Offshore wind energy forecasting sensitivity to sea surface temperature input in the Mid-Atlantic

Stephanie Redfern[1], Mike Optis[1], Geng Xia[1], and Caroline Draxl[1]

[1]National Renewable Energy Laboratory, Golden, Colorado, USA

Correspondence: Stephanie Redfern (stephanie.redfern@nrel.gov)

Abstract. As offshore wind farm development expands, accurate wind resource forecasting over the ocean is needed. One important yet relatively unexplored aspect of offshore wind resource assessment is the role of sea surface temperature (SST). Models are generally forced with reanalysis data sets, which employ daily SST products. Compared with observations, significant variations in SSTs that occur on finer time scales are often not captured. Consequently, shorter-lived events such as sea breezes and low-level jets (among others), which are influenced by SSTs, may not be correctly represented in model results. The use of hourly SST products may improve the forecasting of these events. In this study, we examine the sensitivity of model output from the Weather Research and Forecasting Model (WRF) 4.2.1 to ~~two~~ different SST products~~—~~. We first evaluate three different data sets: the Multiscale Ultrahigh Resolution (MUR25) SST analysis, a daily, ~~spatially coarser resolution data set (0.25° x 0.25° resolution product;~~ the Operational Sea Surface Temperature and Ice Analysis ~~, or~~ (OSTIA), ~~and an hourly, spatially finer resolution product~~(a daily, 0.054° x 0.054° resolution product; and SSTs from the Geostationary Operational Environmental Satellite 16 ~~, or~~ (GOES-16). ~~We find that~~, an hourly, 0.02° x 0.02° resolution product. GOES-16 is not processed at the same level as OSTIA and MUR25; therefore, the product requires gap-filling using an interpolation method to create a complete map with no missing data points. OSTIA and GOES-16 SSTs validate markedly better against buoy observations than MUR25, so these two products are selected for use with model simulations while MUR25 is at this point removed from consideration. We run the model for June and July of 2020 and find that for this time period, in the Mid-Atlantic, although OSTIA SSTs overall validate better against in situ observations taken via a buoy array in the area, the two products result in comparable hub-height (140-m) wind characterization performance on monthly time scales. Additionally, during ~~flagged events~~ hours-long flagged events (< 30 h each) that show statistically significant wind speed deviations between the two simulations, ~~the GOES-16-forced simulation outperforms that forced by OSTIA~~both simulations once again demonstrate similar validation performance (differences in bias, earth mover's distance, correlation, and root mean square error on the order of 10^{-1} or less), with GOES-16 winds validating nominally better than OSTIA winds. With a more refined GOES-16 product, which has been not only gap-filled but also assimilated with in situ SST measurements in the region, it is likely that hub-height winds characterized by GOES-16-informed simulations would definitively validate better than those informed by OSTIA SSTs.

1   Introduction

The United States Atlantic coast is a development site for upcoming offshore wind projects. There are 15 leasing areas located throughout the Atlantic Outer Continental Shelf, where a number of offshore wind farms are planned to be developed (Bureau of Ocean Energy Management, 2018). Therefore, characterizing offshore boundary layer winds in the region has risen in
35   importance. Accurate forecasting will provide developers with a better understanding of local wind patterns, which can inform wind farm planning and layout decisions (Banta et al., 2018). Additionally, improved weather prediction will allow for real-time adjustments of turbine operation to increase their operating efficiency and protect them against unnecessary wear and tear (Gutierrez et al., 2016, 2017; Debnath et al., 2021).

The Mid-Atlantic Bight (MAB) is an offshore cold pool region spanning the eastern United States coast from North Carolina
40   up through Cape Cod, Massachusetts, and it overlies the offshore wind leasing areas. The cold pool forms during the summer months, when the ocean becomes strongly stratified and the thermocline traps colder water near the ocean floor. During the transition to winter, as sea surface temperatures (SSTs) drop, the stratification weakens and the cold pool breaks down. Thus, the cold pool generally persists from the spring through the fall. Southerly winds that drive surface currents offshore will result in coastal upwelling of this colder water. And, at times, strong winds associated with storm development can mix the cold pool
45   upward, cooling the surface and influencing near-surface temperatures and winds (Colle and Novak, 2010; Chen et al., 2018; Murphy et al., 2021).

