EVALUATING AND IMPROVING NWP FORECAST MODELS FOR THE FUTURE

How the Needs of Offshore Wind Energy Can Point the Way

Robert M. Banta, Yelena L. Pichugina, W. Alan Brewer, Eric P. James, Joseph B. Olson, Stanley G. Benjamin, Jacob R. Carley, Laura Bianco, Irina V. Djalalova, James M. Wilczak, R. Michael Hardesty, Joel Cline, and Melinda C. Marquis

Future needs of both the offshore wind energy industry and the NWP modeling community can be met with high-quality profile measurement arrays from new-generation remote sensors.

The lack of high-quality measurements above Earth's surface through the lower troposphere has been a serious impediment to progress in understanding, predicting, and modeling atmospheric processes and weather. Over coastal and oceanic environs, the situation is especially dire. This is because of the great difficulty and expense of deploying profile measurement systems of sufficient vertical resolution and accuracy over a water surface. The dearth of such profile measurements in the offshore environment has meant that offshore wind and weather systems are not well characterized and the skill of numerical weather prediction (NWP) models is poorly known at heights above the ocean surface.

The present study takes advantage of a dataset that exists from a monthlong summertime research cruise in the Gulf of Maine, as part of the 2004 New England Air Quality Study (NEAQS-04; Fehsenfeld et al. 2006; Pichugina et al. 2012). The Earth System Research Laboratory (ESRL)'s high-resolution Doppler lidar (HRDL) made high-quality wind profile measurements, in the form of 15-min-averaged wind profiles

AFFILIATIONS: Banta, Brewer, Benjamin, Wilczak, and Marquis—National Oceanic and Atmospheric Administration/Earth System Research Laboratory, Boulder, Colorado; Pichugina, James, Olson, Bianco, Djalalova, and Hardesty—Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and National Oceanic and Atmospheric Administration/Earth System Research Laboratory, Boulder, Colorado; Carley—I.M. Systems Group, Inc., and National Oceanic and Atmospheric Administration/National Weather Service/Environmental Modeling Center, College Park, Maryland; Cline—Office of Energy Efficiency and Renewable Energy, Department of Energy, Washington, D.C.
CORRESPONDING AUTHOR: Robert M. Banta, robert.banta@noaa.gov

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-16-0310.1

In final form 15 December 2017

at 10-m vertical resolution, with precisions better than 0.1 m s^{-1}. HRDL is a scanning, pulsed Doppler lidar (Grund et al. 2001) equipped for ship-motion compensation, which operated from the deck of the National Oceanic and Atmospheric Administration (NOAA) Research Vessel *Ronald H. Brown* (*RHB*) during NEAQS-04. The HRDL profiles provide the quality of measurements required to advance understanding and modeling of offshore weather systems and winds (Banta et al. 2013).

Concentrated populations and resulting human activities along the coastal areas of the world demand access to the most accurate environmental information and predictions (National Research Council 1992). Such activities include the recent development of offshore wind energy (WE; Musial and Ram 2010; Archer et al. 2014). Measurements and model validation, relative to offshore wind energy, is a focus of this paper. Obtaining accurate environmental information often involves forecasting and modeling weather and atmospheric properties (winds, clouds, temperature, pollutants, moisture, and precipitation) in the offshore marine atmospheric boundary layer (MABL; e.g., Kraus and Businger 1994). Frequent and strong horizontal cross-shore gradients spawn many different types of hard-to-forecast weather systems, which constitute the meteorological challenges of the offshore coastal zone (National Research Council 1992). Coastal weather and wind systems occur on multiple time and space scales. Local atmospheric processes form small-scale nearshore weather phenomena that interact with synoptic-scale and larger-mesoscale weather systems, with the diurnal cycle of surface heating and cooling, with shoreline topography, and with the land–water and air–sea interfaces to produce complex flows at the coast. These influences may extend several hundred kilometers offshore. Characterizing, modeling, and predicting these interactions are key to more accurate atmospheric information and forecast products.

In this ever-changing meteorological environment, the emerging U.S. offshore wind energy industry requires many forms of information on winds aloft. For WE applications, highly accurate wind information is needed above the surface at the level of wind turbine rotor blades, generally 50–200 m above mean sea level (MSL), where the behavior of wind is often different from that within 10 m of the surface (Banta et al. 2013; Pichugina et al. 2017b), with desired uncertainties of 0.1 m s^{-1}. As is the case for land-based WE, operations and maintenance scheduling need accurate wind forecasts (Ahlstrom et al. 2013; Marquis et al. 2011; Wilczak et al. 2015; Shaw et al. 2009). Resource assessment—that is, determining where to site an offshore wind plant—requires accurate maps of the horizontal distribution of the rotor-layer wind, and turbine and wind-farm design demands accurate assessments of directional and speed wind shear across the rotor layer.

NWP models could provide this information—avoiding the need for expensive overwater field measurement campaigns—if the models could prove sufficiently accurate and reliable. Thus, model accuracy at levels above the surface needs to be evaluated. If models prove not to be accurate enough, then rotor-level wind properties will have to be obtained by extended measurement projects, and simultaneously, work must proceed to improve the models. We argue that these two endeavors overlap significantly: the kind of field measurement campaign datasets required by WE are the same kinds of datasets needed to improve the state of the art in NWP modeling, namely, horizontal arrays of high-quality profiling instrumentation (Banta et al. 2013).

To facilitate the planning and development of offshore WE and to address the scarcity of wind information aloft, the Energy Efficiency and Renewable Energy (EERE) Office of the U.S. Department of Energy (DOE) asked NOAA to make use of the existing NEAQS-04 dataset along with NOAA's modeling resources to address these issues. This project, referred to as the Positioning of Offshore Wind Energy Resources (POWER) project, was designed to address three goals: to analyze the measured vertical and horizontal properties of the offshore wind field over the Gulf of Maine; to perform retrospective model runs of two selected weeklong periods during this campaign using two model versions, each hourly updated trial versions of NOAA NWP modeling systems; and to use the reference profile measurements on board the *RHB* to quantitatively evaluate the skill of these models at heights above the surface, including at the turbine rotor level. The trial models employed were versions of the North American Mesoscale Forecast System (NAM) and the Rapid Refresh (RAP) models plus embedded or nested configurations of each (see "The RAP–HRRR and NAM models" sidebar). They are described by James et al. (2018), Djalalova et al. (2016), and Pichugina et al. (2017b). For this study, these models performed special retrospective runs that assimilated available routine measurements, such as conventional surface observations [aviation routine weather reports (METARs)], rawinsondes, and onboard commercial aircraft observations. Model versus lidar-measured wind error statistics were calculated, which can serve as benchmarks for the skill of the 2012 versions of the models assimilating the routine measurements available in 2004.

Previous studies have used models to estimate the offshore wind resource, generally using surface wind data from buoys, shoreline sites, or other sources (e.g., ships, satellites) as validation for the model-derived wind speed maps (e.g., Schwartz et al. 2010; Musial and Ram 2010; Jimenez et al. 2007; Jiang et al. 2008; Dvorak et al. 2010, 2012; Giannakopoulou and Nhili 2014). A few studies have used offshore-tower datasets to calculate model error statistics from measurements aloft. Drechsel et al. (2012), for example, used instrumented-tower data (up to 70 and 100 m MSL) on two platforms in the North Sea off the coasts of Denmark and Germany [the *Forschungsplattformen in Nord- und Ostsee 1 (FINO-1)* tower) to evaluate model errors over a period of 1 year (results of which will be presented later). Krogsæter and Reuder (2015) also used profile data from the FINO-1 tower to validate model wind values, and Draxl et al. (2014) used data from a coastal tower in western Denmark, 1.7 km inland, for model validation. Although this site was near but not over the ocean, the error values were similar to those we present later.

A few overwater research datasets have been used to evaluate NWP models in different ways. Colle et al.

THE RAP–HRRR AND NAM MODELS

The RAP–HRRR and NAM are NWP models that run operationally (24 hours per day, 7 days per week) at NCEP. The NAM provides four forecasts per day over an 84-h forecast period. It is currently NCEP's regional, multiday continental forecast system. The operational NAM also features five additional, comparatively higher-resolution nest domains. Most notable for POWER is the 4-km nest domain that covers the CONUS, having a forecast period of 60 h. The trial NAM system was the hourly updated, 12-km NAMRR over North America with its fine-mesh (4 km) NAMRR-CONUS-nest concentrated over the contiguous United States (Fig. 3a). This NAMRR-CONUS-nest domain was one-way nested into the 12-km NAMRR, also referred to as the NAMRR-parent. Both assimilate available measurements hourly (similar to the RAP and HRRR), with a procedure described in Banta et al. (2014) and Djalalova et al. (2016). To perform the model validation, model profile values for the "nearest neighbor" model grid point to the ship position were used to compare with the HRDL-measured profile data. For the ocean surface relevant to this study, sea surface temperatures (SSTs) were updated once daily at 1800 UTC based upon the most recent real-time global 0.5° SST analysis created by the NCEP/Environmental Modeling Center Marine Modeling and Analysis Branch (Thiébaux et al. 2003). At the same time, snow cover and sea ice were updated from the most recent data available from the National Ice Center's Interactive Multisensor Snow and Ice Mapping System. For completeness, the Northern Hemisphere snow depth was also updated based upon analyses provided by Air Force Weather Agency (now U.S. Air Force 557th Weather Wing). Other relevant features of the NAM models are given in Table 2.

The RAP (Benjamin et al. 2016) was developed to provide short-range hourly updated forecasts out to 18 h for applications such as aviation that need frequently updated information. Model skill at these short lead times depends highly on initial conditions with focus on effective assimilation of real-time, frequent datasets such as onboard aircraft, satellite, and weather radar. A particular concern for the RAP was predicting warm-season deep, moist convection. Therefore, the RAP system includes a smaller-domain, finer-resolution nested model—the HRRR—run on a 3-km mesh where this convection is explicitly simulated rather than parameterized. The features, rationale, and history of RAP and HRRR development have recently been detailed by Benjamin et al. (2016). The 2004 retrospective study described here used the earliest version (version 1) of RAP–HRRR (see Table 2 in Benjamin et al. 2016) without ensemble data assimilation and for which no radar data were available. This RAP version used for the POWER simulations, the RAP–HRRR-2012P, was performed over a much-reduced domain (Fig. 3b) because of computer time and data storage restrictions. For model validation, RAP–HRRR-simulated profile values were interpolated horizontally to the position of the ship using a parabolic interpolation scheme. Over the ocean surface, the RAP and HRRR systems (both real-time and retrospective POWER versions) do not have a coupled ocean model. Sea ice coverage and SSTs are specified from daily updated analyses, as described by Benjamin et al. (2016, their section 4). Sea ice is based on a satellite-based National Environmental Satellite, Data, and Information Service (NESDIS) dataset, and SST is based upon the NCEP global high-resolution SST analysis.

Increasingly upgraded versions of the RAP and HRRR models were run experimentally in near–real time starting in 2010, and archived HRRR model output fields were used for subsequent wind and solar climatology studies (James et al. 2017, 2018). This archive was used to generate rotor-layer wind speed composite maps covering the U.S. East Coast over a 3-yr period [2013–15; see Fig. 4 here, reproduced from James et al. (2018)]. The embedded HRRR was one-way nested with lateral boundary conditions and model fields simply interpolated to the HRRR grid from the parent RAP. Compared to the 2012 POWER retrospective experiment configurations, the NCEP HRRR (not implemented until 2014) uses a 1-h spinup period at 3-km resolution combined with ensemble hybrid data assimilation. Relevant features of the RAP–HRRR-2012P models for the POWER experiments are given in Table 2.

(2016) looked at model output from individual case study periods, which they compared with profile or aircraft measurement analyses. Darby et al. (2002b) used long-range Doppler lidar measurements taken from the shore of Monterey Bay, California, to give confidence to idealized model findings. In this case, they studied the sensitivity of vertical flow structure at the shoreline to the presence of the coastal ranges and the Sierra Nevada. The NEAQS dataset has been used in other POWER studies. Pichugina et al. (2017a) studied properties of the offshore low-level jet (LLJ) in the Gulf of Maine and also determined power-law wind profile exponents α and their horizontal distribution. James et al. (2018) performed High-Resolution Rapid Refresh (HRRR) modeling studies to evaluate the long-term (3 years) mean distribution of hub-height winds off the U.S. East Coast.

