR also generally underpredicted daytime horizontal wind speeds in the lowest levels (< 30 m) at all times where peak speed-up flows were observed. Note that 30 m generally coincides with the lowest extent of a turbine rotor disk. This analysis considers the potential for measurement error of ±10% associated with lidar usage in complex terrain (Bingöl et al. 2010) (see error bars for observation vertical profiles in Figure 3), although this error magnitude is not expected to change conclusions regarding HRRR bias relative to lidar measurements.

At midnight (0:00 LT), the average observed wind speed ranged from $10\,\mathrm{m\,s^{-1}}$ at 10 m agl to approximately $7\,\mathrm{m\,s^{-1}}$ at 150 m agl, following a decreasing profile with respect to height. Average model wind speeds were $2\,\mathrm{m\,s^{-1}}$ lower than observations at 10 m, although the model vertical profile demonstrated an increase in wind speed with height, following a quasi-logarithmic profile due to the combination of a coarse vertical grid and the Monin-Obukhov boundary condition imposed at the surface. This resulted in an underprediction of wind speed in the surface layer reaching 10%, with the remainder of the vertical wind profile being overpredicted by as much as 30%. By early morning (6:00 LT, immediately before sunrise), observations show that surface layer winds have lessened with near-constant average wind speeds of 5 to $6\,\mathrm{m\,s^{-1}}$ throughout the vertical observational profile. On average, the model predicted the magnitude and vertical profile of winds similarly to overnight hours, with relative errors ranging from 10% underprediction at the surface to a 30% overprediction at 150 m agl. By midday (12:00 LT), average observed winds resumed a reverse shear profile, with 10 m winds averaging $6.4\,\mathrm{m\,s^{-1}}$ and decreasing to approximately $5\,\mathrm{m\,s^{-1}}$ at 150 m. Average model winds resumed a quasi-logarithmic boundary-layer profile, with winds ranging from $4.4\,\mathrm{m\,s^{-1}}$ at 10 m to $5\,\mathrm{m\,s^{-1}}$ at 150 m. This resulted in underpredictions of horizontal wind speed exceeding 20% at the surface, with decreasing underprediction through the observed layer, reaching zero bias at 150 m agl. Daytime biases throughout the observed layer persisted through the early evening (18:00 LT) with surface winds underpredicted by up to 30%, although relative errors throughout most of the observed layer reduced to < 5%. The persistence of strong near-surface bias through the afternoon and evening indicates an underprediction of speed-up events that are characteristic of boundary layer flows in the study area.

Model bias in horizontal wind speed prediction follows a diurnal pattern at both sites, as shown in Figure 4. Nocturnal winds above the surface (> 25 m agl) are overpredicted, with peak overpredictions occurring during the decay of the evening speed-up events. At sunrise, model bias decreases throughout the observed layer to $< 1\,\mathrm{m\,s^{-1}}$ at both sites. However, a negative model bias (model underprediction) develops throughout the morning, with peak underpredictions reaching $4\,\mathrm{m\,s^{-1}}$ near the surface (< 25 m agl), with underprediction magnitudes lessening with height. Model biases reach greater magnitudes for over- and underpredictions at EOP than at WOP, which may be a result of predominantly westerly flows reaching the WOP observation site relatively unobstructed by prominent topographical features upstream of the observation site. In contrast, EOP is downstream of WOP during westerly flows and is at a lower height, potentially subject to flow perturbations at scales that are unresolved by HRRR.

It is noted that the diurnal pattern of wind speed bias suggests a correlation between atmospheric stability and model performance that could be investigated in future work. However, this analysis is not pursued here due to a combination of observational constraints (i.e., the lack of high-frequency temperature observations at the lidar sites) and the limitations of conventional stability estimates in complex terrain (Albornoz et al. 2022; Peterson and Hennessey Jr 1978; Touma 1977).

2) Wind direction

Model performance between sites for wind direction followed similar composite mean diurnal profiles among sites througout the depth of the observed layer, as shown in Figure 4c-d. Positive composite mean wind direction model biases were typical throughout the overnight and early morning hours, which suggest a more westerly and northwesterly component in modeled flows relative to observed flows, given that observed winds are primarily westerly and southwesterly during these times. Throughout the day, model biases become negative, with strongest negative biases exceeding 60° during the early afternoon at both sites. Given that wind directions shift northwesterly during the daytime, the negative wind direction biases during the early afternoon suggest that HRRR continues predicting primarily westerly flow, and may not resolve local daytime shifts in wind direction. Into the evening hours, composite mean wind direction model biases become positive again, with vertically-averaged values of approximately 30% at both sites during hours of observed mean westerly and southwesterly flow, again suggesting a westerly bias in HRRR predictions of flow direction.

The difference in bias characteristics between WOP and EOP can be attributed to the observed differences in composite mean flow directions among these sites. The modeled composite mean wind directions are similar between sites, given that they are in neighboring cells. However, as seen Figure 2c-d,the observed wind direction composite means show a disparity between sites. Namely, WOP demonstrates a relatively higher cumulative frequency of winds with a southerly component during the daytime than EOP (see Figure 2c), while EOP shows a higher portion possessing a northerly component (Figure 2d). Therefore, it can be deduced that bias characteristics are different among sites due to effects of complex terrain that are unresolved by HRRR.

Due to the complex terrain surrounding the observation sites, flow properties are likely to be strongly dependent on the direction of the prevailing wind. To investigate the relationship of horizontal wind speed model bias with the direction of the flow, the mean absolute errors of HRRR wind speed predictions relative to observations are shown by direction in Figure 5 at 40, 80, and 150 m agl. At 40 m agl, the largest errors in horizontal wind prediction occur for winds coming from the southeast at both sites, with relative errors reaching 50%, whereas small errors occur for winds coming from the west and southwest, with errors reaching 30%. Similar patterns are evident at 80 and 150 m agl, with southeasterly and easterly winds being associated with the largest horizontal wind speed errors and westerly winds being associated with the smallest. Note that sample sizes are considerably larger for winds with a westerly component than for winds with an easterly component, which may partially explain the difference in mean error values between the different directions. However, error distributions were found to be significantly ($p < 0.01$) different using a 2-sample Kolmogorov-Smirnov test, indicating that errors from the different directions are characteristically different.

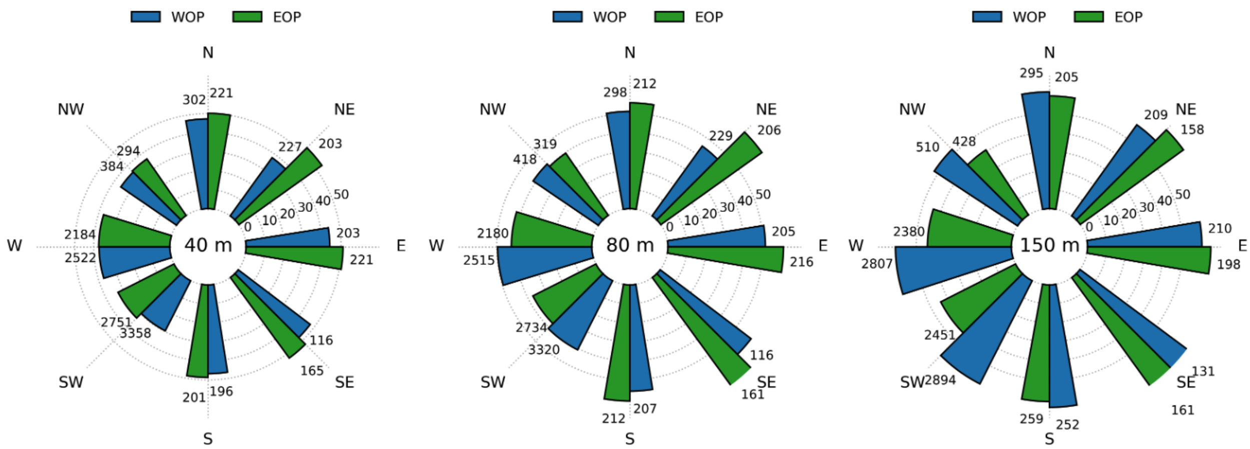

FIG. 5: Wind rose plots showing time-averaged mean absolute errors of horizontal wind speed predictions relative to observations at both lidar observation sites (EOP and WOP) at 40, 80, and 150 m agl, respectively. Model data used is HRRR output at forecast hour 1, which corresponds to a 1 h lead time. Concentric circles denote percentage error (error labels located in the E-NE sector of the plot), while numbers next to the bars indicate number of unique observations for each wind direction bin. Bar width is proportional to the number of unique observations for each wind direction bin.