Accurate representation of the MAB in forecasting models is important because SSTs are closely tied to offshore winds. Horizontal temperature gradients between land and the ocean, as well as vertical temperature gradients over the ocean—which can form, for example, when SSTs are anomalously cold, as with the MAB—help define offshore airflow. In particular, vari-
50   ations in temperature can lead to or impact short-lived offshore events occurring on hourly time scales, such as sea breezes and low-level jets (LLJs). Sea breezes are driven by the land-sea temperature difference, which, if strong enough (around 5°C or greater), can generate a circulation between the water and the land (Stull et al., 2015). With a relatively colder ocean, as during summer months, this leads to a near-ground breeze blowing landward, with a weak recirculation toward the ocean aloft (Miller et al., 2003; Lombardo et al., 2018). Similarly, the near-surface horizontal and air-sea temperature differences
55   dictate the strength of stratification over the ocean. Studies have found a robust link between atmospheric stability and LLJ development, so accurately representing SSTs is key to modeling near-surface stability and, accordingly, LLJs (Gerber et al., 1989; Kaellstrand, 1998; Kikuchi et al., 2020; Debnath et al., 2021). ~~Both of these phenomena~~ A WRF analysis of the region conducted by Aird et al. (2022) found that in the MAB leasing areas specifically, LLJs, occurred on 12% of the hours in June,

2010-2011. Both LLJs and sea breezes can affect individual wind turbine and whole farm operation, so forecasting them correctly can improve power output and turbine reliability (Nunalee and Basu, 2014; Pichugina et al., 2017; Murphy et al., 2020; Xia et al., 2021).

Typical climate and weather model initialization and forcing inputs are reanalysis products, such as ERA5 and ~~MERRA2~~MERRA-2, which are global data sets that assimilate model output with observations to create a comprehensive picture of climate at each time step considered (Gelaro et al., 2017; Hersbach et al., 2020). These data sets primarily include global SST products that are produced at lower temporal and spatial resolutions than what can be available via regional, geostationary satellites. These coarser-resolution data sets, therefore, do not capture observed hourly and, in many cases, diurnal fluctuations in SSTs, which may influence their ability to properly force sea breezes and LLJs. Some preliminary comparisons between weather simulations, forced with different SST products, indicate that this particular input can have a significant impact on modeled offshore wind speeds (Byun et al., 2007; Chen et al., 2011; Dragaud et al., 2019; Kikuchi et al., 2020).

Few studies have examined the impact of finer-temporal resolution SST products specifically on wind forecasting, and to the authors' knowledge, none so far have focused on the Mid-Atlantic. There have been studies looking at numerical weather prediction model (NWP) sensitivity to SST, but they have considered other regions or different, often coarser spatial and temporal resolution, products ~~(Chen et al., 2011; Park et al., 2011; Shimada et al., 2015; Dragaud et al., 2019; Kikuchi et al., 2020)~~ (Chen et al., 2011; Park et al., 2011; Shimada et al., 2015; Dragaud et al., 2019; Kikuchi et al., 2020; Li et al., 2021). In this article, we explore the effects of forcing the Weather Research and Forecasting Model (WRF), a ~~numerical weather prediction model~~ NWP used for research and operational weather forecasting, with different SST data sets characterized by different spatial and temporal resolutions, in the Mid-Atlantic region during the summer months. Specifically, we address differences in model performance on monthly time scales and then contrast characterization effectiveness during shorter wind events. Section 2 lays out the data, model setup, and methods used in this study. Section 3 explains the findings of our simulations, and Section 4 explores their implications. Finally, Section 5 summarizes the intent of the study as well as its findings.

2 Methods

We ~~run two model simulations with identical setups,~~ first validate three different SST data sets against observations taken at an array of buoys off the Mid-Atlantic coast during the months of June and July, 2020. Following validation, we select two of the three data sets for use as inputs to two different model simulations in the MAB region, which are identically configured aside from the ~~input SST data, off the Mid-Atlantic coast for June and July of 2020.~~ SST data. August is not considered due to data availability constraints at the time of the study. The output data are compared with in situ measurements taken at buoys (SSTs) and floating lidars (winds) in the region. A 140-m hub height is assumed, based on the analysis of regions with moderate wind resource by Lantz et al. (2019). We evaluate performance primarily via a set of validation metrics calculated on monthly time scales. We then flag specific events during which the model generally captures regional winds, but output from the two simulations deviate significantly (defined in this study as one or more standard deviations from their mean differences) from one another. Again, validation analysis is performed for these periods.