Also as part of the POWER project, Djalalova et al. (2016) and Pichugina et al. (2017b) investigated the impacts on 0–12-h forecast skill over the ocean, resulting from assimilating 11 land-based, 915-MHz wind-profiling radars (profilers), deployed as a part of the original campaign. The results, in brief, were that assimilating the profiler data produced an improved initialization by ~0.2 m s^{-1} and forecast improvement by as much as 5%–10% early (generally in the first 2 h), with positive improvement indicated out to 3–4 h, when validated against the independent offshore profile data (HRDL and a profiler aboard the *RHB* that were not assimilated into the model runs). These studies also performed model evaluations for a rainy period in the middle of one of the weeks chosen for study.

Table 1. Technical characteristics of NOAA/ESRL Doppler lidar.

Characteristic	HRDL
Wavelength	2.02 μm
Pulse energy	2.0 mJ
Pulse rate	<300 Hz
Range resolution	60 m
Velocity resolution	~0.5 cm s^{-1}
Time resolution	0.3 s
Minimum range	90 m
Maximum range	5–12 km
Scanner height	8 m MSL

Because the *RHB* traversed the Gulf of Maine repeatedly during NEAQS-04, we were able to document the horizontal variability of the winds for the time period and locale of the cruise. Based on the observed variability, we discuss considerations in deploying a network of offshore wind-profiling devices and provide recommendations for possible coastal offshore sampling-array layouts. Proposed offshore profiling devices could include floating-buoy lidar systems currently being built and tested. By using profiling arrays of this kind, we could test the ability of NWP models to characterize the vertical structure of the flow and its horizontal variability.

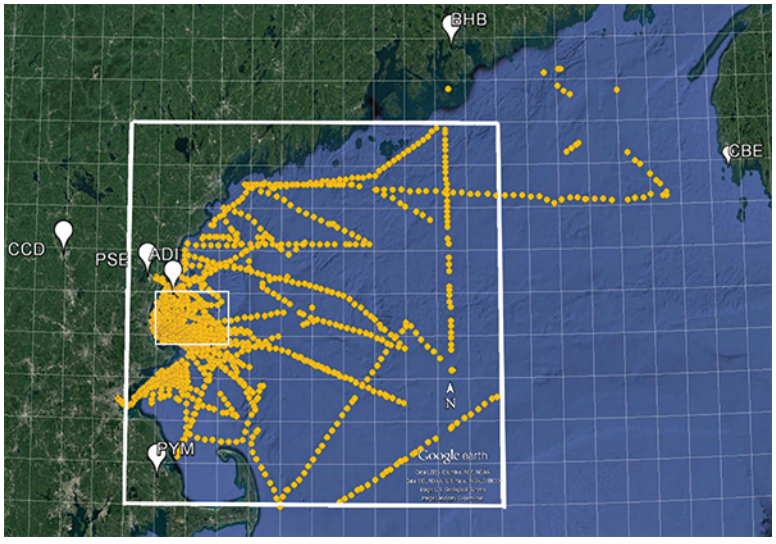

Fig. 1. Gulf of Maine region (background from Google Earth). Ship tracks for entire monthlong campaign are traced by yellow dots, which designate the locations of the 15-min wind profiles used in this paper. The larger white box outlines the region analyzed in this paper, and a 1° latitude–longitude grid is superimposed. Profiler sites identified are Appledore Island, Maine (ADI); Bar Harbor, Maine (BHB); Chebogue Point, Nova Scotia, Canada (CBE); Concord, New Hampshire (CCD); Pease Tradeport, New Hampshire (PSE); and Plymouth, Massachusetts (PYM).

BACKGROUND, MEASUREMENTS, AND ANALYSIS PROCEDURES.

The offshore, coastal, and inland NEAQS-04 datasets used in the present study were originally obtained to characterize local pollution sources in the New England region (Fehsenfeld et al. 2006; Angevine et al. 2006; Darby et al. 2007; White et al. 2007), so they are very useful but not optimized for offshore WE or NWP improvement studies. Land-based, airborne, and shipborne instrumentation all contributed to the dataset (Fehsenfeld et al. 2006). The major offshore measurement platform

was the *RHB*, which took meteorological, air chemistry, and oceanographic data over the Gulf of Maine from 9 July to 12 August 2004 (see Fig. 1 for ship tracks). Instrumentation on board pertinent to the present study included HRDL, a 915-MHz profiler (Wolfe et al. 2007), and NOAA's in situ surface-flux measurement package, mounted on a boom protruding out from the bow (Fairall et al. 2006). The focus of this study is the use of HRDL.

The two intensive study and numerical simulation periods were chosen based on HRDL data, which are available except during rainy or foggy conditions. The two case study periods were mostly during dry conditions, which were hard to find because the summer of 2004 was abnormally cool, cloudy, and wet for the region (White et al. 2007). The first study period selected was 6–12 August, corresponding to the only 1-week-long lull between frontal passages (White et al. 2007). For the second study period (10–17 July), two of the shorter periods separated by a day and a half of rain (1600 UTC 13 July–1400 UTC 15 July) were selected so that the model runs could be performed through the week without having to restart the models. All models were run for both study periods.

The high-resolution Doppler lidar. HRDL is a scanning, coherent, pulsed Doppler lidar designed, built, and operated by NOAA/ESRL for atmospheric boundary layer research (Grund et al. 2001; see Table 1). Deployed on board the *RHB*, HRDL was operated over the Gulf of Maine 24 h day^{-1} during the NEAQS-04 field campaign, for a total of 28 days of data (Pichugina et al. 2012, 2017a,b). This lidar employs a motion-compensation system to provide precise wind profile measurements at high vertical resolution from the deck of the *RHB* (Pichugina et al. 2012). To address the objectives of the POWER study, HRDL scan data were analyzed to calculate mean profiles of wind speed and direction from the deck of the *RHB* at 15-min intervals, which was the repeat time of the scan sequences. Conclusions drawn in this paper based on the 15-min averages would apply equally to 10-min averaging commonly used in WE.

The HRDL scan sequences consisted of azimuth (conical) scans at three elevation angles (Fig. 2), elevation (vertical slice) scans at two perpendicular

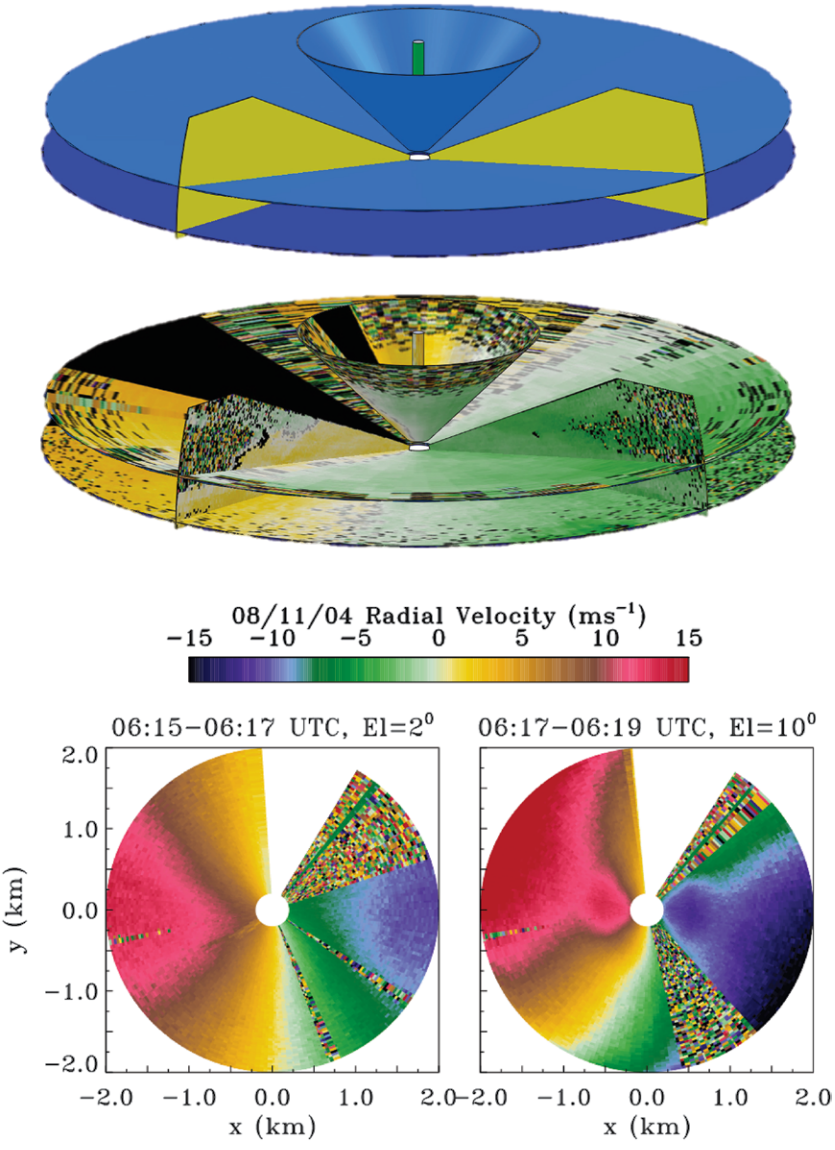

Fig. 2. Conical scan examples from 11 Aug 2004. (Top) schematic depiction of scan pattern during NEAQS: three nested conical (azimuth) scans indicated in blue, two vertical-slice or range–height (elevation) scans in yellow, and vertically staring in green. (middle) Geometry as in the top panel, but showing actual scan data on a scan image. (Bottom) Two examples of full 360° azimuth (conical) scans at fixed elevations of (left) 2° and (right) 10°.

azimuths (oriented to avoid the ship superstructure), and vertical staring for vertical velocity measurements. For this study, we used only the three nested conical scans, which were taken mostly at 2°, 10°, and 30° elevation. The conical scans at the higher elevations reach higher into the atmosphere than lower-elevation scans but have coarser resolution in the vertical, so we form a single wind profile from the three conical scans by collecting the velocity data into vertical bins having vertical intervals that increase gradually with height, as described by Banta et al. (2015). Typical vertical analysis bin intervals are 9 m near the surface, 20 m at 100 m MSL, and 45 m at 1,000 m MSL. The calculated profiles thus represent a 15-min mean, averaged vertically within the bin and horizontally around the lidar's velocity–azimuth display (VAD) sampling ring and along the ship's track, extending from near the ocean surface to heights of ~2 km or more. The atmosphere over the ocean in general varies slowly enough over 15 min and smoothly enough over the spatial-averaging dimensions involved to regard these as representative mean profiles over these temporal and vertical scales.

Many of the NEAQS-04 measurements were assimilated into the experimental model runs, but the HRDL measurements were withheld, so they constitute an independent verification dataset. For comparison with model profiles, which were extracted at the top of each hour, HRDL data were averaged for 1 h centered on the top of the hour.

NWP forecast models used. This study used special trial versions of two models that run operationally at NOAA/National Centers for Environmental Prediction (NCEP), the NAM and RAP, covering a North American domain at horizontal grid spacings of 12 and 13 km, respectively. Both models feature finer-mesh models embedded or nested within, as described in "The RAP–HRRR and NAM Models" sidebar.

The trial NAM system (Table 2) consisted of an hourly updated, 12-km model, the NAM Rapid Refresh (NAMRR), over North America, run on

Table 2. Domain configurations for NWP forecast models.

Model	Configuration
12-km NAMRR-parent	
Points in *x*, *y*, and *z* directions	954, 835, and 60
Microphysics parameterization	Ferrier et al. (2002, 2011)
Boundary layer parameterization	Janjić (2001)
Convective parameterization	Janjić (1994)
Longwave and shortwave radiation parameterization	Iacono et al. (2008) and Mlawer et al. (1997)
Land surface model	Ek et al. (2003)
Gravity wave drag parameterization	Alpert (2004)
4-km CONUS-nest	
Points in *x*, *y*, and *z* directions	1,371, 1,100, and 60
Convective parameterization	Janjić (1994), modified to be less active for higher resolution
Gravity wave drag parameterization	None
13-km RAP-2012P, CONUS domain reduced from whole domain	
Points in *x*, *y*, and *z* directions	Reduced to 758, 567, and 51
Cloud microphysics parameterization	Thompson et al. (2008)
Boundary layer parameterization	Janjić (2001)
Convective parameterization	Grell 3D/Grell shallow cumulus
Longwave and shortwave radiation parameterization	Chou and Suarez (1994) and Mlawer et al. (1997)
Land surface model	Smirnova et al. (1997, 2000)
3-km HRRR-2012P, reduced from the whole domain	
Points in *x*, *y*, and *z* directions	Reduced to 520, 450, and 51
Convective parameterization	Turned off

the operational domain (Fig. 3a). Nested within the NAMRR is a fine-mesh (4 km) model concentrated over the conterminous United States (CONUS), the NAMRR-CONUS-nest (blue box in Fig. 3a). The trial RAP version, the RAP-2012P (see Table 2), ran on the reduced domains shown in Fig. 3b. The fine-mesh version is the 3-km HRRR, and the domain of the POWER version, the HRRR-2012P, is the blue box in Fig. 3b.