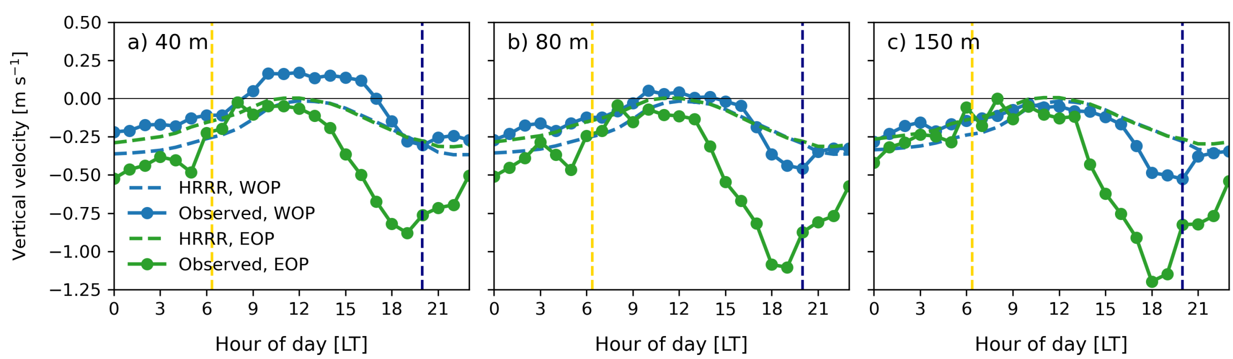

FIG. 6: Diurnal composite means of vertical velocity (w) at WOP and EOP at (a) 40, (b) 80, and (c) 150 m agl, respectively. Diurnal composite means of modeled w are shown in dashed lines, while diurnal compositemean of observed w are shown in solid lines with circle markers. Model data used is HRRR output at forecast hour 1.

c. Vertical velocity

Composite mean diurnal profiles of observed and modeled vertical velocities at both observation points are shown in Figure 6. Note that modeled and observed vertical velocity should be considered qualitatively due to limitations of vertical velocity measurements in complex terrain using the VAD scanning method described in Section 2c (Bingöl et al. 2010). Despite agreement in diurnal trends of observed composite mean w, magnitudes of w may be impacted by the effects of terrain-induced flow that is not properly captured by this measurement procedure, thus preventing conclusions of speed-up events based on measurements of w to be made.

The diurnal profile of insolation can be inferred from patterns of w in both observed and modeled data, with a cycle of near-neutral and positive w (i.e., upward motion) at the surface during the day and negative (i.e., downward motion) near-surface w overnight. Vertical velocities at higher vertical levels generally follow a similar profile, although rising motions are weaker during the day while strong subsidence occurs during the early evening hours before sunset. The coincidence of mean downdraft peak magnitudes and near-surface horizontal speed-up events (see Figure 2) suggests that localized surface divergence connects these phenomena. At EOP and WOP, w magnitude maxima occur during downdrafts above 50 m agl, with average subsidence values reaching -1 $\mathrm{m\,s^{-1}}$, as compared to maximum mean vertical velocities of 0.3 $\mathrm{m\,s^{-1}}$ at WOP during the early afternoon. The difference in observed composite mean profiles of w between sites is not reflected in the model composite mean profiles, further indicating that HRRR does not resolve heterogeneous flow properties in a region with complex terrain.

Notable differences between observation sites are evident in composite mean diurnal profiles of observed w. With regards to intersite differences in diurnal profiles of observed composite mean w, peak differences occur during the late afternoon and evening hours. Magnitudes reach 0.6 $\mathrm{m\,s^{-1}}$ at approximately 18:00 LT, which is coincident with times of strongest horizontal winds and speed-up events, and persist but decrease overnight. Interestingly, intersite differences in diurnal profiles of observed composite mean w decrease with increasing height outside of speed-up events, which provides further evidence of the effects of intersite terrain variability on near-surface dynamics.

In the observed dataset, it is evident that EOP experiences much stronger composite mean downdrafts at sunset relative to WOP. This may be a result of high terrain variability, such that peaks upwind of EOP (including the hill upon which WOP is situated) generate lee effects and lead to stronger downdrafts downwind, such as at EOP. This is further evidenced by WOP experiencing stronger mean updrafts than EOP, which may be a result of terrain-driven flow due to its steeper grade, its topographical prominence leading to unobstructed insolation and subsequent surface heating, as well as weaker effects from neighboring peaks. In contrast, EOP experiences stronger mean downdrafts, which may be terrain-driven due to its position in the lee of the Diablo Range and its lower prominence relative to surrounding peaks.

d. Synoptic-scale atmospheric conditions associated with wind speed bias

Reanalysis and NWP models are skilled at representing synoptic-scale phenomena, such as synoptic-scale dynamics. Given that synoptic-scale processes influence those at smaller scales (i.e., mesoscale and local scales), analyzing synoptic-scale processes may provide insight into patterns influencing local-scale NWP biases. This approach is taken to explore the relationship between synoptic-scale and regional wind patterns, with the goal of identifying a relationship between synoptic patterns and HRRR forecast bias magnitudes at Site 300. This relationship between patterns at different horizontal scales is investigated in this portion of the analysis using the methodology outlined in Section 2e for geopotential heights at 500 hPa (ϕ_{500}) and 850 hPa (ϕ_{850}), respectively.

Contours of composite mean geopotential heights at ϕ_{500} and ϕ_{850} during identified maximal and minimal bias magnitude days are provided in Figure 7. Accordingly, analysis of reanalysis data and model performance is discussed in terms of synoptic-scale and mesoscale conditions. Additionally, standardized anomalies of ϕ_{500} and ϕ_{850} are derived to investigate synoptic and mesoscale patterns associated with days of maximal and minimal model bias magnitude.

At Site 300, NARR-derived mean ϕ_{500} was 5872 m with a standard deviation of 53 m over the study period. During days with maximal model bias magnitude, ϕ_{500} featured a composite mean of 5901 m with a standard deviation of 30 m, which corresponds to a standardized anomaly of +0.53σ relative to mean ϕ_{500} over the duration of the study period. The synoptic setup of ϕ_{500} shown in Figure 7a shows highest ϕ_{500} values situated over the southwestern United States with decreasing ϕ_{500} towards the Pacific coast, suggesting ridging over the western United States during days with highest model bias magnitudes at Site 300. The composite mean standardized anomalies of ϕ_{500} show further anomalously high ϕ_{500} over the Pacific coast during days of maximal model bias magnitude (see Figure 7c), which indicates the presence of anomalously high pressure near Site 300 during days when bias magnitude is largest.

During days with minimal model bias magnitude, NARR-derived mean ϕ_{500} was 5826.4 m with a standard deviation of 78 m, which corresponds to a standardized anomaly of -0.61σ relative to mean ϕ_{500} over the duration of the study period. Figure 7b shows the synoptic setting at 500 hPa, revealing low values of ϕ_{500} over the Pacific coast relative to zonal means, suggesting a trough over the western United States during days with minimal model bias magnitude at Site 300. Composite mean standardized anomalies of ϕ_{500} show anomalously low ϕ_{500} to the northwest of Site 300, indicating anomalously low pressure near Site 300 during these days. The standardized anomaly pattern during minimal model days is of a similar location, similar magnitude, and opposite in sign to the pattern shown for standardized anomalies during maximal model bias magnitude days. Note that the composite mean values for ϕ_{500} exceed the NARR 40-year July-September mean during days of maximal model bias magnitude and are below the NARR 40-year July-September mean during days of minimal model bias magnitude in the region surrounding Site 300 (Brewer and Mass 2016).

The association of anomalously high ϕ_{500} values and ridging with weaker model performance, as well as troughing with stronger model performance, suggests that synoptic regimes play a role in HRRR predictive skill for low-level winds at Site 300. 500 hPa ridging is often associated with anomalously-weak horizontal winds and a relative increase in the contributions of thermodynamically-induced multiscale effects on local-scale dynamics. The primary regional factor contributing to local-scale dynamics is a strengthening of the sea breeze circulation, while local factors include stronger vertical motion and heat transport due to weakened horizontal winds and increased insolation (Banta et al. 2021; Brewer et al. 2012).