Table 1. Buoy and Lidar Locations

| Buoy or Lidar	Name	Latitude (°)	Longitude (°)	SST Depth (m)	Owner
Lidar & Buoy	E05	40.1614	-72.7396	0.8	NYSERDA
Lidar & Buoy	E06	39.6273	-73.4123	0.8	NYSERDA
Lidar & Buoy	ASOW-6	39.2717	-73.8892	1	Atlantic Shores
Buoy	44008	40.496	-69.250	1	NDBC
Buoy	44017	40.693	-72.049	1.5	NDBC
Buoy	44025	40.251	-73.164	1.5	NDBC
Buoy	44065	40.369	-73.703	1.5	NDBC
Buoy	44075	40.363	-70.883	1	Ocean Observatories Initiative
Buoy	44076	40.137	-70.775	1	Ocean Observatories Initiative
Buoy	44077	39.940	-70.883	1	Ocean Observatories Initiative
Buoy	44091	39.768	-73.770	0.46	U.S. Army Corps of Engineers
Buoy	44097	40.967	-71.126	0.46	Ocean Observatories Initiative

2.1 In Situ and Lidar Data

This study makes use of both SST and wind profile observational data for model validation. SSTs are provided by the National Buoy Data Center (NBDC) at several locations along the Mid-Atlantic Coast, as listed in Table 1 and shown in Fig. 1. Buoy data located at the Atlantic Shores Offshore Wind ASOW-6 location are also used (Fig. 1b). Wind data have been taken from the Atlantic Shores Offshore Wind floating lidar and the two New York State Energy Research & Development Authority (NYSERDA) floating lidars, whose locations are listed in Table 1. These lidars provide wind speed and wind direction at 10-minute intervals from either 10 m (Atlantic Shores) or 40 m (NYSERDA) up through 250 m above sea level. There are periods of missing data for all buoys and lidars.

2.2 Model Setup

WRF Version 4.2.1 is the NWP employed in this study (Powers et al., 2017). WRF is a fully compressible, non-hydrostatic model that is used for both research and operational applications. Our model setup, including key physics and dynamics options, are outlined in Table 2.

The study area spans the majority of the MAB, with the nested domain (grid spacing of 2 km x 2 km) running from the mid-Virginia coast to the south up through Cape Cod to the north (Fig. 1a).
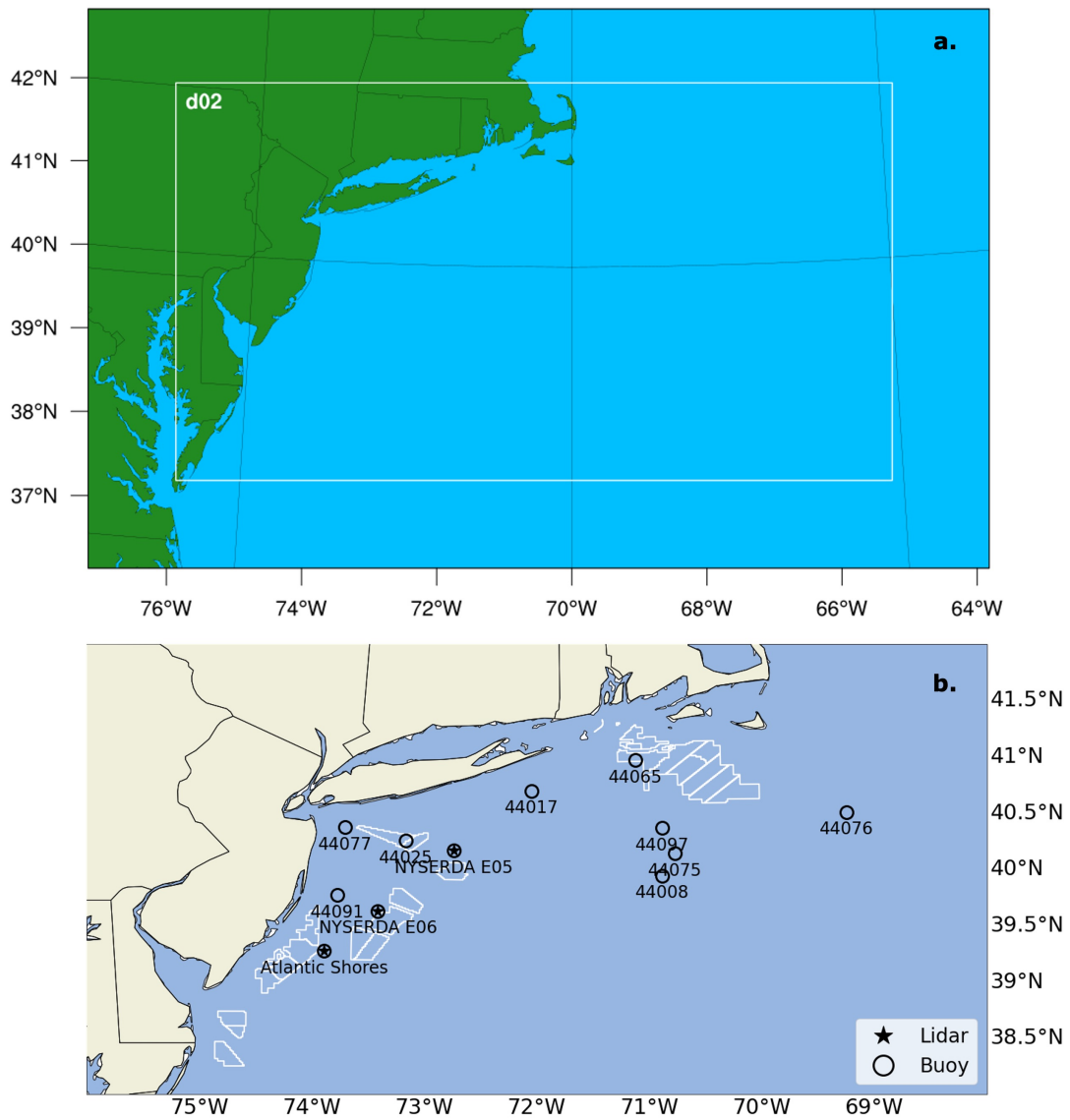
Figure 1. The ~~selected study area~~ WRF domains (a.) and the nested domain only ~~)~~—with markers indicating the lidars ~~depicted by~~ (stars) and buoys ~~by~~ (hollow circles) that were used in this study (b.). The Atlantic Shores location is the site of both a buoy and a lidar. ~~Leasing~~ Planning areas are outlined in ~~brown and~~ white.