In the wind resource assessment process, the horizontal distribution of turbine-level wind speeds is of critical importance for the siting of wind plants and wind turbines. NWP models have been used to produce maps showing the horizontal variability of the offshore wind resource (e.g., Schwartz et al. 2010; Musial and Ram 2010; Jimenez et al. 2007; Jiang et al. 2008; Dvorak et al. 2012). As a part of POWER, a multiyear archive of trial HRRR runs made at ESRL was used to calculate composite maps of 80-m wind speeds averaged over a 2-yr period (Banta et al. 2014) and then a 3-yr period (James et al. 2018) over the CONUS. The values going into this map were an hourly dataset of 1-h lead-time forecast values—close to the initial times but after the model initialization transients had time to subside, as explained by James et al. (2017). The 3-yr composite, 80-m wind speed map for the Gulf of Maine (Fig. 4) shows that wind speeds increase significantly with distance from the coast, a characteristic that makes offshore WE attractive.

This project provided an opportunity to evaluate the skill of these models in an offshore setting by exploiting the rare availability of high-quality wind measurements aloft. To accomplish the model–lidar comparisons, the gridded model wind profile values were determined for the location of the ship as described in the "The RAP–HRRR and NAM Models" sidebar. These values were then linearly interpolated to the heights of the lidar measurements, where the differences between model and lidar measurements at each vertical level were used to calculate model error statistics, such as root-mean-square (rms) error or mean bias. The error values could then be further averaged over deeper vertical layers. Pichugina et al. (2017b) have shown that the magnitude of this vertically averaged error is sensitive to the depth of the averaging layer, where averages over deeper layers have larger sample sizes and give smaller errors.[1] Figure 5, reproduced from that paper, quantifies this effect for the August study period: for short lead times up to 3–4 h, averaging

[1] It will also be shown later that the error values decreased aloft, so the deeper averaging layers also benefited from including vertical levels where the model errors were smaller.

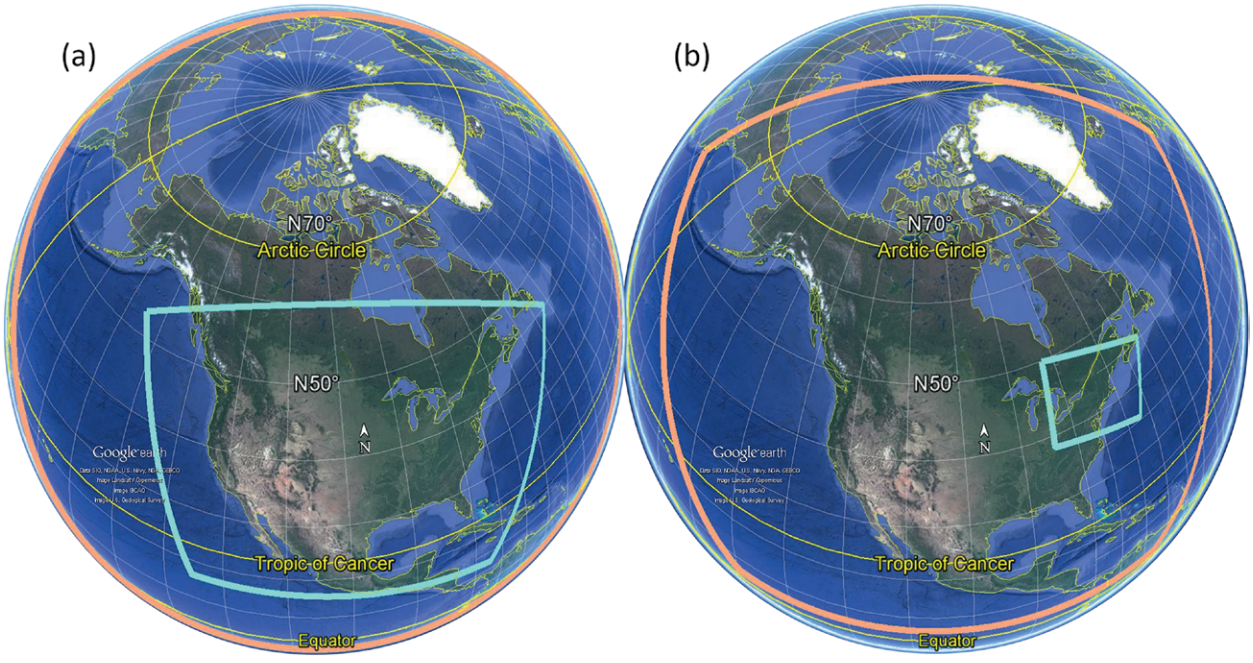

Fig. 3. (a) Domains for NAMRR-parent (orange, covering most of image) and NAMRR-CONUS-nest (light blue), as used both in POWER and for real-time operational forecast runs, superimposed on a Google Earth map image. (b) Reduced-size trial domains for RAP-2012P (orange) and HRRR-2012P (light blue) models for POWER retrospective experiments. From Pichugina et al. (2017), copyright American Meteorological Society, used with permission.

over a 500-m-deep layer produces an error of 2.0 m s^{-1}, whereas the wind speed error at the 100-m hub-height level, representing a ~20-m-deep layer average about that height, is 2.5 m s^{-1}, or 25% larger. In calculating model errors appropriate for WE applications, it is thus important to consider a shallow-enough layer of the atmosphere at the right heights for this industry.

RESULTS. *HRDL–NWP comparisons.* In this section, we validate wind forecasts from the models against HRDL measurements. For context, we first present time–height cross sections of wind speed for a sample day, which give an overview of the time dependence of the vertical structure of the flow. Figure 6 shows an example for 11 August, with hourly HRDL data (Fig. 6a) and interpolated model values for the NAMRR-parent (Fig. 6c) and NAMRR-CONUS-nest (Fig. 6b). RAP and HRRR results (not shown) were similar. LLJ structure—higher wind-speed layers having lower wind speeds above and below—is seen for many of the hours. The overall patterns between measurements and models are similar, indicating that the models simulated the existence of the LLJ and other structures. Quantitative details of the vertical structure at the beginning and end of the period, however, as well as the penetration of strong winds below 200 m between 0700 and 0900 UTC, were not well modeled.

On the same day, time series of HRDL-measured wind speeds at seven vertical levels from 10 to 500 m (Fig. 7), including three levels within the hypothetical turbine rotor layer (50, 100, and 150 m), show how the flows and their evolution at various vertical levels are related—or not. For the first 8 h, the near-surface flow at 10 m MSL (dark blue line) lies mostly between 7 and 11 m s^{-1}, whereas in the rotor layer, the wind speeds at 50, 100, and 150 m (red, black, and lilac curves), for example, range from 17 m s^{-1} down to 7 m s^{-1} (the 50- and 100-m curves intersecting the 10-m wind speed line just before 0500 UTC) then back up to nearly 15 m s^{-1}. These increases and decreases in wind speed, or *ramp* events, are of great consequence to wind energy and thus are very important to be able

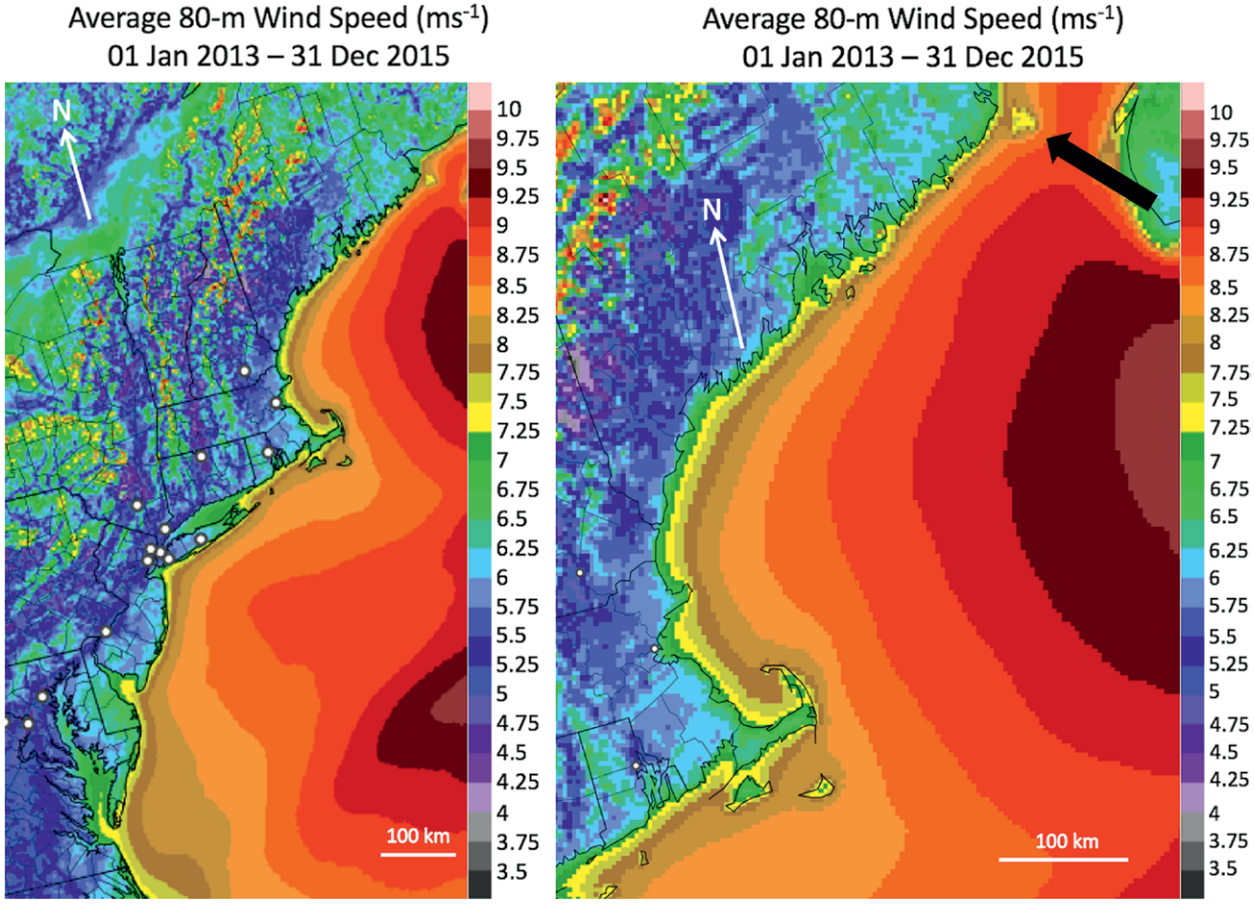

Fig. 4. Average 80-m wind speed from 1-h HRRR forecasts over the 2013–15 period, showing (left) the northeastern coast of the United States and (right) the Gulf of Maine (James et al. 2018). Thick arrow in the upper-right corner of the right panel indicates a region of slower offshore wind speeds associated with the island of Grand Manan, New Brunswick, Canada.

to predict accurately (Ahlstrom et al. 2013; Marquis et al. 2011). The changes in wind speed at 100 m relative to those at 10 m illustrate how differently wind speed at 10 m can behave in time from the speeds at higher levels, even within the rotor layer, consistent with other studies over the ocean (Pichugina et al. 2012, 2017a) and land (Banta et al. 2013).

Individual hourly profile comparisons in Fig. 8 also from 11 August show HRDL-measured wind data during one of the LLJ periods along with NAMRR and NAMRR-CONUS-nest values, for the model initial conditions (forecast 0; top row), and for the 2-h forecast (forecast 2; bottom row) valid at each hour. For this example, the finer-mesh model values better

EFFECTS OF FINER GRID SPACING ON MODEL SKILL

Some previous studies have found decreases in NWP model skill as model horizontal grid spacing decreases below 10 km (e.g., Mass et al. 2002; Mittermaier 2014). The reasons for this are unknown, but speculations can be grouped as validation issues, appropriateness of subresolution-scale (SRS) parameterizations, and numerical. Subresolution scale refers to the fact that models having a horizontal grid interval of Δx do not properly resolve the behavior of atmospheric features smaller than $4–5\Delta x$, sometimes larger.

An example of validation issues is that significant finescale structure may exist in the atmosphere at length scales corresponding to grid spacings less than 10 km. The predicted structure may actually be captured reasonably well overall by the models, but relatively small horizontal displacements of the model fields relative to the atmosphere could produce significant calculated model versus atmosphere differences, which would be interpreted as large errors. Errors in timing of otherwise well-predicted events could also produce a similar inflation of calculated errors (Mass et al. 2002).