In contrast, 500 hPa troughs are associated with stronger and less variable onshore winds from the Pacific, resulting in cool air intrusion over the western United States that heightens the effect of the dynamical contribution to wind speeds relative to the effects of thermodynamic contributions (Banta et al. 2021). These findings imply that above-average localized HRRR performance occurs during periods with a synoptic pattern associated with uniform winds (i.e., low temporal variability in wind speed and direction) over Site 300, while below-average HRRR performance occurs during periods with a synoptic pattern associated with weaker winds and heightened regional-to-local scale thermodynamic contributions. This aligns with findings in Banta et al. (2021) in the Columbia River basin over the northwestern United States, which showed that HRRR performance improved during days with stronger synoptic-scale wind speeds and reduced contributions from diabatic heating processes and warm-air advection.

To provide a more direct connection between synoptic-scale atmospheric conditions and hub-height winds at Site 300 (i.e., local scale), the 850 hPa level was also evaluated to roughly approximate the interface between the free troposphere and the boundary layer. At Site 300, NARR-derived mean ϕ_{850} was 1519 m with a standard deviation of 18 m over the study period. Days with maximal model bias magnitude featured a composite mean ϕ_{850} of 1522 m with a standard deviation of 19 m, presenting a standardized anomaly of +0.32σ relative to mean ϕ_{850} over the study period. The mesoscale distribution of ϕ_{850} shown in Figure 7e shows a strong ϕ_{850} gradient to the west of the Pacific coast with a weakening gradient over land, suggesting strong offshore winds with slower flow over central California. The composite mean standardized anomalies of ϕ_{850} show slightly above-average high ϕ_{850} over Site 300. This anomaly pattern indicates that horizontal flow near the boundary layer interface is somewhat weaker than the study period mean (see Figure 7g). Similar to composite anomalies of ϕ_{500}, the ϕ_{850} anomaly pattern further suggests the presence of anomalously high pressure near Site 300 during days when model bias magnitude is largest. Similar to composite anomalies of ϕ_{500}, the ϕ_{850} anomaly pattern further suggests the presence of anomalously high pressure near Site 300 during days when model bias magnitude is largest.

On days with minimal model bias magnitude, NARR-derived mean ϕ_{850} was 1510 m with a standard deviation of 23 m, corresponding to a standardized anomaly of -0.61σ relative to composite mean ϕ_{850} over the duration of the study period. Figure 7f shows the composite mean mesoscale distribution of ϕ_{850} on days with minimal model bias magnitude over Site 300, revealing a stronger gradient of ϕ_{850} relative to days with maximal model bias magnitude and the surrounding region. In contrast to the pattern of ϕ_{850} during days with maximal model bias magnitude, the gradient magnitude implies stronger and more uniform flow (i.e., less temporal variability in wind speed and direction) at 850 hPa during days with minimal model bias magnitude. This is reinforced by the spatial distribution of composite mean standardized anomalies of ϕ_{850} in the area surrounding Site 300. As shown in Figure 7f, Site 300 is flanked by negative anomalies of 850 hPa to the north and positive anomalies of 850 hPa to the south, indicating a favorable dynamical setup for enhanced horizontal flows relative to the study period mean. As for composite mean values of ϕ_{500}, values of ϕ_{850} exceed the NARR 40-year July-September mean during days of maximal model bias magnitude, and are below the NARR 40-year July-September mean during days of minimal model bias magnitude in the region surrounding Site 300 (Brewer and Mass 2016).

During days of maximal HRRR bias magnitude relative to lidar observations, the ϕ_{850} composite mean shows northerly flow across Site 300 (see Figure 7c). Composite mean ϕ_{850} at Site 300 reached 1523 m, which exceeds mean ϕ_{850} values corresponding to the monthly mean conditions from July to September at Site 300 in the NARR dataset (1979 to 2019) (Brewer and Mass 2016). During days of minimal HRRR bias magnitude, the ϕ_{850} composite shows stronger northwesterly onshore flow at 850 hPa over central California (see Figure 7d). Composite mean ϕ_{850} at Site 300 reached 1510 m, which is near (within 5 m) the mean ϕ_{850} values corresponding to the monthly mean conditions from July to September at Site 300 in the NARR dataset (1979 to 2019).

Two notable differences arise in comparing the ϕ_{850} setup between maximal and minimal HRRR bias magnitude days: the (1) direction and (2) magnitude of the 850-hPa geopotential height gradient. Regarding (1), days with maximal HRRR bias magnitude show meridionally-oriented contours, suggesting mean northerly flow over Site 300. In contrast, days with minimal HRRR bias magnitude show both zonal and meridional components, resulting in mean northwesterly flow over Site 300. Assuming flow at 850 hPa follows the geopotential contours, the composite analysis demonstrates the role of wind direction in model skill for forecasting winds. Results suggest that the more westerly the flow, the shorter the path for an air parcel to take over land, reducing the opportunity for frictional and topographic effects to perturb the prevailing flow. Regarding (2), days with maximal HRRR bias magnitude show a lesser ϕ_{850} gradient compared to days with minimal HRRR bias magnitude, indicating that the pressure gradient over Site 300 is weaker and consequently, that horizontal winds over Site 300 are weaker.

To further investigate the relationship between the ϕ_{850} gradient and HRRR bias magnitude, the gradient of ϕ_{850} along a given path s_i (where the subscript i denotes an individual path) normal to the composite-mean contours was analyzed for individual days identified as maximal and minimal bias magnitude days, respectively. This approach has previously been used to evaluate numerical model performance by using the connection between surface layer dynamics and larger-scale factors (Collins et al. 2024a,b; Goutham et al. 2021). Twelve paths s were selected at approximately 0.5° latitude intervals along the California coast with path lengths of 500 km, oriented from the west-southwest (247.5° heading) direction to the east-northeast (67.5° heading) direction, roughly normal to ϕ_{850} contours composited over all identified days (see Figure 7e and f for an overlay of transects on the region). The distribution of the resultant gradients, $\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i$ (i.e., the geopotential gradient evaluated at a path s_i), are shown for maximal (red) and minimal (blue) days in Figure 8. Values of $\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i$ during maximal HRRR bias magnitude days followed an approximately-normal distribution, with a mean value of -0.03 m km^{-1} and standard deviation of 0.02 m km^{-1} (negative gradient denotes decreasing geopotential height moving eastward). In comparison, values of $\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i$ during minimal HRRR bias magnitude days followed a wider distribution, with a mean value of -0.05 m km^{-1} and standard deviation of 0.02 m km^{-1}.

Overall, days with minimal HRRR bias magnitude featured mean gradient values with magnitudes 1σ greater than those on days with maximal HRRR bias magnitudes, where σ is the standard deviation of the distributions of $\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i$. Moreover, several instances of gradients during maximal bias magnitude days show a reversal of gradient direction $\left[\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i > 0\right]$, which does not occur during minimal HRRR bias magnitude days, highlighting the association between westerly flow and improved HRRR performance.

Note that a potential shortcoming of using ϕ_{850} in this analysis is presented by higher elevations to the east of the San Joaquin valley, which may intersect the 850 hPa pressure level. Despite this potential issue, we note that transects used for gradient evaluation do not intersect areas with elevations that are high enough to cross ϕ_{850}.