Table 2. ~~WRF Physics Options Used in This Study~~Model Setup

WRF Parameter	Selection
Number of Domains	2
Domain Resolution	6 km (Parent), 2 km (Nest)
Output Time Resolution	5 minutes
Vertical Levels	61
Reanalysis Data	ERA5 (Hersbach et al., 2020)
Microphysics	Ferrier (Schoenberg Ferrier, 1994)
Radiation Scheme	RRTMG (longwave & shortwave) (Iacono et al., 2008)
Planetary Boundary Layer	Nakanishi and Niino (MYNN) ~~, 2006~~ (Nakanishi and Niino, 2006)
Surface Layer Parameterization	MYNN (Nakanishi and Niino, 2009)
Land Surface Scheme	Unified Noah land-surface model (Tewari et al., 2004)
Cumulus Parameterization	Kain-Fritsch (Kain and Fritsch, 1993)
Upper-Level Damping	Rayleigh at 5km depth

2.3 Sea Surface Temperature Data

We compare how well three different SST datasets validate against buoy observations and subsequently select the two best-performing data sets to force our simulations (Table 3). Aside from these different SST product inputs, the rest of the model parameters in the simulations remain identical.

110   We have selected the ERA5 global reanalysis data set to force our simulations. OSTIA is the SST data set native to this product (Hersbach et al., 2020). As such, when included as part of ERA5, OSTIA's resolution has been adjusted to match ERA5's 31 km spatial resolution and hourly temporal resolution. For our simulations, however, we overwrite these SSTs with OSTIA data at its original resolution of 0.05°, MUR25 data at its 0.25° resolution, and GOES-16 data at its resolution of 0.02°.

115   ~~Our coarser-resolution~~ The coarsest-resolution product we consider is the 0.25° Multi-scale Ultra-high Resolution SST analysis (MUR25). MUR25 is a daily product that assimilates observations and model data to render a complete global grid with SSTs. MUR25 has undergone pre- and post-processing and the SSTs in the dataset are foundation temperatures, which are measured deep enough in the water to discount diurnal temperature fluctuations. MUR25 has a spatial resolution of 0.25° x 0.25°, which is significantly coarser than either of the other two SST data sets we evaluate (Chin et al., 2017).