SRS flux parameterizations, as model resolutions decrease from coarser to finer, pass from a "mesoscale" or larger-scale regime to a small-scale regime. In the larger-scale regime, model grid spacing is much larger than the turbulence length scales (turbulence scale) that contain the turbulence kinetic energy. In the smaller-scale regime, the grid intervals are much smaller than the turbulence scale, such that the energy-containing eddies are resolved—the large-eddy simulation (LES) and direct numerical simulation (DNS) approaches. For the mesoscale, grid intervals are larger than ~10 km, and those for LES are smaller than 50–100 m. Without going into great detail here, Wyngaard (2004) argues that the parameterization methodology in the mesoscale case can be tuned to perform reasonably well for mesoscale and larger-scale dynamics and that, in the LES case, the role of SRS transport processes is small, so they can be modeled as diffusive without significant error. These parameterizations thus each seem to work acceptably at the scales for which they were intended. However, at intermediate resolutions, dubbed "terra incognita" by Wyngaard (2004), the energy-containing turbulence scales are partially resolved, and in this regime, neither parameterization approach may be appropriate. Thus, as the grid size decreases into these terra incognita scales, the inappropriateness of the mesoscale SRS flux parameterizations may lead to increased model errors versus the atmosphere.

Another possibility is a numerical analog to the terra incognita of model parameterizations. Successful finite-difference modeling relies on some key assumptions, one of which is that the derivatives in the primitive equations can be represented by first-order terms in their Taylor series expansions (or at least low order). This requires that the prognostic variables vary smoothly in time and space and that gradients, curvatures, and higher-order differences of these variables are not large. Numerical and explicitly modeled diffusion ensure that model variable fields are smooth enough to prevent numerical instability. At larger scales (perhaps grid intervals >10 km or so), the success demonstrated by NWP forecast models implies that the smoothed fields reasonably well represent the dynamics at those scales, seemingly justifying the neglect of higher-order terms. At LES scales, finescale SRS features exist in the atmosphere, but by definition, they should have little impact on the overall simulation and should not be a significant source of model errors (i.e., departure in time of modeled fields from those in the atmosphere). For simulations having intermediate mesh sizes, perhaps 0.5–5-km horizontal grid spacing, however, the wind fields in the atmosphere at equivalently resolved scales are often not smooth, and large gradients can exist. This is known to several authors of this paper through decades of experience with Doppler lidar datasets and several dozen studies of flows having strongly varying atmospheric structure on scales up to 30 km (e.g., Angevine et al. 2003; Banta et al. 1993, 1996, 1997, 2004, 2005, 2012; Clark et al. 1994; Darby et al. 1999, 2002a,b, 2006, 2007; Flamant et al. 2002; Gohm et al. 2010; Langford et al. 2010; Levinson and Banta 1995; Rothermel et al. 1998a,b; Rucker et al. 2008; Weissmann et al. 2004). Atmospheric dynamics at subresolution scales may be dominated by these kinds of features, which are often not turbulence, often not diffusive, and often unrelated to local resolved gradients. Furthermore, higher-order terms of such fields that would be calculated for the real atmosphere are likely to be too large to neglect. In this case, processes in the atmosphere that would be subresolution in a model of this grid size may be responsible for SRS transfer that is large compared with the other terms in the governing equations. In the model these processes would at least be smoothed and much less effective, and the SRS terms could even be of the wrong sign. As a result, the model solutions would diverge in time from the evolution of the true fields in the atmosphere.

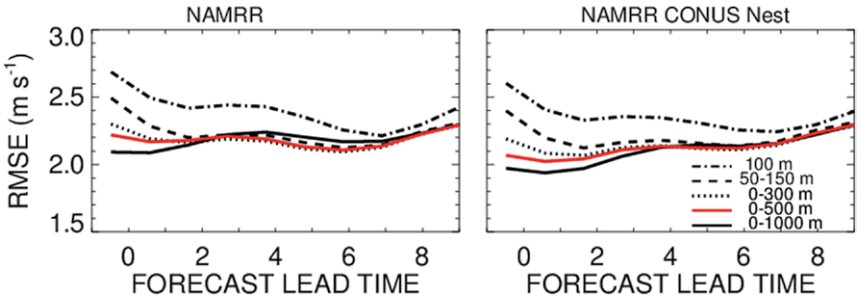

Fig. 5. Model rms error (difference from HRDL-measured values) vs forecast lead time for averaging over different vertical layers (dash–dotted: 100-m level; dashed: 50–150-m layer; dotted: 0–300-m layer; red solid: 0–500-m layer; and black solid: 0–1,000-m layer). (left) NAMRR-parent and (right) NAM-CONUS-nest [from Pichugina et al. (2017b), copyright American Meteorological Society, used with permission].

approximate the measured data in general, especially below 800 m. The model profiles captured the overall LLJ occurrence, but some aspects of the vertical structure and its evolution were not well represented at low levels, as, for example, the very shallow jet below 100 m at 0400–0500 UTC and the acceleration of the flow below 400 m at 0600 UTC, an hour after local midnight. Both of these discrepancies, often 3 m s^{-1} to as much as 6 m s^{-1} in the rotor layer, are within the height and wind speed ranges that would produce large errors in wind power calculations.

High-quality wind profile measurements make it possible to evaluate statistically the skill of model forecasts as a function of height for extended time periods. Figures 9a and 9d show profiles of the model rms wind speed error calculated for the initial conditions and 2-h forecasts over the August study period for each of the four model versions: RAP–HRRR (RAP-2012P and HRRR-2012P in this study) in the top row, and the two NAM configurations in the bottom row. Near 500 m, all models indicate errors of 2 m s^{-1}. For the RAP models, this decreases downward to 1.5 m s^{-1} in the 100–200-m layer, but in the NAM models, the error stays near 2 m s^{-1}. All models show a significant degradation in skill below 100 m MSL, with the rms error jumping up by nearly 0.5 m s^{-1}, which amounts to near-surface increases of 28% in the RAP error, 53% in the HRRR, 32% in the NAM, and 32% in the NAM-CONUS. For the HRRR-2012P, this relatively poor performance of the initial values compared to the others can be attributed to the use of the simple-interpolation initialization (see "The RAP-HRRR and NAM Models" sidebar) rather than the full assimilation of available hourly measurements on the fine-mesh HRRR grid, as used in the operational RAP–HRRR (Benjamin et al. 2016). Note that these large initial HRRR-2012P error values have adjusted after 2 h of simulated time, decreasing to RAP values. Overall, the errors in the fine-mesh models were the same as or less than those of their parent versions except for this HRRR-2012P initialization imbalance. Errors in the 2-h forecasts were less than those of the initial model state, confirming that the mutual adjustment of the model fields to each other alleviated initialization errors early in the simulations. Figure 9 also shows the coefficient of determination R^2 (Figs. 9b,e) and bias profiles (Figs. 9c,f). As with the rms error, these statistics indicate a degradation of skill in the lowest 100 m MSL—smaller correlations and larger biases. Thus, the statistical profiles indicate that the forecast models

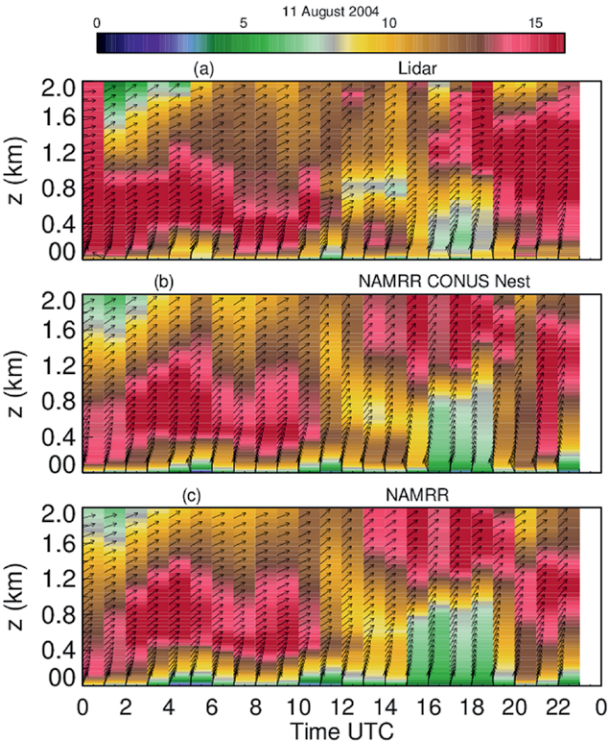

Fig. 6. Time–height cross sections of wind speed up to 2 km MSL on 11 Aug 2004 for (a) HRDL-measured profiles interpolated to form hourly values, (middle) NAMRR-CONUS-nest, and (bottom) NAMRR. Model values indicated for each hour were extracted from the time step at the beginning of each hour. Local standard time (LST) is 5 h behind UTC (0000 LST = 0500 UTC and 1200 LST = 1700 UTC).

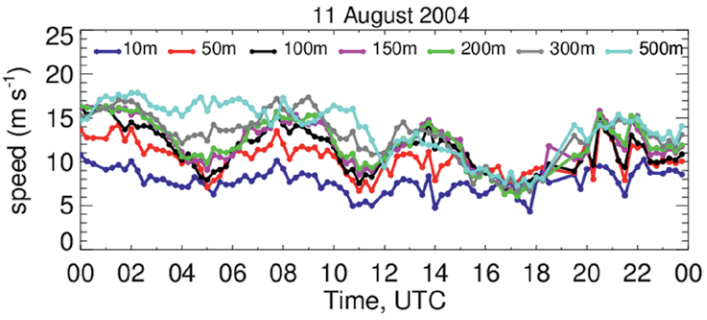

Fig. 7. Time series of wind speed from lidar measurements on 11 Aug 2004 are shown for several heights according to the legend.

exhibit poorer forecast skill near the ocean surface than above 100 m MSL, which is likely indicative of the complexities of atmosphere–surface interactions and the difficulties that models have in properly representing them.

The growth of model errors at longer lead times is illustrated in Fig. 10, again averaged for the August study period. HRDL wind data, also averaged vertically over the layer from 20 to 200 m MSL, were used to verify model forecasts (averaged over the same layer) out to 12-h lead time. Figure 10 shows the rms error for the four model configuration runs over the August study period. The plot shows the scalar-wind (i.e., mean wind speed) error and the magnitude of the vector-wind error, which includes directional deviations. Except for the initialization errors for the HRRR-2012P model described previously, the traces for the higher-resolution models in blue (HRRR-2012P and NAM-CONUS-nest) lie at or below those of the lower-resolution "parent" models for a short time after initialization by as much as 0.3 m s^{-1}. After this interval, the finer-resolution embedded models exhibit worse error statistics for this dataset. This kind of error behavior has been noted before (see "Effects of finer grid spacing on model skill" sidebar). During the "nowcasting" lead times (up to 3 h or so), wind speed errors were near 2 m s^{-1}, increasing to >2.5 m s^{-1} after 8-h lead time. The vector-wind errors were larger, starting with initial errors of 2.5 m s^{-1}, growing to 4 m s^{-1} for the 12-h predictions. The magnitudes of the vector-wind errors are similar to annual RAP model 6-h lead-time values found by Benjamin et al. (2016) over the United States.

The behavior of the model wind speed forecast at ~100 m MSL for each model as a function of time of day is shown in Fig. 11. The 100-m height is near the typical hub heights of contemporary turbines. For comparison, the mean HRDL-measured wind at 100 m MSL is also plotted for each hour of the day (black line, which remains the same in all 12 panels). In the NAMRR results, the initial wind (forecast 0) is underestimated by about 2 m s^{-1} when the model is initialized at night, especially after local midnight (0500 UTC), but better represented when initialized during daytime (1200–2200 UTC), when the differences between measurements and model output are generally less than 1 m s^{-1}.