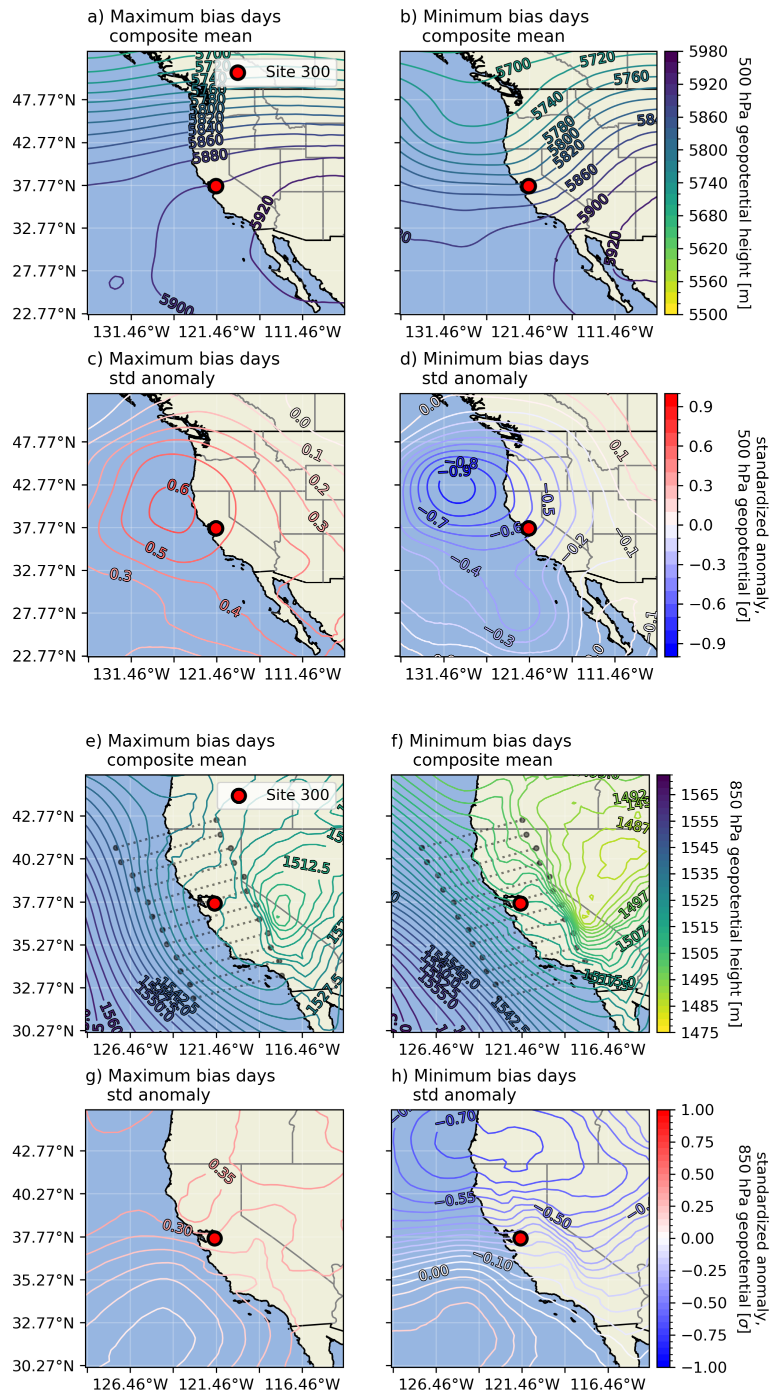

FIG. 7: Composite means of 500 hPa and 850 hPa geopotential heights are shown in panels a-b and e-f, respectively, corresponding to days with maximal and minimal HRRR bias magnitude. All geopotential height contours are in units of meters. Similarly, composite mean standardized anomalies of 500 hPa and 850 hPa geopotential heights are shown in panels c-d and g-h, respectively. Note that a composite is generated using a synthesis of observations from both sites (WOP and EOP), as the distance between WOP and EOP is negligible relative to the spatial resolution of NARR. Site 300 is demarcated by the red dot. The coloring of each geopotential contour corresponds to the geopotential height, as denoted in the contour labels and the colorbar. Transects for the analysis at the 850 hPa level are represented in panels e-f by the dotted lines approximately normal to the contours. Note that the region considered in composite means of 500 hPa geopotential heights (panels a-d) features a larger spatial extent than the region considered in composite means of 850 hPa geopotential heights (panels e-h).

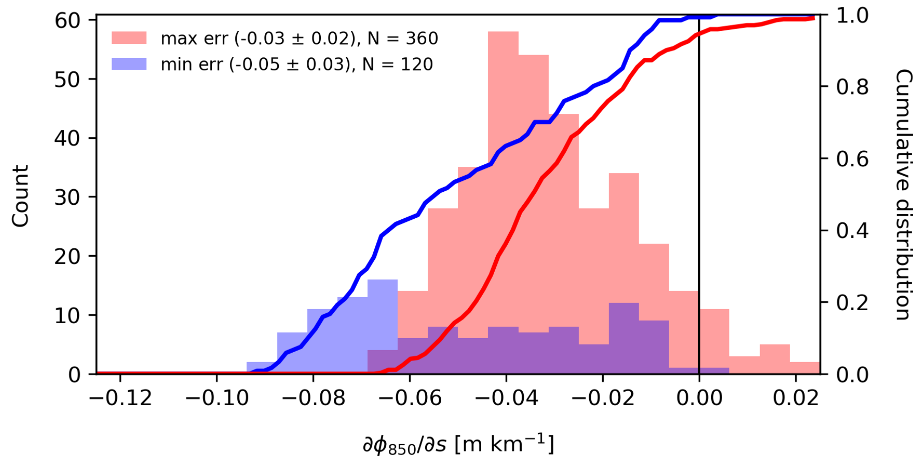

FIG. 8: Histogram (bars) and cumulative distributions (curves) for gradients $\partial\left(\phi_{850}|_{s_i}\right)/\partial s_i$ of 850 hPa geopotential height ϕ_{850} along a transect s_i normal to geopotential contours over Site 300 for days with maximal (red) and minimal (blue) HRRR bias magnitude (sample size $N = 360$ and $N = 180$, respectively). The distributions are different to a statistically-significant degree (p ¡ 0.01) using a 2-sample Kolmogorov-Smirnov test. Note that samples are synthesized from observations from both sites given the synoptic-scale analyses.

e. Wind energy forecasting performance

To assess the ability of HRRR to forecast wind power generation in the nearby APWRA, power curves were obtained from generic turbine models provided by the National Renewable Energy Laboratory (NREL) (see Figure 9). These curves are scaled from International Energy Agency (IEA) turbine models developed through IEA Wind Task 37. Specifically, NREL models NREL-1.7-103 (1.7 MW) and NREL-2.3-107 (2.3 MW) [both downscaled from IEA-3.4-130-RWT, Bortolotti et al. (2019)] are used, which match the turbine characteristics of most of the turbines in APWRA. Power curves are obtained as a function of hub-height wind speed, although rotor-equivalent wind speed (U_{eq}, see definition in Equation 3) is used as the metric for the analysis herein. This metric is intended to provide a more representative measure of horizontal wind speeds over the full vertical extent of the turbine rotor region (Wagner et al. 2009, 2014).

The power curves and wind forecasting analysis presented here are intended to illustrate the potential effect of HRRR wind speed biases on wind energy forecasting, rather than serving as a precise representation of forecast biases in APWRA. Note that the analysis does not consider the horizontal variability in winds over the entire APWRA, nor the variability between APWRA and Site 300. Rather, the analysis uses characteristic wind profiles from Site 300 that are assumed to be representative of the conditions in the region surrounding APWRA. Note that this analysis will focus on the NREL-1.7-103 power curve for turbines rated at 1.7 MW given the similarity in power curves (see Figure 9), as the primary difference between curves is rating magnitude.

To begin understanding model biases in power generation forecasting, model bias in rotor-equivalent wind speed predictions were analyzed using a composite hourly-averaged mean (see Figure 10). At both sites, a diurnal trend in model bias exists, with model overprediction of rotor-equivalent wind speeds during overnight hours and underprediction during daytime hours. Overprediction magnitudes are greater at WOP than EOP, with biases exceeding $3\,\mathrm{m\,s^{-1}}$ at 02:00 LT, whereas EOP biases reached $2\,\mathrm{m\,s^{-1}}$ around the same time. Bias magnitudes decreased toward 0 shortly after sunrise at both sites, and increased again through the mid-afternoon, with WOP underpredictions reaching $-1\,\mathrm{m\,s^{-1}}$ and EOP exceeding $-3\,\mathrm{m\,s^{-1}}$. These biases decreased again toward 0 shortly after sunset, before increasing to overpredictions again into the nighttime. Variance in observed and model composite mean wind speeds followed similar diurnal profiles at both sites, with modest increases in wind speed variance during periods of stronger winds (notably at sunset, when speed-up flows occur) and decreases in wind speed variance during daytime hours.

The NREL power curves are used to generate estimates for composite diurnal power generation from lidar observations and HRRR predictions. As shown in Figure 11, estimated power generation based on observed winds at midnight (00:00 LT) was approximately 0.70 MW at WOP for the 1.7 MW NREL curve, while estimated generation at EOP at midnight (00:00 LT) was approximately 0.90 MW. This decreases overnight through the morning to near-zero values at both sites, before increasing to its diurnal peak after sunset at approximately 1.25 MW at WOP, and at EOP to 1.60 MW. Estimated power generation based on HRRR winds, and correspondingly the model biases, follow a similar diurnal profile. Substantial overpredictions occur overnight, with model estimates of power generation exceeding observational predictions by up to 0.50 MW at both sites. As shown in Figure 12, daytime model bias magnitudes decrease to near-zero at WOP, whereas underpredictions reach 0.70 MWduring the mid-afternoon at EOP.