120   Our next SST selection, the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, is a daily global product that combines in situ observations taken from buoys and ships, model output, and multiple remotely sensed SST data sets ~~. Because of this~~(Stark et al., 2007; Donlon et al., 2012). Similar to MUR25, it is a complete product in that it has no missing data. OSTIA has a spatial resolution of 0.05° x 0.05°. Additionally, as a daily product like MUR25, OSTIA provides

Table 3. SST Data Sets Considered in This Study

Parameter	MUR25	OSTIA	GOES-16
Satellite Coverage	Global	Global	Regional
Temporal Resolution	Faily	Daily	Hourly
Spatial Resolution	0.25°	0.054°	0.02°
Processing Level	Gridded & assimilated with in situ observations	Gridded & assimilated with in situ observations	Gridded only
Gap-Filling	Released product is filled	Released product is filled	DINEOF needed

a foundation SST, which is measured deep enough in the water to discount diurnal temperature fluctuations foundation SSTs

(Stark et al., 2007; Donlon et al., 2012; Fiedler et al., 2019).

Our finer-resolution finest-resolution product is taken via GOES-16, which is a geostationary, regional satellite with a spatial resolution of 0.02° x 0.02° (Schmit et al., 2005, 2017). This product does not assimilate its measurements with in situ data. While GOES-16 does not offer global coverage and, therefore, cannot be used for certain world regions, it does cover the Mid-Atlantic Bureau of Ocean Energy Management (BOEM) offshore wind lease areas, which is our region of interest. Because GOES-16 has an hourly resolution, it can capture diurnal changes in SST and, therefore, measures surface temperature (Schmit et al., 2005, 2008).

We have selected the ERA5 global reanalysis data set to force our simulations. OSTIA is the SST data set native to this product (Hersbach et al., 2020). As such, when included as part of ERA5, OSTIA's resolution has been adjusted to match ERA5's 31 km spatial resolution and hourly temporal resolution. For our simulations, however, we overwrite these SSTs with OSTIA data at its original resolution of 0.05and with GOES-16 data at its resolution of 0.02.

Due to the lesser level of processing in the GOES-16 data set, it contains numerous data gaps that must be filled. This missing data arises due to is the result of a post-processing algorithmapplied prior to release of the data set, which flags pixels that fall below a specified temperature threshold, that is applied prior to release of the data set. This filter is in place to remove cloud cover. While this method is effective with regard to its set defined intent, it can also erroneously discard valid pixels that capture the cold water upwelling typical to the MAB region during the warmer months. And, although this cold-pixel filter is also a common practice in global SST data sets (OSTIA), the high level of post-processing applied in those products interpolates over and fills the missing grid cells prior to release.

To gap-fill the GOES-16 data so that we may use it in our application, we employ the Data INterpolating Emperical Orthogonal Function algorithm, or DINEOF, which is an open-source application that applies empirical orthogonal function (EOF) analysis to reconstruct incomplete data sets (Ping et al., 2016). The program was originally designed to specifically gap-fill remotely observed SSTs that contain missing data due to cloud-flagging and removal algorithms (as is the case with the GOES-16 data), and has demonstrated strong results in past studies (Alvera-Azcárate et al., 2005; Ping et al., 2016).

7

We additionally ~~apply to~~ include in the GOES-16 SST data set the sensor-specific error statistics (SSES) bias field that is included with the distributed product. This component accounts for retrieval bias ~~via~~ using a statistical algorithm designed to

150 correct for errors in the SST field. Compared with the SST values alone, the bias-corrected GOES-16 data offsets an inherent warm bias ~~, bringing the values down to be more consistent with those of OSTIA~~ in the raw data.

2.4 Event Selection

We are particularly interested in ~~examining model performance for forecasting~~ understanding how well the model forecasts shorter wind events, as LLJs and sea breezes occur on hourly time scales. ~~Therefore, we~~ We have created a set of parameters

155 that, when met, detect ~~briefer events~~ relatively brief time periods (on the order of hours to days) during which one simulation may be outperforming the other. ~~Once the events have been collected, we~~ , which we then more closely examine ~~them to characterize the wind profiles and SSTs at those times~~ to evaluate differences in their wind profile characterization and SST validation.

For an event to be flagged, it must meet the following criteria:

160 1. Correlation for both models is above 0.5 at hub height, for two of the three lidar locations.

2. Differences in wind speeds between the two models must be greater than one standard deviation from the monthly mean difference.

3. Gaps during which the wind speed difference drops below one standard deviation must not persist for more than 2 hours during a single event.

165 4. Events must last for at least one hour.