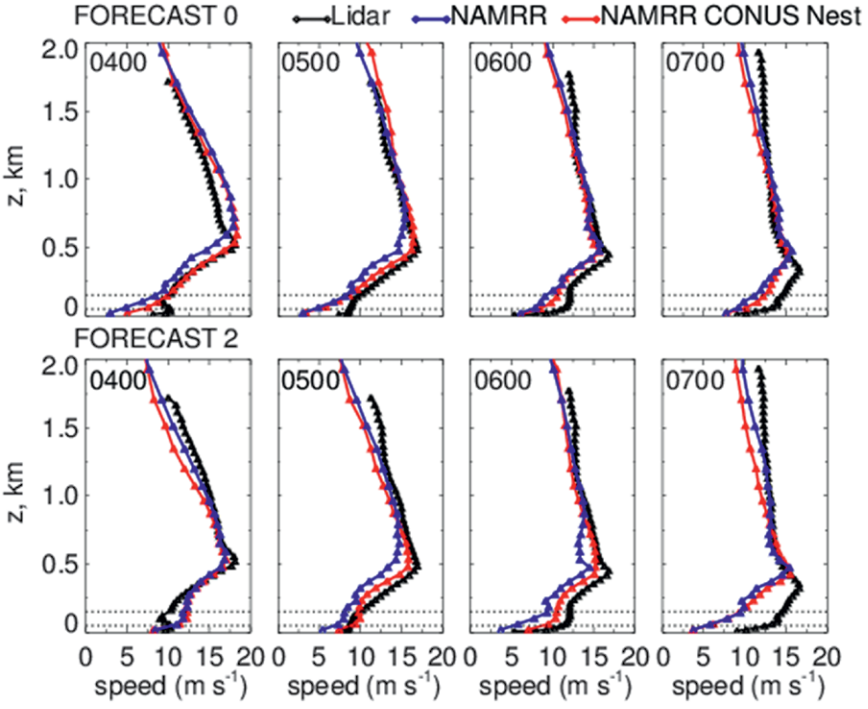

Fig. 8. Sampling of HRDL vs model profiles for four consecutive hours on 11 Aug 2004. (top) Initial model conditions (forecast hour 0) and (bottom) 2-h lead-time forecast valid at the indicated hour. Black lines show HRDL indicated-hour profiles. Blue lines are for the NAMRR-parent model, and the red lines are for the NAMRR-CONUS-nest. Dotted horizontal lines at 50 and 150 m delimit hypothetical turbine rotor layer.

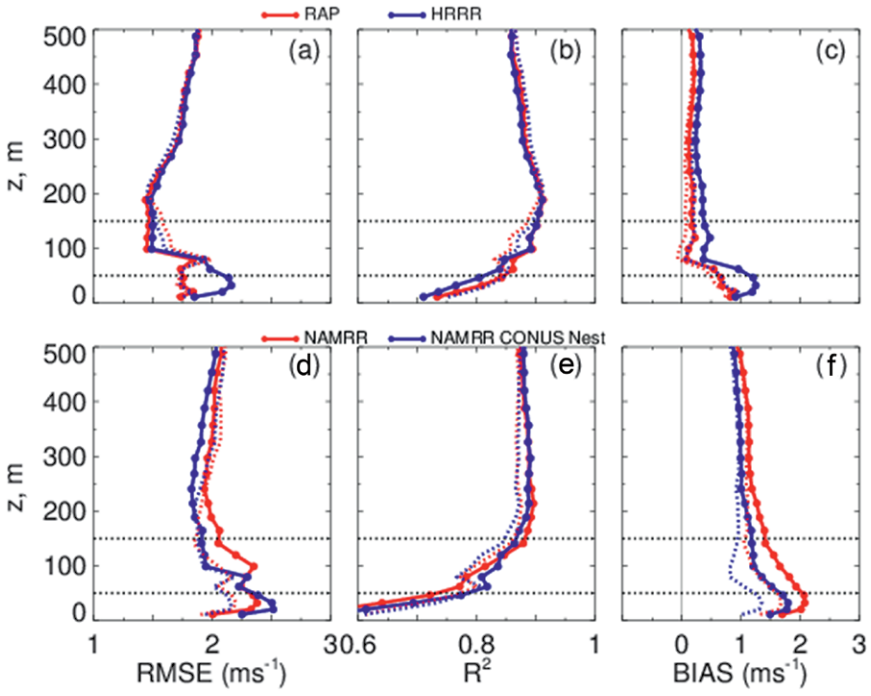

Fig. 9. Vertical profiles of (a),(d) rms error (m s^{-1}; horizontal axis), (b),(e) coefficients of determination, and (c),(f) mean bias (m s^{-1}; horizontal axis) between HRDL-measured and modeled wind speed averaged over the Aug study period. Results are shown for (a)–(c) the RAP-2012P (red lines) and HRRR-2012P (blue lines) models and (d)–(f) the NAMRR-parent (red lines) and NAMRR-CONUS-nest (blue lines). Solid lines are for the initial conditions, and dotted lines are for the 2-h forecast. Dotted horizontal lines at 50 and 150 m delimit hypothetical turbine rotor layer, as in Fig. 8.

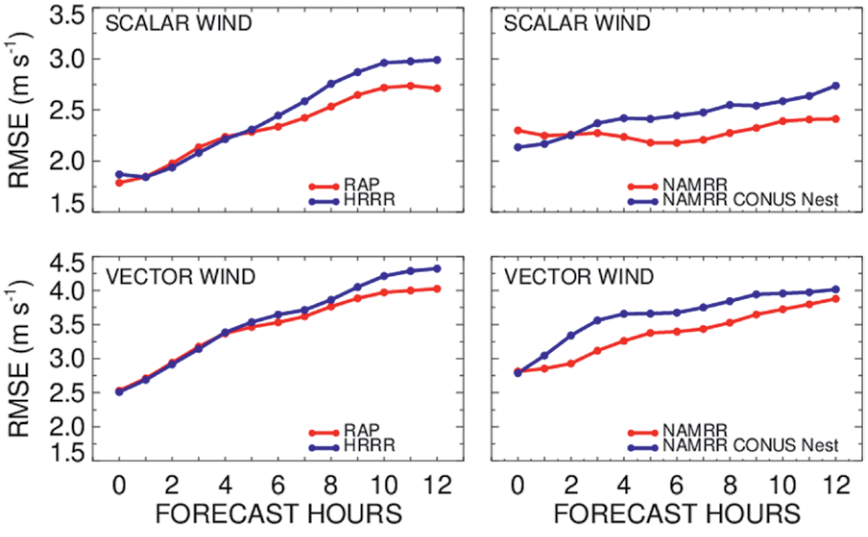

Fig. 10. The rms error between observed and modeled wind for 6–12 Aug, averaged over the 20–200-m MSL layer. (left) RAP-2012P (red line) and HRRR-2012P (blue line) models and (right) NAMRR-parent (red line) and NAMRR-CONUS-nest (blue line) models. (top) The scalar wind, or wind speed, rms error and (bottom) the vector-wind rms error, which is $\sigma_v^2 = (\sigma_u^2 + \sigma_v^2)^{1/2}$, where σ_u^2 and σ_v^2 are the rms errors in the two orthogonal wind components and σ_v^2 is the plotted vector-wind error.

The 1-h forecasts (forecast 1) valid after local midnight continue to show low biases of 2 m s^{-1} on average, sustaining the underestimates in the initial conditions, as do the 2- and 3-h forecasts. During daylight hours, on the other hand, the 1–3-h forecasts are mostly within 1 m s^{-1} of the measured winds, also in line with the error characteristics of the daytime initial conditions. Longer term, the 12-h forecasts valid in the afternoon after 1800 UTC overpredict the measured winds by 2–4 m s^{-1}, extending into the evening hours through 0800 UTC. The RAP–HRRR-2012P results on the right are similar. Thus, model errors depended on hour of the day: the model initial state and short-term forecasts, on average, had the largest errors during the nighttime hours after local midnight, whereas the longer-term predictions (12 h) had the largest errors in the late afternoon and early evening hours.

The errors in this section showed that the models tended to underpredict (low bias) wind speeds for forecasts of less than 3 h and overpredict them at 12 h; rms wind speed errors were larger in the lowest 100 m (Fig. 9), grew with forecast lead time (Fig. 10), and were largest at night (Fig. 11), especially at short lead times. Quantitatively, Table 3 compares results for this study with others where offshore winds aloft were available for extended time periods, including

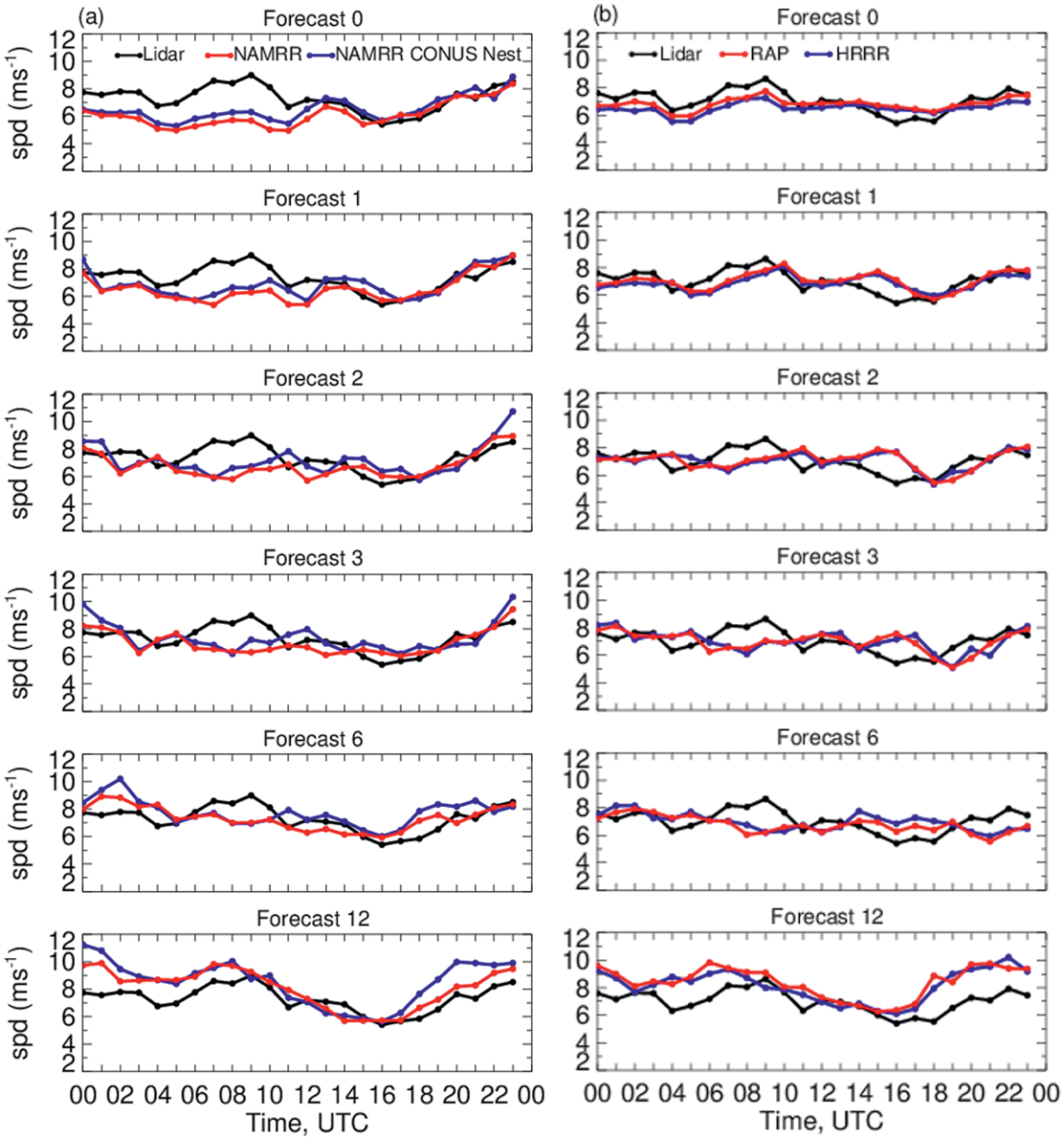

Fig. 11. Time series of lidar-measured and modeled wind plotted against hour of the day (UTC, which leads eastern standard time by 5 h) averaged for the Aug study period. (a) NAMRR and (b) RAP–HRRR-2012P model results. Black trace is mean HRDL wind speed data at 100 m MSL, averaged for each hour. (top) The model initial wind speed (forecast hour 0, or forecast 0) is plotted vs hour of the day; the blue lines are for the parent runs (NAMRR and RAP-2012P) and the red lines are for the finer-mesh, embedded models (NAMRR-CONUS-nest and HRRR-2012P). (second row) The 1-h wind speed forecast averaged for each hour and plotted at the valid time of the forecast. (third row)–(sixth row) As in the second row, but for forecast hours 2, 3, 6, and 12, respectively.

the two other studies using the present NEAQS-04 dataset, where errors were calculated in different ways. Wind speed errors were generally 2–2.5 m s^{-1} and vector-wind errors were 4–4.5 m s^{-1}. For reference, these results are consistent with Benjamin et al. (2016), who verified 1 year of RAP output between 1,000 and 100 hPa against all available routine rawinsonde launches over the CONUS and found vector-wind errors of ~4 m s^{-1} for 6-h forecasts over 50-hPa vertical layers through most of the troposphere. The statistics for the Gulf of Maine are for weeklong samples in summer. Annual error values could be expected to be smaller for the larger sample size, but even if they were half the weekly value, they

would still be large compared to the 0.1 m s^{-1} desired by the industry.