It can also be seen that estimates of generated power are most sensitive to changes in wind speeds during periods of wind speeds between 6 and 8 $\mathrm{m\,s^{-1}}$ (refer to Figure 9), which may explain why periods with temporally-variable wind speeds but low wind speed bias magnitude (such as the period between 14:00 and 17:00 LST for WOP and 21:00 to 1:00 LST for EOP) have moderate to high errors for estimated power generation. Despite these biases, estimated power generation profiles based on observed and modeled wind speeds are similar at both sites, given that the diurnal profile of wind speed is captured in HRRR and composite mean hub-height winds are often simulated within 1σ of observed winds, as shown in Figure 10.

Analysis of model bias in estimated power generation was also performed over HRRR's 18 h forecast horizon. Although HRRR forecasts are initialized on an hourly basis, the analysis of model bias over the forecast horizon samples each forecast at 3 h intervals. The intent of this analysis is to determine HRRR prediction skill in forecasting power generation relative to available power from observed winds. As shown in Figure 12, several trends in prediction skill are apparent. With respect to the diurnal cycle, a diurnal trend in model bias is persistent throughout the forecast horizon, with strong overpredictions during overnight hours and minimal bias at WOP to moderate underpredictions at EOP during the daytime hours. With respect to the forecast horizon, overpredictions become greater with increasing forecast hour, as overpredictions reach their maxima for both sites at 18 h. The ratio of model bias relative to the turbine power ratings reaches approximately 70% at WOP and 50% at EOP, respectively, for both power ratings, suggesting that HRRR tends to overpredict power at all forecast horizons, especially overnight. The lower biases during daytime hours suggests skillful daytime forecasts, which are critical due to common temperature-driven load increases during the day. However, most of the diurnal cycle exhibits large overpredictions at both sites, indicating a need for improved modeling of boundary layer winds to improve short-term wind energy forecasting. For 2.3 MW-rated turbines, similar trends were found for all analyses performed in related to power generation using the NREL-2.3-107 power curve.

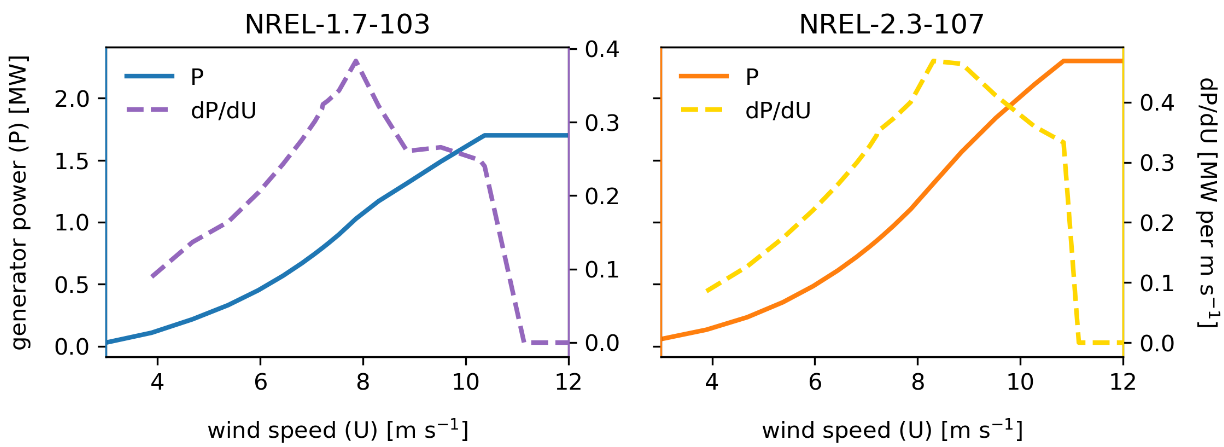

FIG. 9: Power curves generated for generic turbine models using the NREL-1.7-103 (1.7 MW rated generator power, left) and NREL-2.3-107 (2.3 MW rated generator power, right) curves. Solid lines denote generator power P as a function of horizontal wind speed U and dashed lines denote the sensitivity of generator power to changes in wind speed (dP/dU).

4. Summary and conclusions

This study used observational profiling Doppler lidar data to evaluate performance of the HRRR model in predicting lower atmospheric boundary layer winds at two complex-terrain sites near the APWRA. This region is characterized by recurring local-scale speed-up flows that occur as summertime westerly winds are channeled through the Altamont Pass, a gap in the Diablo Range. Over the study period in mid-to-late summer 2019, model biases of horizontal wind speed exhibited a dependence on time of day and height. The diurnal variability of horizontal wind speed bias was made apparent by HRRR overprediction during overnight and early morning hours above the

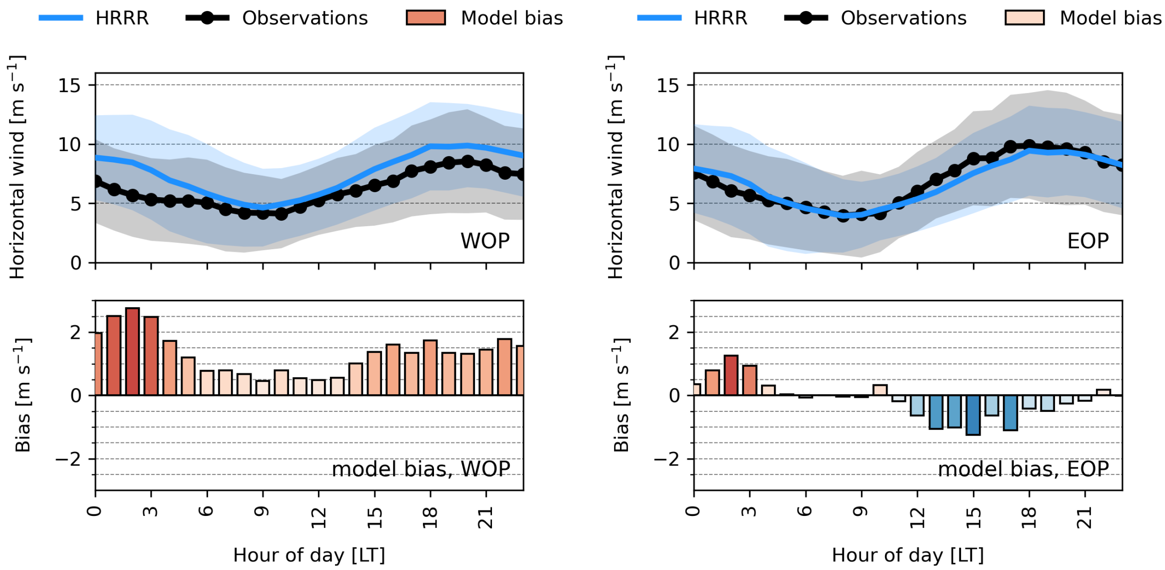

FIG. 10: Composite diurnal hourly means of derived rotor-equivalent wind speeds using observed winds (black line with circle markers) and HRRR (forecast hour 1) winds (blue line, no markers) at WOP (left) and EOP (right). Grey and blue shading denote one standard deviation from the observed and model means, respectively. Note that composite mean model bias (HRRR - observations) are shown by the bars, with red bars indicating HRRR overprediction and blue bars indicating underprediction.