HORIZONTAL DISTRIBUTION OF MEAN WIND AT HUB HEIGHT AND SURFACE. Characterizing the horizontal variability of atmospheric flows is important in several ways. Over shorter time periods (hours), horizontal variability is fundamental to understanding and NWP modeling of the atmosphere, whereas variations over longer times (mean annual, seasonal, monthly) are important for wind plant siting.

On the shorter time scales, horizontal variability often implies flow divergence,[2] which, in turn, implies vertical motion and vertical transport. The depth, location, extent, and rate of transfer are key attributes of atmospheric flow structures that perform vertical exchange and that therefore make essential contributions to the budgets of many atmospheric quantities, such as heat, momentum, and moisture. Future states of the atmosphere are determined by these four attributes or properties of vertical-exchange mechanisms. To improve forecast skill, NWP models must more accurately represent these properties of the mechanisms (over the entire simulation domain, to be sure). Being able to determine whether model updates have improved this representation will require measurements—multiscale measurement arrays that capture these attributes (Banta et al. 2013). Characterizing, understanding, and accurate modeling of horizontal-flow variability (with its associated vertical transports) therefore have a fundamental significance.

Over longer time periods, horizontal variability of the mean wind distribution may result from persistent or recurrent flows, and in the coastal zone, a number of mechanisms generate this kind of horizontal-flow variability. Alongshore variability may result from coastline shape, including headlands that locally block or accelerate flow parallel to the coastline

[2] Horizontal variations in the wind field are a necessary but not sufficient condition for flow divergence, since purely rotational flow has horizontal variability but is nondivergent.

Table 3. Error table. POWER data: 6-h forecasts for Aug study period. Data are from Drechsel et al. (2012, their Figs. 6 and 8), Djalalova et al. (2016, their Fig. 16), Pichugina et al. (2017, their Fig. 18), and Benjamin et al. (2016, their Fig. 7). An asterisk indicates that the mean bias is removed (unbiased rmse).

Location	Data source	Model	Averaging depth	Instrument	Averaging period	Lead time (h)	Error (m s^{-1})	Relative error
Rms error—Wind speed								
Gulf of Maine	Current paper	RAP NAM	20–200 m	HRDL	1 week	6–12	2.5 (2.0*) 2.1 (1.6*)	
FINO-1	Drechsel et al. (2012)	ECMWF	100 m	Tower	1 year	0–9	1.4*	14%
Horns Rev	Drechsel et al. (2012)	ECMWF	70 m	Tower	1 year	0–9	1.4*	16%
Gulf of Maine	Djalalova et al. (2016)	RAP NAMRR	120–2,000 m	915 profiler	2 weeks	1, 12	2.6, 2.8 2.4, 2.9	
Gulf of Maine	Pichugina et al. (2017)	RAP NAMRR	10–500 m	HRDL	1 week	1 and 12	1.7, 2.4 2.2, 2.4	
Rms error—Vector wind								
Gulf of Maine	Current paper	RAP NAMRR	20–200 m	HRDL	1 week	6–12	4 4	
Gulf of Maine	Djalalova et al. (2016)	RAP NAMRR	120–2,000 m	915 profiler	2 weeks	1 and 12	3.3, 3.9 3.2, 3.7	
Gulf of Maine	Pichugina et al. (2017)	RAP NAMRR	10–500 m	HRDL	1 week	1 and 12	2.4, 4.0 2.7, 3.8	
United States	Benjamin et al. 2016	RAP	975–925 hPa	Rawinsonde	1 year	1 and 12	3.4, 4.0	
Mean absolute error—Wind speed								
Gulf of Maine	Djalalova et al. (2016)	RAP NAMRR	10–2,000 m	HRDL	2 weeks	3	1.5 1.5	

(e.g., Rothermel et al. 1998a; Jiang et al. 2008), or terrain effects, such as gaps in coastal topography that focus wind speed maxima, which can eject out over the sea. Diurnal sea-breeze circulations, which often form an important contribution to the offshore wind resource, are driven by large diurnal changes in surface heating inland. Along-coast variation in coastal topography, surface roughness (including vegetation type), soil moisture, soil type, or land use may generate alongshore variability in the wind resource through their modification of local surface heat fluxes over land, with a resultant change in the local intensity of the sea-breeze circulation. Such localized phenomena may recur often enough to be part of the long-term-average wind resource at offshore locations and thus to appear in long-term (seasonal, annual, or longer) maps of mean wind speed.

One of the most important decisions in developing a wind plant is where to put it. The essential problem for siting is to ensure an adequate wind resource at the proposed location, enough that the plant will show a profit over the lifetime of the installed turbines. Differences of 1–2 m s^{-1} in annual average wind speed can translate to tens of millions of dollars (if not hundreds of millions) over several years (Marquis et al. 2011). Observational studies (e.g., Rothermel et al. 1998a,b; Jiang et al. 2008; Pichugina et al. 2017a,b; Hasager et al. 2015) suggest that horizontal differences of this magnitude and greater exist in the offshore environment. Even model output, though smoothed, can show such alongshore variability, for example, the thick arrow in the upper right of Fig. 4.

What can the NEAQS-04 dataset tell us about horizontal variability of the wind during the July–August 2004 sampling period? Figure 12 shows color-coded, HRDL-measured wind speeds near the turbine hub height (80 m) and at 10 m MSL plotted on the ship tracks for the entire monthlong cruise. The data show considerable horizontal and temporal variability in the winds along the track due to spatial variability of the wind flow in the region (Pichugina et al. 2012, 2017a; Banta et al. 2014) but also partly due to sampling in different locations and at different times. A key question can be addressed by this dataset: Can the horizontal *distribution* of near-surface winds be used to infer the horizontal distribution of winds at the hub height? Near-surface horizontal wind speed distributions are often obtained from buoy, ship, or satellite (e.g., synthetic-aperture radar, scatterometer, radiometer) data. Careful inspection of the 10 and 80 m MSL plots in Fig. 12 shows these distributions to be dissimilar. The variations can be highlighted by plotting the wind speed differences between 80 and 10 m (right panel). Quantitatively, the right panel shows that the differences span a range from near 0 to more than 5 m s^{-1} depending on location. A consequence of this large differential across the region is that extrapolating wind speeds upward from near the surface, using profile shape parameters optimized for one part of the domain, would lead to large errors in other parts of the domain: 5 m s^{-1} is likely a range too large to be accommodated by standardized-profile extrapolation with any accuracy (Pichugina et al. 2012, 2017a), even if adjusted for stability or other predetermined environmental factors. Pichugina et al. (2017a) have shown with this NEAQS-04 dataset that α, the exponent in the power-law wind profile, does in fact vary significantly over the NEAQS domain (refer to Fig. 11 of that paper).

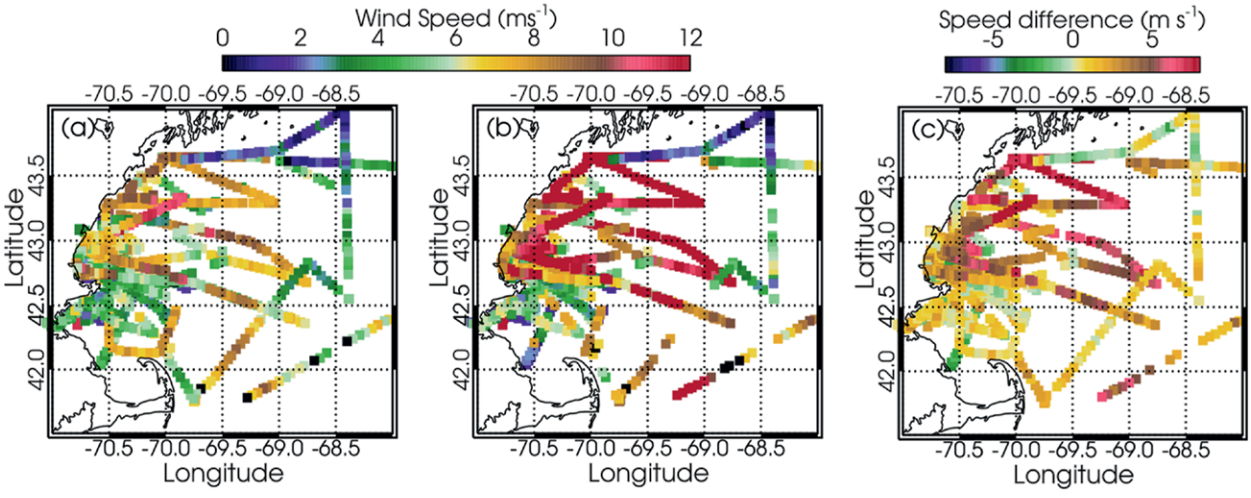

Fig. 12. Horizontal distributions of lidar-measured wind speed within the large white box region in Fig. 1 at (a) 10 and (b) 80 m MSL along ship tracks from 9 Jul to 12 Aug. (c) Wind speed difference between (b) and (a).

DISCUSSION AND RECOMMENDATIONS. Significant differences appear between the horizontal complexities of the measured versus modeled wind fields. Figure 12 reflects the strong spatial viability observed in other studies (Pichugina et al. 2012, 2017a,b; Djalalova et al. 2016). For models, Fig. 4 shows a much smoothed wind speed distribution in the multiyear average. Even in weekly averaged model fields for the POWER study periods, the modeled wind speed distribution appears smooth (Djalalova et al. 2016; see their Fig. 5). A critical question for resource assessment becomes will all this observed detail (Fig. 12) be averaged out of long-term (annual, multiyear) horizontal wind distributions in the atmosphere? Or do the long-term-mean patterns also show detail to their structure, with areas of higher and lower wind speeds? Answering this question will again require measurements: an array of profiling instruments distributed over the horizontal area of interest.

As part of this project, DOE asked NOAA/ESRL to use the results of this study to propose designs for such a sampling array, assuming the availability of offshore profiling technology, such as the buoy-based lidar systems mentioned previously. These systems so far have not been extensively tested and verified on location in the harsh marine environment, but if they do prove successful in being robust and able to remove sensor motion to accurately measure the wind profile, how should they be deployed?

A large-scale, East Coast sampling array to measure horizontal variability at alongshore distances of 500 km is shown in Fig. 13 (top left). The distribution of wind speeds has some noteworthy characteristics, including 1) a strong perpendicular-to-coastline, or *cross-shore*, wind speed gradient in the first 75–100 km from shore at all locations; 2) the variability of this gradient in the cross-shore direction, with the strongest gradients closest to the coastline and weaker gradients farther from the shore; and 3) the variability of the gradient in the alongshore direction (e.g., the 9 m s^{-1} isotach is much closer to shore in the Gulf of Maine than off the Maryland–Delaware coastline). To investigate these kinds of features, a proposed array of wind-profiling sites, most likely buoy based, has been added to this wind resource map (top panels of Fig. 13). This network consists of four cross-shore lines of three to five profiling instruments (three shown here). Each line is depicted in a region of strong, moderate, or weak cross-shore gradient.

Variability over smaller alongshore distances on the order of 100 km or less is probably more critical for making offshore WE siting decisions; for example, which of two candidate wind plant sites separated by 100 km (or 50 or 30 km) in the alongshore direction should be preferred for development? And, for model error diagnosis investigations, what are the important distance scales of meteorological variability in the alongshore direction, and are they accurately represented in the models? What are their amplitudes and potential impacts on WE or NWP model skill? Examples of more localized array designs are shown in Fig. 13 in the bottom panels for the Gulf of Maine. A variation having one fewer sensor is shown in the bottom-right panel. These kinds of arrangements better characterize the local alongshore variability in wind speed, which could be scaled down to 50- or 100-km sensor intervals as needed.

The important scales of horizontal variability for each location are unknown, and the necessary sampling intervals to detect, for example, regions of low wind resource, so as to avoid siting a wind farm in an unfavorable location, are also unknown. Long-term sampling arrays would be deployed based on the best-available information for the region and on available resources. (How many sensors can the project afford?) But if the sampling array is not of high-enough density, important phenomena will be missed without knowing it. The problem could be addressed by monthlong IOPs during different seasons, incorporating mobile platforms with profiling capability, such as the *RHB* in this study, or instrumented aircraft [including unmanned aerial vehicles (UAVs)]. The tracks of the mobile sensing platforms will have to be carefully designed to avoid potential nonuniform sampling issues described previously and illustrated in Fig. 12. Thus, studies of spatial variability will most likely require multiple approaches (e.g., a mixture of sensors on fixed platforms and on mobile platforms; Banta et al. 2012). If the mobile platform were a ship, sailing past fixed-location profiling devices is important for sorting out the temporal versus spatial ambiguities in the variability of the measured flows. Information from these monthlong IOPs could then be used to adjust or reposition the long-term array instruments for optimum sensing of the long-term local flow structure.

CONCLUSIONS. The rotor-layer wind speed errors in NWP were larger than required for siting decisions for offshore wind farms. Even if measurements were available to adjust for bias, rms errors are still large, and furthermore, the horizontal complexity of the hub-height winds may not be well represented by the models. The approach with the highest likelihood of reducing risks of overestimating (or underestimating) the wind resource would be

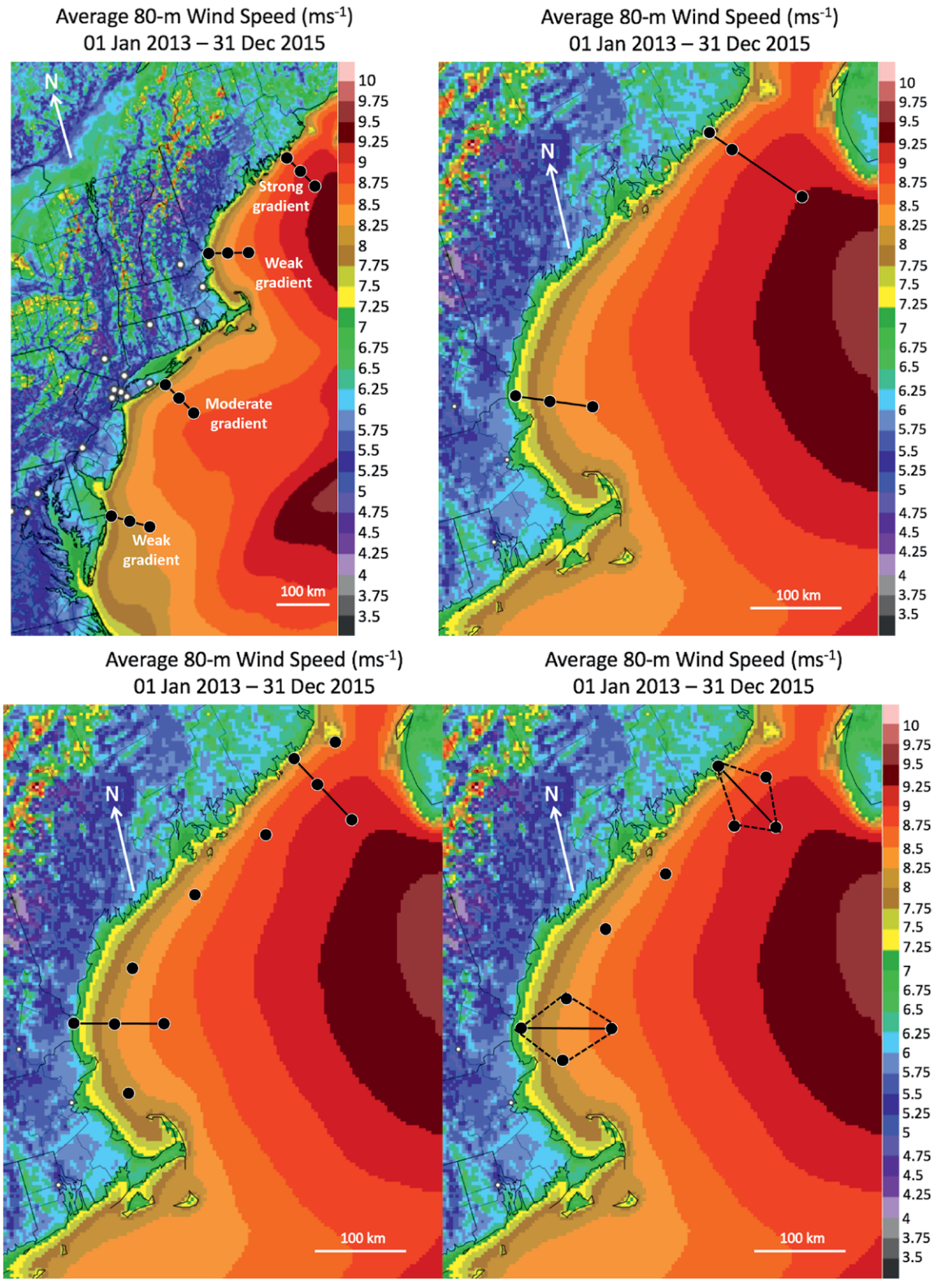

Fig. 13. Proposed layouts for profiling sensor arrays superimposed on maps of wind speed from Figs. 4a and 4b. (top) Wind profile measurement-array configuration focused on measuring cross-shore gradients of wind speed: (left) larger-scale view and (right) centered on Gulf of Maine and New England. (bottom left) Example of a deployment of offshore wind profiling sites (black dots) having both alongshore and cross-shore sampling. (bottom right) As in the bottom-left panel, but cross-shore sampling is accomplished in a diamond pattern.

to deploy offshore arrays of wind-profiling sensors, such as lidars, as described in the previous section. Such arrays would also provide datasets to improve NWP models.

In models as well as in the atmosphere, the lowest several hundred meters of the atmosphere is a layer where important vertical and horizontal transports occur that have a critical role in vertical exchange between Earth's surface and the lower atmosphere. These processes have traditionally been poorly characterized because of measurement limitations (Banta et al. 2016). As mentioned, model skill in this layer is also not well known. For the most part, interest in this layer and its dynamics has been largely academic. Wind energy has changed that. By virtue of having its hardware embedded in that layer, WE has drawn attention to the serious gap in our understanding and modeling of the layer's structure, dynamics, and evolution. New instrumentation, such as Doppler lidar, gives promise of characterizing flows in this layer at much finer resolution, frequency, and precision than previously, with a goal of improved understanding and modeling of the critical vertical exchange processes through this layer.

Conceptually, the pathways for these exchange processes are known. The surface is often the most important source or sink of dynamic and thermodynamic quantities, as well as other atmospheric constituents. Sources and sinks constitute major components of budgets. Atmospheric processes distribute atmospheric quantities vertically and horizontally, and future states of the atmosphere depend on where the quantities are transported and how they become distributed. Vertical transport occurs throughout the atmosphere but tends to be concentrated at lower levels over mountainous and coastal regions, as well as other zones of strong horizontal gradient. What is not well known is how well NWP models handle these mechanisms. Accuracy requires that models faithfully represent the budgets at each vertical level, especially in regions where vertical transfer is strong, such as coastal zones. Whether they do so, and how to fix them if they do not, can only be determined by profile measurement-array deployments as described above.

Reliable information about the lower atmosphere thus impacts WE both directly and indirectly. Directly, accurately measured data in the turbine rotor layer will contribute to more reliable resource assessment and nowcasting. Indirectly, better forecasts by the "foundational models," those used by the WE industry as a starting point for tailored WE forecast products, would save significant money and make WE more efficient and effective (Ahlstrom et al. 2013). Improved forecasts would also, of course, benefit the public at large and all applications that use meteorological data and forecasts. Stated more succinctly, solving WE's problems solves the problems of everyone who needs accurate and reliable atmospheric information and forecasts in the coastal atmospheric boundary layer.

ACKNOWLEDGMENTS. This study was sponsored by the U.S. Department of Energy Energy Efficiency and Renewable Energy Office Wind and Water Power Program, by the NOAA Atmospheric Science for Renewable Energy Program, and by the NOAA/ESRL Air Quality Program. We thank our colleagues Raul Alvarez, Dan Law, Janet Machol, Richard Marchbanks, Brandi McCarty, Rob K. Newsom, Sara Tucker, Scott Sandberg, and Ann Weickmann for their long hours collecting HRDL data on the ship during the NEAQS 2004 campaign and Lisa Darby for her careful review of this manuscript.

REFERENCES

Ahlstrom, M., and Coauthors, 2013: Knowledge is power: Efficiently integrating wind energy and wind forecasts. *IEEE Power Energy Mag.*, **11**, 45–52, https://doi.org/10.1109/MPE.2013.2277999.

Alpert, J., 2004: Sub-grid scale mountain blocking at NCEP. *Proc. 20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., P2.4., https://ams.confex.com/ams/pdfpapers/71011.pdf.

Angevine, W. M., A. B. White, C. J. Senff, M. Trainer, R. M. Banta, and M. A. Ayoub, 2003: Urban–rural contrasts in mixing height and cloudiness over Nashville in 1999. *J. Geophys. Res.*, **108**, 4092, https://doi.org/10.1029/2001JD001061.

——, J. E. Hare, C. W. Fairall, D. E. Wolfe, R. J. Hill, W. A. Brewer, and A. B. White, 2006: Structure and formation of the highly stable marine boundary layer over the Gulf of Maine. *J. Geophys. Res.*, **111**, D23S22, https://doi.org/10.1029/2006JD007465.

Archer, C., and Coauthors, 2014: Meteorology for coastal/offshore wind energy in the United States: Recommendations and research needs for the next 10 years. *Bull. Amer. Meteor. Soc.*, **95**, 515–519, https://doi.org/10.1175/BAMS-D-13-00108.1.

Banta, R. M., L. D. Olivier, and D. H. Levinson, 1993: Evolution of the Monterey Bay sea-breeze layer as observed by pulsed Doppler lidar. *J. Atmos. Sci.*, **50**, 3959–3982, https://doi.org/10.1175/1520-0469(1993)050<3959:EOTMBS>2.0.CO;2.

——, ——, P. H. Gudiksen, and R. Lange, 1996: Implications of small-scale flow features to

modeling dispersion over complex terrain. *J. Appl. Meteor.*, **35**, 330–342, https://doi.org/10.1175/1520-0450(1996)035<0330:IOSSFF>2.0.CO;2.

——, and Coauthors, 1997: Nocturnal cleansing flows in a tributary valley. *Atmos. Environ.*, **31**, 2147–2162, https://doi.org/10.1016/S1352-2310(96)00359-7.

——, L. S. Darby, J. D. Fast, J. O. Pinto, C. D. Whiteman, W. J. Shaw, and B. D. Orr, 2004: Nocturnal low-level jet in a mountain basin complex. Part I: Evolution and implications to other flow features. *J. Appl. Meteor.*, **43**, 1348–1365, https://doi.org/10.1175/JAM2142.1.

——, and Coauthors, 2005: A bad air day in Houston. *Bull. Amer. Meteor. Soc.*, **86**, 657–669, https://doi.org/10.1175/BAMS-86-5-657.

——, and Coauthors, 2012: Observational techniques: Sampling the mountain atmosphere. *Mountain Weather Research and Forecasting*, F. K. Chow, S. De Wekker, and B. Snyder, Eds., Springer, 409–530, https://doi.org/10.1007/978-94-007-4098-3.

——, Y. L. Pichugina, N. D. Kelley, W. A. Brewer, and R. M. Hardesty, 2013: Wind energy meteorology: Insight into wind properties in the turbine rotor layer of the atmosphere from high-resolution Doppler lidar. *Bull. Amer. Meteor. Soc.*, **94**, 883–902, https://doi.org/10.1175/BAMS-D-11-00057.1.

——, and Coauthors, 2014: NOAA study to inform meteorological observation for offshore wind: Positioning of Offshore Wind Energy Resources (POWER). NOAA Tech. Rep., 145 pp., www.esrl.noaa.gov/gsd/renewable/AMR_DOE-FinalReport-POWERproject-1.pdf.

——, and Coauthors, 2015: 3D volumetric analysis of wind turbine wake properties in the atmosphere using high-resolution Doppler lidar. *J. Atmos. Oceanic Technol.*, **32**, 904–914, https://doi.org/10.1175/JTECH-D-14-00078.1.

——, and Coauthors, 2016: Renewable energy, measurements, and forecast improvement: Illuminating the atmosphere's "dark zone." *Forum on Observing the Environment from the Ground Up*, Washington, DC, Amer. Meteor. Soc., https://ams.confex.com/ams/GroundUp/webprogram/meeting.html#Tuesday1.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Chou, M.-D., and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, 85 pp., https://ia600502.us.archive.org/23/items/nasa_techdoc_19950009331/19950009331.pdf.

Clark, T. L., W. D. Hall, and R. M. Banta, 1994: Two- and three-dimensional simulations of the 9 January 1989 severe Boulder windstorm: Comparison with observations. *J. Atmos. Sci.*, **51**, 2317–2343, https://doi.org/10.1175/1520-0469(1994)051<2317:TATDSO>2.0.CO;2.