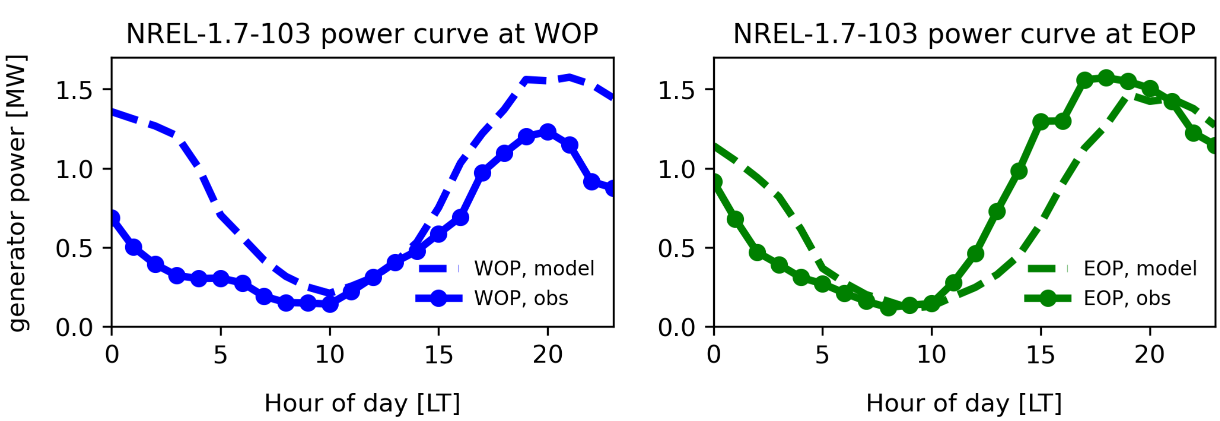

FIG. 11: Estimated composite mean hourly generated power for wind turbines at both sites using observed (solid line) and model (dashed line) winds based on rotor-equivalent wind speeds, provided for turbines with a 1.7 MW rating. The left column corresponds to estimates at WOP, while the right corresponds to estimates at EOP. Power curves provided by the International Energy Agency (IEA) and the National Renewable Energy Laboratory (NREL) (Bortolotti et al. 2019). Model data used is HRRR output at forecast hour 1.

surface layer, with an underprediction of lesser magnitude occurring during the daytime. The

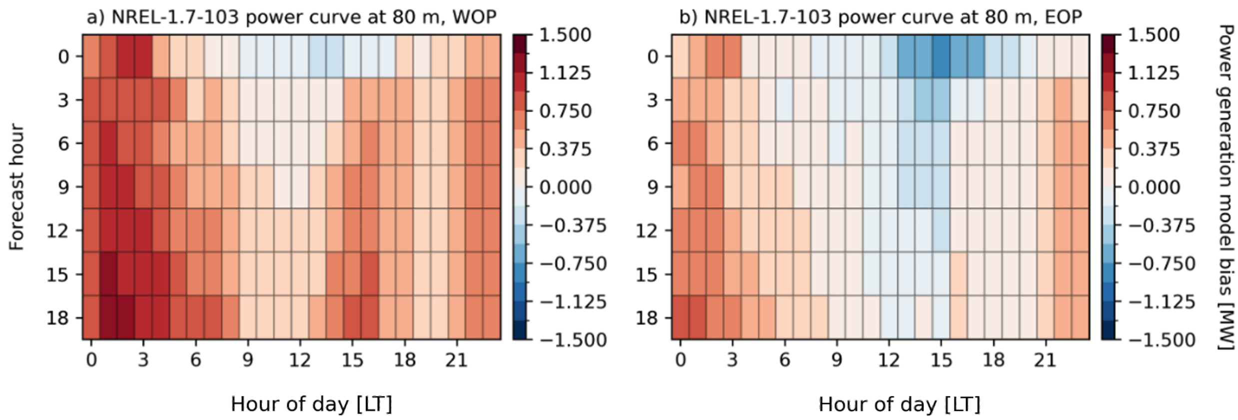

FIG. 12: Power generation forecast biases (HRRR - observations) for wind turbines at WOP (left column) and EOP (right column), provided for turbines with 1.7 MW ratings. Forecasts are based on composite hourly mean wind speeds at hub-height (80 m agl). Note that forecast horizon increases downward along the y-axis, with red-shaded cells indicating HRRR overprediction and blue-shaded cells indicating underprediction.

diurnal variability in model biases was largely dependent on height, with model underprediction maxima occurring within the lowest 30 m agl and overprediction occurring above 100 m agl.

These dependencies are related to near-surface speed-up events, which were consistently observed at the study site but were not captured by the model. At both lidar sites, a near-surface jet-let like flow with a peak wind speed around 10 m agl develops during the evening and continues into the night. Due to a combination of factors, HRRR is generally unable to capture this non-logarithmic flow profile. These factors include limited horizontal resolution of topographic effects, limited vertical resolution of near-surface gradients, and a surface boundary condition based on Monin-Obukhov similarity theory, which assumes a logarithmic flow profile. In the absence of increased resolution, which would be computationally expensive, these results suggest that HRRR could benefit from a modified boundary condition that is able to parameterize terrain-driven non-logarithmic flows. Such a parameterization could substantially improve near-surface wind speed (and thus wind energy) predictions.

Investigation of additional factors related to forecast bias for horizontal winds was performed by evaluating wind speed bias based on prevailing wind direction and synoptic-scale conditions. Bias magnitudes were generally highest during periods with non-westerly flows at both lidar observation sites. Locally, maximum wind speed biases occurred during periods of southerly and easterly flows at all heights. On the synoptic-scale, days with maximal HRRR bias magnitude coincided with days during which ridging occurred over Site 300. Connecting findings from the local- and synoptic-scales, it can be inferred that weaker wind speeds and more variable wind directions are associated with increased HRRR wind speed bias magnitudes. In contrast, horizontal wind speed bias magnitudes were minimal during periods when the prevailing flow over Site 300 had a westerly component. This onshore flow pattern was more constant in time and maintained higher wind speeds than days with maximal HRRR bias magnitude, at both local- and synoptic scales. Synoptic-scale analyses showed that days with minimal wind speed bias magnitude were associated with 500 hPa troughs and strong 850 hPa geopotential height gradients occurred with the presence of strong onshore winds. These findings indicate that HRRR performance (and therefore wind energy forecasting performance) can be linked to synoptic-scale conditions, which are generally predicted more accurately and at longer lead times than boundary layer conditions in NWP models. Given that the prevailing wind direction is westerly at Site 300 throughout the observed layer, this analysis provides evidence that HRRR can be a useful forecasting resource for wind energy applications in the APWRA.

Several similiarities were found between results in this study and those from the WFIP2 field campaign, despite differences in site terrain and composite mean conditions. Pichugina et al. (2019) found that HRRR underpredicted the strongest wind speeds at all observation sites, with the greatest underpredictions occuring during the summer, due in part difficulty capturing the diurnal profile of observed horizontal winds. Several studies analyzing WFIP2 observations and corresponding HRRR runs (Bianco et al. 2019; Pichugina et al. 2019, 2020) noted that HRRR wind speed biases were largest during the nighttime over observed periods (often exceeding 2 $\mathrm{m\,s^{-1}}$ at 80 m agl) which is also found in this study. Moreover, these biases were often amplified during summertime months due to the occurrence of speed-up events during the evening transition. Additionally, it was noted that results were highly variable between sites over the study region, stressing the need for a dense observational network in complex terrain. Banta et al. (2021) noted that HRRR wind speed biases in the rotor layer were lower during periods of westerly flow driven by synoptic-scale forcing, while biases increased during periods with dominant thermal forcing fostered by upper-level ridging.

The findings in this study lead to several potential avenues for future research near the APWRA and other complex-terrain regions. The primary avenue is to employ numerical models with higher spatial resolution in an attempt to capture processes that are hypothesized to be occurring at scales smaller than 3 km. With increased resolution, the observed speed-up events and associated turbulence might be captured in the model, thus reducing bias. Schemes that account for increased horizontal flow variability, such as the three-dimensional planetary boundary layer (3DPBL) scheme developed by Juliano et al. (2022a), or large-eddy simulation (LES) approaches, would likely be favorable for such a study. A second future direction involves further investigating the the link between local model performance and synoptic-scale meteorological conditions, extending the analysis presented in Section 3d. Such a study could aim to more robustly classify HRRR bias using a series of characteristic mesoscale regimes, similar to the characterization process performed in Banta et al. (2021). Such studies would allow for an improved understanding of the factors that modulate local HRRR performance, potentially leading to improved local predictions.

Acknowledgments. The authors would like to thank Kathryn Foster for support with data preparation and access, as well as three anonymous reviewers for suggestions that have considerably improved manuscript quality. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship (DOE CSGF) under award number DE-SC0024386. This research has also been supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Wind Energy Technologies Office. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

Data availability statement. Data used for this study are available through the U.S. Department of Energy website for the HilFlowS project at https://a2e.energy.gov/projects/wfip2.hilflows, as well as through the U.S. Department of Energy Data Archive and Portal (DAP; http://a2e.energy.gov/data). For additional information about the data used in this study, please refer to Wharton and Foster (2022).