Colle, B. A., M. A. Sienkiewicz, C. L. Archer, D. Veron, F. Veron, W. Kempton, and J. E. Mak, 2016: Improving the Mapping and Prediction of Offshore Wind Resources (IMPOWR): Experimental overview and first results. *Bull. Amer. Meteor. Soc.*, **97**, 1377–1390, https://doi.org/10.1175/BAMS-D-14-00253.1.

Darby, L. S., W. D. Neff, and R. M. Banta, 1999: Multiscale analysis of a meso-β frontal passage in the complex terrain of the Colorado Front Range. *Mon. Wea. Rev.*, **127**, 2062–2081, https://doi.org/10.1175/1520-0493(1999)127<2062:MAOAMF>2.0.CO;2.

——, and Coauthors, 2002a: Vertical variations in O_3 concentrations before and after a gust front passage. *J. Geophys. Res.*, **107**, 4321, https://doi.org/10.1029/2001JD000996.

——, R. M. Banta, and R. A. Pielke Sr., 2002b: Comparisons between mesoscale model terrain sensitivity and Doppler lidar measurements of the sea breeze at Monterey Bay. *Mon. Wea. Rev.*, **130**, 2813–2838, https://doi.org/10.1175/1520-0493(2002)130<2813:CBMMTS>2.0.CO;2.

——, K. J. Allwine, and R. M. Banta, 2006: Nocturnal low-level jet in a mountain basin complex. Part II: Transport and diffusion of tracer under stable conditions. *J. Appl. Meteor. Climatol.*, **45**, 740–753, https://doi.org/10.1175/JAM2367.1.

——, and Coauthors, 2007: Ozone differences between near-coastal and offshore sites in New England: Role of meteorology. *J. Geophys. Res.*, **112**, D16S91, https://doi.org/10.1029/2007JD008446.

Djalalova, I. V., and Coauthors, 2016: The POWER experiment: Impact of assimilation of a network of coastal wind profiling radars on simulating offshore winds in and above the wind turbine layer. *Wea. Forecasting*, **31**, 1071–1091, https://doi.org/10.1175/WAF-D-15-0104.1.

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

Drechsel, S., G. J. Mayr, J. W. Messner, and R. Stauffer, 2012: Lower boundary wind speeds: Measurements and verification of forecasts. *J. Appl. Meteor. Climatol.*, **51**, 1602–1617, https://doi.org/10.1175/JAMC-D-11-0247.1.

Dvorak, M. J., C. L. Archer, and M. Z. Jacobson, 2010: California offshore wind energy potential. *Renewable Energy*, **35**, 1244–1254, https://doi.org/10.1016/j.renene.2009.11.022.

——, E. D. Stoutenburg, C. L. Archer, W. Kempton, and M. Z. Jacobson, 2012: Where is the ideal location for a US East Coast offshore grid? *Geophys. Res. Lett.*, **39**, L06804, https://doi.org/10.1029/2011GL050659.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, https://doi.org/10.1029/2002JD003296.

Fairall, C. W., and Coauthors, 2006: Turbulent bulk transfer coefficients and ozone deposition velocity in the International Consortium for Atmospheric Research into Transport and Transformation. *J. Geophys. Res.*, **111**, D23S20, https://doi.org/10.1029/2006JD007597.

Fehsenfeld, F. C., and Coauthors, 2006: International Consortium for Atmospheric Research on Transport and Transformation (ICARTT): North America to Europe—Overview of the 2004 summer field study. *J. Geophys. Res.*, **111**, D23S01, https://doi.org/10.1029/2006JD007829.

Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. Preprints, *19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., 280–283.

——, W. Wang, and E. Colon, 2011: Evaluating cloud microphysics schemes in nested NMMB forecasts. *Proc. 24th Conf. on Weather Analysis and Forecasting/20th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14B.1., https://ams.confex.com/ams/91Annual/webprogram/Paper179488.html.

Flamant, C., and Coauthors, 2002: Gap flow in an Alpine valley during a shallow south föhn event: Observations, numerical simulations and hydraulic analog. *Quart. J. Roy. Meteor. Soc.*, **128**, 1173–1210, https://doi.org/10.1256/003590002320373256.

Giannakopoulou, E.-M., and R. Nhili, 2014: WRF Model methodology for offshore wind energy applications. *Adv. Meteor.*, **2014**, 319819, https://doi.org/10.1155/2014/319819.

Gohm, A., G. J. Mayr, L. S. Darby, and R. M. Banta, 2010: Evolution and structure of a cold front in an Alpine valley as revealed by a Doppler lidar. *Quart. J. Royal Meteor. Soc.*, **136**, 962–977, https://doi.org/10.1002/qj.609.

Grund, C. J., R. M. Banta, J. L. George, J. N. Howell, M. J. Post, R. A. Richter, and A. M. Weickmann, 2001: High-resolution Doppler lidar for boundary layer and cloud research. *J. Atmos. Oceanic Technol.*, **18**, 376–393, https://doi.org/10.1175/1520-0426(2001)018<0376:HRDLFB>2.0.CO;2.

Hasager, C. B., and Coauthors, 2015: Offshore wind climatology based on synergistic use of Envisat ASAR, ASCAT and QuikSCAT. *Remote Sens. Environ.*, **156**, 247–263, https://doi.org/10.1016/j.rse.2014.09.030.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

James, E. P., S. G. Benjamin, and M. Marquis, 2017: A unified high-resolution wind and solar dataset from a rapidly updating high-resolution numerical weather prediction model. *Renewable Energy*, **102B**, 390–405, https://doi.org/10.1016/j.renene.2016.10.059.

——, ——, and ——, 2018: Offshore wind speed estimates from a high-resolution rapidly updating numerical weather prediction model forecast dataset. *Wind Energy*, **21**, 264–284, https://doi.org/10.1002/we.2161.

Janjić, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

——, 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP meso model. NCEP Office Note 437, 61 pp., www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf.

Jiang, Q., J. D. Doyle, T. Haack, M. J. Dvorak, C. L. Archer, and M. Z. Jacobson, 2008: Exploring wind energy potential off the California Coast. *Geophys. Res. Lett.*, **35**, L20819, https://doi.org/10.1029/2008GL034674.

Jimenez, B., F. Durante, B. Lange, T. Kreutzer, and J. Tambke, 2007: Offshore wind resource assessment with WAsP and MM5: Comparative study for the German Bight. *Wind Energy*, **10**, 121–134, https://doi.org/10.1002/we.212.

Kraus, E. B., and J. A. Businger, 1994: *Atmosphere–Ocean Interaction*. 2nd ed. Oxford University Press, 384 pp.

Krogsæter, O., and J. Reuder, 2015: Validation of boundary layer parameterization schemes in the Weather Research and Forecasting Model under the aspect of offshore wind energy applications—Part I: Average

wind speed and wind shear. *Wind Energy*, **18**, 769–782, https://doi.org/10.1002/we.1727.

Langford, A. O., C. J. Senff, R. J. Alvarez II, R. M. Banta, and R. M. Hardesty, 2010: Long-range transport of ozone from the Los Angeles Basin: A case study. *Geophys. Res. Lett.*, **37**, L06807, https://doi.org/10.1029/2010GL042507.

Levinson, D. H., and R. M. Banta, 1995: Observations of a terrain-forced mesoscale vortex and canyon drainage flows along the Front Range of the Colorado Rockies. *Mon. Wea. Rev.*, **123**, 2029–2050, https://doi.org/10.1175/1520-0493(1995)123<2029:OOATFM>2.0.CO;2.

Marquis, M., J. M. Wilczak, M. Ahlstrom, J. Sharp, A. Stern, J. C. Smith, and S. Calvert, 2011: Forecasting the wind to reach significant penetration levels of wind energy. *Bull. Amer. Meteor. Soc.*, **92**, 1159–1171, https://doi.org/10.1175/2011BAMS3033.1.

Mass, C., D. Owens, K. Westrick, and B. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? The results of two years of real-time numerical weather prediction over the Pacific Northwest. *Bull. Amer. Meteor. Soc.*, **83**, 407–430, https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.

Mittermaier, M., 2014: A strategy for verifying near-convection-resolving model forecasts at observing sites. *Wea. Forecasting*, **29**, 185–204, https://doi.org/10.1175/WAF-D-12-00075.1.

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.

Musial, W., and B. Ram, 2010: Large-scale offshore wind power in the United States: Assessment of opportunities and barriers. National Renewable Energy Laboratory Tech. Rep. NREL/TP-500-40745, 221 pp.

National Research Council, 1992: Coastal meteorology: A review of the state of the science. National Research Council Rep., 112 pp., https://doi.org/10.17226/1991.

Pichugina, Y. L., R. M. Banta, W. A. Brewer, S. P. Sandberg, and R. M. Hardesty, 2012: Doppler lidar–based wind-profile measurement system for offshore wind-energy and other marine boundary layer applications. *J. Appl. Meteor. Climatol.*, **51**, 327–349, https://doi.org/10.1175/JAMC-D-11-040.1.

——, and Coauthors, 2017a: Properties of offshore low-level jet and rotor layer wind shear as measured by scanning Doppler lidar. *Wind Energy*, **20**, 987–1002, https://doi.org/10.1002/we.2075.

——, and Coauthors, 2017b: Assessment of NWP forecast models in simulating offshore winds through the lower boundary layer by measurements from a ship-based scanning Doppler lidar. *Mon. Wea. Rev.*, **145**, 4277–4301, https://doi.org/10.1175/MWR-D-16-0442.1.

Rothermel, J., and Coauthors, 1998a: The Multi-Center Airborne Coherent Atmospheric Wind Sensor. *Bull. Amer. Meteor. Soc.*, **79**, 581–599, https://doi.org/10.1175/1520-0477(1998)079<0581:TMCACA>2.0.CO;2.

——, and Coauthors, 1998b: Remote sensing of multi-level wind fields with high-energy airborne scanning coherent Doppler lidar. *Opt. Express*, **2**, 40–50, https://doi.org/10.1364/OE.2.000040.

Rucker, M., R. M. Banta, and D. G. Steyn, 2008: Along-valley structure of daytime thermally driven flows in the Wipp Valley. *J. Appl. Meteor. Climatol.*, **47**, 733–751, https://doi.org/10.1175/2007JAMC1319.1.

Schwartz, M., D. Heimiller, S. Haymes, and W. Musial, 2010: Assessment of offshore wind energy resources for the United States. National Renewable Energy Laboratory Tech. Rep. NREL/TP-500-45889, 96 pp.

Shaw, W. J., J. K. Lundquist, and S. J. Schreck, 2009: Research needs for wind resource characterization. *Bull. Amer. Meteor. Soc.*, **90**, 535–538, https://doi.org/10.1175/2008BAMS2729.1.

Smirnova, T. G., J. M. Brown, and S. G. Benjamin, 1997: Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Mon. Wea. Rev.*, **125**, 1870–1884, https://doi.org/10.1175/1520-0493(1997)125<1870:PODSMC>2.0.CO;2.

——, ——, and ——, 2000: Validation of long-term precipitation and evolved soil moisture and temperature fields in MAPS. Preprints, *15th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 43–46.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Thiébaux, J., E. Rogers, W. Wang, and B. Katz, 2003: A new high-resolution blended real-time global sea surface temperature analysis. *Bull. Amer. Meteor. Soc.*, **84**, 645–656, https://doi.org/10.1175/BAMS-84-5-645.

Weissmann, M. D., G. J. Mayr, R. M. Banta, and A. Gohm, 2004: Observations of the temporal evolution and spatial structure of the gap flow in the Wipp Valley on 2 and 3 October 1999. *Mon. Wea. Rev.*, **132**, 2684–2697, https://doi.org/10.1175/MWR2817.1.

White, A. B., and Coauthors, 2007: Comparing the impact of meteorological variability on surface ozone

during the NEAQS (2002) and ICARTT (2004) field campaigns. *J. Geophys. Res.*, **112**, D10S14, https://doi.org/10.1029/2006JD007590.

Wilczak, J. M., and Coauthors, 2015: The Wind Forecast Improvement Project (WFIP): A public–private partnership addressing wind energy forecast needs. *Bull. Amer. Meteor. Soc.*, **96**, 1699–1718, https://doi.org/10.1175/BAMS-D-14-00107.1.

Wolfe, D. E., and Coauthors, 2007: Shipboard multisensory merged wind profiles from the New England Air Quality Study 2004. *J. Geophys. Res.*, **112**, D10S15, https://doi.org/10.1029/2006JD007344.

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.