References

Adler, B., J. M. Wilczak, L. Bianco, I. Djalalova, J. B. Duncan, and D. D. Turner, 2021: Observational case study of a persistent cold pool and gap flow in the Columbia River Basin. *Journal of Applied Meteorology and Climatology*, https://doi.org/10.1175/JAMC-D-21-0013.1.

Adler, B., J. M. Wilczak, J. Kenyon, L. Bianco, I. V. Djalalova, J. B. Olson, and D. D. Turner, 2023: Evaluation of a cloudy cold-air pool in the Columbia River basin in different versions of the High-Resolution Rapid Refresh (HRRR) model. *Geoscientific Model Development*, **16 (2)**, 597–619, https://doi.org/10.5194/gmd-16-597-2023.

Albornoz, C. P., M. E. Soberanis, V. R. Rivera, and M. Rivero, 2022: Review of atmospheric stability estimations for wind power applications. *Renewable and Sustainable Energy Reviews*, **163**, 112 505.

Arthur, R. S., T. W. Juliano, B. Adler, R. Krishnamurthy, J. K. Lundquist, B. Kosović, and P. A. Jiménez, 2022: Improved representation of horizontal variability and turbulence in mesoscale simulations of an extended cold-air pool event. *Journal of Applied Meteorology and Climatology*, https://doi.org/10.1175/JAMC-D-21-0138.1.

Banta, R. M., and Coauthors, 2021: Doppler-Lidar Evaluation of HRRR-Model Skill at Simulating Summertime Wind Regimes in the Columbia River Basin during WFIP2. *Weather and Forecasting*, https://doi.org/10.1175/WAF-D-21-0012.1.

Banta, R. M., and Coauthors, 2023: Measurements and model improvement: Insight into nwp model error using doppler lidar and other wfip2 measurement systems. *Monthly Weather Review*, **151 (12)**, 3063–3087.

Bauweraerts, P., and J. Meyers, 2019: On the feasibility of using large-eddy simulations for real-time turbulent-flow forecasting in the atmospheric boundary layer. *Boundary-Layer Meteorology*, **171**, 213–235.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, **144 (4)**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Bianco, L., and Coauthors, 2019: Impact of model improvements on 80 m wind speeds during the second wind forecast improvement project (wfip2). *Geoscientific Model Development*, **12 (11)**, 4803–4821.

Bianco, L., and Coauthors, 2022: Comparison of Observations and Predictions of Daytime Planetary-Boundary-Layer Heights and Surface Meteorological Variables in the Columbia River Gorge and Basin During the Second Wind Forecast Improvement Project. *Boundary-Layer Meteorology*, **182 (1)**, 147–172, https://doi.org/10.1007/s10546-021-00645-x.

Bingöl, F., J. Mann, and G. C. Larsen, 2010: Light detection and ranging measurements of wake dynamics part i: one-dimensional scanning. *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, **13 (1)**, 51–61.

Bortolotti, P., H. C. Tarres, K. Dykes, K. Merz, L. Sethuraman, D. Verelst, and F. Zahle, 2019: Iea wind task 37 on systems engineering in wind energy – wp2.1 reference wind turbines. Tech.

rep., NREL/TP-73492, International Energy Agency. URL https://www.nrel.gov/docs/fy19osti/73492.pdf.

Brewer, M. C., and C. F. Mass, 2016: Projected changes in western us large-scale summer synoptic circulations and variability in cmip5 models. *Journal of Climate*, **29 (16)**, 5965–5978.

Brewer, M. C., C. F. Mass, and B. E. Potter, 2012: The west coast thermal trough: Climatology and synoptic evolution. *Monthly Weather Review*, **140 (12)**, 3820–3843.

Carvalho, D., A. Rocha, M. Gómez-Gesteira, and C. Santos, 2012: A sensitivity study of the wrf model in wind simulation for an area of high wind energy. *Environmental Modelling & Software*, **33**, 23–34.

Castellani, F., D. Astolfi, M. Mana, M. Burlando, C. Meißner, and E. Piccioni, 2016: Wind power forecasting techniques in complex terrain: Ann vs. ann-cfd hybrid approach. *Journal of Physics: Conference Series*, IOP Publishing, Vol. 753, 082002.

Cheng, W. Y., Y. Liu, A. J. Bourgeois, Y. Wu, and S. E. Haupt, 2017: Short-term wind forecast of a data assimilation/weather forecasting system with wind turbine anemometer measurement assimilation. *Renewable Energy*, **107**, 340–351.

Christiansen, M. B., and C. B. Hasager, 2005: Wake effects of large offshore wind farms identified from satellite sar. *Remote Sensing of Environment*, **98 (2-3)**, 251–268.

Clifton, A., S. Barber, A. Stökl, H. Frank, and T. Karlsson, 2022: Research challenges and needs for the deployment of wind energy in hilly and mountainous regions. *Wind Energy Science*, **7 (6)**, 2231–2254.

Collins, E., Z. J. Lebo, R. Cox, C. Hammer, M. Brothers, B. Geerts, R. Capella, and S. McCorkle, 2024a: Forecasting High Wind Events in the HRRR Model over Wyoming and Colorado. Part I: Evaluation of Wind Speeds and Gusts. *Weather and Forecasting*, **39 (5)**, 705–723, https://doi.org/10.1175/WAF-D-23-0036.1.

Collins, E., Z. J. Lebo, R. Cox, C. Hammer, M. Brothers, B. Geerts, R. Capella, and S. McCorkle, 2024b: Forecasting High Wind Events in the HRRR Model over Wyoming and Colorado. Part II:

Sensitivity of Surface Wind Speeds to Model Resolution and Physics. *Weather and Forecasting*, **39 (5)**, 725–743, https://doi.org/10.1175/WAF-D-23-0037.1.

Coppin, P., E. F. Bradley, and J. Finnigan, 1994: Measurements of flow over an elongated ridge and its thermal stability dependence: the mean field. *Boundary-layer meteorology*, **69 (1-2)**, 173–199.

Danielson, J. J., and D. B. Gesch, 2011: Global multi-resolution terrain elevation data 2010 (gmted2010). Tech. rep., US Geological Survey.

Djalalova, I. V., and Coauthors, 2020: Wind Ramp Events Validation in NWP Forecast Models during the Second Wind Forecast Improvement Project (WFIP2) Using the Ramp Tool and Metric (RT&M). *Weather and Forecasting*, **35 (6)**, 2407–2421, https://doi.org/10.1175/WAF-D-20-0072.1.

Draxl, C., and Coauthors, 2021: Mountain waves can impact wind power generation. *Wind Energy Science*, **6 (1)**, 45–60.

Fitch, A. C., J. K. Lundquist, and J. B. Olson, 2013: Mesoscale influences of wind farms throughout a diurnal cycle. *Monthly Weather Review*, **141 (7)**, 2173–2198.

Giebel, G., and G. Kariniotakis, 2017: Wind power forecasting—a review of the state of the art. *Renewable energy forecasting*, 59–109.

Goutham, N., and Coauthors, 2021: Using machine-learning methods to improve surface wind speed from the outputs of a numerical weather prediction model. *Boundary-Layer Meteorology*, **179**, 133–161.

Heppelmann, T., A. Steiner, and S. Vogt, 2017: Application of numerical weather prediction in wind power forecasting: Assessment of the diurnal cycle. *Meteorol. Z. DOI*, **10**.

Hoen, B., J. E. Diffendorfer, J. Rand, L. A. Kramer, C. P. Garrity, A. D. Roper, and H. Hunt, 2018: United states wind turbine database. U.S. Geological Survey, URL https://www.sciencebase.gov/catalog/item/57bdfd8fe4b03fd6b7df5ff9, https://doi.org/10.5066/F7TX3DN0.

Hyvärinen, A., G. Lacagnina, and A. Segalini, 2018: A wind-tunnel study of the wake development behind wind turbines over sinusoidal hills. *Wind Energy*, **21 (8)**, 605–617.

James, E. P., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part II: Forecast Performance. *Weather and Forecasting*, **37 (8)**, 1397–1417, https://doi.org/10.1175/WAF-D-21-0130.1.

Juliano, T. W., B. Kosović, P. A. Jiménez, M. Eghdami, S. E. Haupt, and A. Martilli, 2022a: "gray zone" simulations using a three-dimensional planetary boundary layer parameterization in the weather research and forecasting model. *Monthly Weather Review*, **150 (7)**, 1585–1619.

Juliano, T. W., and Coauthors, 2022b: Smoke from 2020 united states wildfires responsible for substantial solar energy forecast errors. *Environmental Research Letters*, **17 (3)**, 034 010.

Kariniotakis, G., G. Stavrakakis, and E. Nogaret, 1996: Wind power forecasting using advanced neural networks models. *IEEE transactions on Energy conversion*, **11 (4)**, 762–767.

Li, J., S. Zhang, and Z. Yang, 2022: A wind power forecasting method based on optimized decomposition prediction and error correction. *Electric Power Systems Research*, **208**, 107 886.

Liu, M.-K., and M. A. Yocke, 1980: Siting of wind turbine generators in complex terrain. *Journal of Energy*, **4 (1)**, 10–16.

Liu, Y., Y. Qiao, S. Han, T. Tao, J. Yan, L. Li, G. Bekhbat, and E. Munkhtuya, 2021: Rotor equivalent wind speed calculation method based on equivalent power considering wind shear and tower shadow. *Renewable Energy*, **172**, 882–896.

Lubitz, W. D., and B. R. White, 2007: Wind-tunnel and field investigation of the effect of local wind direction on speed-up over hills. *Journal of Wind Engineering and Industrial Aerodynamics*, **95 (8)**, 639–661.

McClung, B., and C. F. Mass, 2020: The strong, dry winds of central and northern california: Climatology and synoptic evolution. *Weather and Forecasting*, **35 (5)**, 2163–2178.

Mesinger, F., and Coauthors, 2006: North american regional reanalysis. *Bulletin of the American Meteorological Society*, **87 (3)**, 343–360.

Mickle, R., N. Cook, A. Hoff, N. Jensen, J. Salmon, P. Taylor, G. Tetzlaff, and H. Teunissen, 1988: The askervein hill project: Vertical profiles of wind and turbulence. *Boundary-Layer Meteorology*, **43**, 143–169.

Mirocha, J., B. Kosovic, M. Aitken, and J. Lundquist, 2014: Implementation of a generalized actuator disk wind turbine model into the weather research and forecasting model for large-eddy simulation applications. *Journal of Renewable and Sustainable Energy*, **6 (1)**.

Olson, J. B., and Coauthors, 2019: Improving Wind Energy Forecasting through Numerical Weather Prediction Model Development. *Bulletin of the American Meteorological Society*, **100 (11)**, 2201–2220, https://doi.org/10.1175/BAMS-D-18-0040.1.

Peterson, E. W., and J. P. Hennessey Jr, 1978: On the use of power laws for estimates of wind power potential. *Journal of Applied Meteorology and Climatology*, **17 (3)**, 390–394.

Pichugina, Y. L., and Coauthors, 2019: Spatial Variability of Winds and HRRR–NCEP Model Error Statistics at Three Doppler-Lidar Sites in the Wind-Energy Generation Region of the Columbia River Basin. *Journal of Applied Meteorology and Climatology*, **58 (8)**, 1633–1656, https://doi.org/10.1175/JAMC-D-18-0244.1.

Pichugina, Y. L., and Coauthors, 2020: Evaluating the wfip2 updates to the hrrr model using scanning doppler lidar measurements in the complex terrain of the columbia river basin. *Journal of Renewable and Sustainable Energy*, **12 (4)**.

Platis, A., and Coauthors, 2018: First in situ evidence of wakes in the far field behind offshore wind farms. *Scientific reports*, **8 (1)**, 2163.

Quon, E. W., P. Doubrawa, J. Annoni, N. Hamilton, and M. J. Churchfield, 2019: Validation of wind power plant modeling approaches in complex terrain. *AIAA Scitech 2019 Forum*, 2085.

Safaei Pirooz, A. A., and R. G. Flay, 2018: Comparison of speed-up over hills derived from wind-tunnel experiments, wind-loading standards, and numerical modelling. *Boundary-Layer Meteorology*, **168**, 213–246.

Santoni, C., E. García-Cartagena, U. Ciri, G. V. Iungo, and S. Leonardi, 2018: Coupling of mesoscale weather research and forecasting model to a high fidelity large eddy simulation. *Journal of Physics: Conference Series*, IOP Publishing, Vol. 1037, 062010.

Sasser, C., M. Yu, and R. Delgado, 2022: Improvement of wind power prediction from meteorological characterization with machine learning models. *Renewable Energy*, **183**, 491–501.

Shaw, W. J., and Coauthors, 2019: The Second Wind Forecast Improvement Project (WFIP2): General Overview. *Bulletin of the American Meteorological Society*, **100 (9)**, 1687–1699, https://doi.org/10.1175/BAMS-D-18-0036.1.

Sideratos, G., and N. D. Hatziargyriou, 2007: An advanced statistical method for wind power forecasting. *IEEE Transactions on power systems*, **22 (1)**, 258–265.

Sisterson, D. L., and P. Frenzen, 1978: Nocturnal boundary-layer wind maxima and the problem of wind power assessment. *Environmental Science & Technology*, **12 (2)**, 218–221.

Tian, W., A. Ozbay, W. Yuan, P. Sarakar, H. Hu, and W. Yuan, 2013: An experimental study on the performances of wind turbines over complex terrain. *51st AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*, 7–10.

Tian, W., K. Zheng, and H. Hu, 2021: Investigation of the wake propagation behind wind turbines over hilly terrain with different slope gradients. *Journal of Wind Engineering and Industrial Aerodynamics*, **215**, 104 683.

Touma, J. S., 1977: Dependence of the wind profile power law on stability for various locations. *Journal of the Air Pollution Control Association*, **27 (9)**, 863–866.

Van Sark, W. G., H. C. Van der Velde, J. P. Coelingh, and W. A. Bierbooms, 2019: Do we really need rotor equivalent wind speed? *Wind Energy*, **22 (6)**, 745–763.

Wagenbrenner, N. S., J. M. Forthofer, B. K. Lamb, K. S. Shannon, and B. W. Butler, 2016: Downscaling surface wind predictions from numerical weather prediction models in complex terrain with windninja. *Atmospheric Chemistry and Physics*, **16 (8)**, 5229–5241.

Wagner, R., I. Antoniou, S. M. Pedersen, M. S. Courtney, and H. E. Jørgensen, 2009: The influence of the wind speed profile on wind turbine performance measurements. *Wind Energy*, **12 (4)**, 348–362, https://doi.org/10.1002/we.297.

Wagner, R., and Coauthors, 2014: Rotor equivalent wind speed for power curve measurement–comparative exercise for iea wind annex 32. *Journal of Physics: Conference Series*, IOP Publishing, Vol. 524, 012108.

Wharton, S., 2019: Wfip2 - hill flow study (hilflows) dataset. URL https://a2e.energy.gov/project/wfip2-hilflows.

Wharton, S., and K. Foster, 2022: Deploying Taller Turbines in Complex Terrain: A Hill Flow Study (HilFlowS) Perspective. *Energies*, **15 (7)**, 2672, https://doi.org/10.3390/en15072672.

Wharton, S., and J. K. Lundquist, 2012: Assessing atmospheric stability and its impacts on rotor-disk wind characteristics at an onshore wind farm. *Wind Energy*, **15 (4)**, 525–546.

Wharton, S., J. Newman, G. Qualley, and W. Miller, 2015: Measuring turbine inflow with vertically-profiling lidar in complex terrain. *Journal of Wind Engineering and Industrial Aerodynamics*, **142**, 217–231.

Wilczak, J., and Coauthors, 2015: The wind forecast improvement project (wfip): A public–private partnership addressing wind energy forecast needs. *Bulletin of the American Meteorological Society*, **96 (10)**, 1699–1718.

Wilczak, J. M., and Coauthors, 2019: The Second Wind Forecast Improvement Project (WFIP2): Observational Field Campaign. *Bulletin of the American Meteorological Society*, **100 (9)**, 1701–1723, https://doi.org/10.1175/BAMS-D-18-0035.1.

Wiser, R., and Coauthors, 2022: Land-based wind market report: 2022 edition. Tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Xia, G., C. Draxl, A. Raghavendra, and J. K. Lundquist, 2021: Validating simulated mountain wave impacts on hub-height wind speed using sodar observations. *Renewable Energy*, **163**, 2220–2230.