This set of event characteristics first acts to filter out periods during which WRF is generally under-performing—possibly due to model shortcomings outside of SST forcing—so that the performance difference in the selected events may be ~~attributed to SSTs~~ with more certainty attributed to SSTs. Then, events are located during which the two simulations forecast statistically significantly different hub-height wind speeds, which persist for a period of time long enough to substantially affect power

170 generation.

2.5 Validation Metrics

To evaluate which simulation performs best during our study period, we calculate sets of validation metrics, as outlined by Optis et al. (2020). Specifically, we look at SSTs and ~~100-m~~ 140-m (hub-height) wind speeds, on both monthly and event time scales. The metrics we calculate are named and defined in Table 4, and they ~~provide~~ give a quantification of the error present

175 in each case.

The bias provides information on the average simulation performance during the evaluation period—specifically, if the model is consistently over- or under-predicting the output variable in consideration. Unbiased root mean square error (RMSE)

Table 4. Error Metrics Used in This Study

Error Metric	Equation
Bias	$\bar{p} - \bar{o}$
Unbiased Root Mean Square Error (cRMSE)	$\left[\frac{1}{N} \sum_{n=1}^{N} [(p_n - \bar{p})(o_n - \bar{o})]^2 \right]^{\frac{1}{2}}$
Square of Correlation Coefficient (R^2)	$\left[\frac{\frac{1}{N} \sum_{n=1}^{N} (p_n - \bar{p})(o_n - \bar{o})}{\sigma_p \sigma_o} \right]^2$
Wasserstein Metric / Earth-Mover's Distance (EMD)	$\sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} d_{ij}$

provides a more nuanced look at the spread of the error in the results. ~~Correlation~~ The square of the correlation coefficient (referred to from here forward as correlation) quantifies how well the simulations' variables change in coordination with those of the observations. And, finally, the Wasserstein Metric, also known as the Earth-Mover's Distance (EMD), measures the difference between the observed and simulated variable distributions.

3   Results

We ~~compare the performance of GOES-16 with OSTIA by validating each data set's SSTs and modeled hub-height winds with in situ observations collected via buoys and lidars in the Mid-Atlantic. First, we focus on the SSTs of each remotely sensed data set and how they compare with~~ first assess how well each SST data set compares with buoy data. Following SST validation, we evaluate the model's wind characterization performance ~~using each input dataset~~ when forced with different SST data. We assume a hub height of ~~100~~ 140 m and compare output winds at this altitude against measurements taken via floating lidars off the Mid-Atlantic coast. Specifically, we analyze monthly performance and then select several shorter periods during June and July ~~of~~ 2020 during which we compare wind characterization accuracy.

~~Overall, we find that on a monthly time frame, OSTIA SSTs validate better than those of GOES-16. Average~~ After evaluation of the SST data sets, only OSTIA- and GOES-16-forced WRF simulations are compared, as during July—one half of our study period—MUR25 SSTs validate significantly worse against buoy measurements than the other two products. OSTIA SSTs validate the best out of the three data sets. Average, simulated hub-height winds ~~at this resolution~~ across monthly periods validate similarly for both ~~products~~ OSTIA and GOES-16. At an event-scale temporal resolution—that is, on the order of hours—GOES-16 ~~outperforms OSTIA for the selected events~~ and OSTIA perform comparably, with GOES-16 marginally outperforming OSTIA.

3.1   Sea Surface Temperature Performance

~~Before narrowing the scope of this study to a comparison between OSTIA and GOES-16 SSTs, we have included in the SST analysis a third data set: the Multi-scale Ultra-high Resolution SST analysis (MUR). The MUR is a daily product that has undergone foundation-level preprocessing (the same level as OSTIA). MUR has a spatial resolution of 0.25° x 0.25°,~~

Table 5. Validation Metrics for Each Remotely Sensed Data Source at Atlantic Shores, NYSERDA E05, and NYSERDA E06 on a 10-min. Output Interval for July 2020.

	Data Source	Bias (°C)	cRMSE (°C)	R^2	EMD
Atlantic Shores	MUR25	0.41	0.90	0.68	0.49
	OSTIA	-0.17	0.71	0.8	0.37
	GOES-16	-0.01	0.76	0.76	0.35
NYSERDA E05	MUR25	0.06	0.80	0.76	0.37
	OSTIA	-0.14	0.59	0.88	0.32
	GOES-16	-0.13	0.63	0.86	0.31
NYSERDA E06	MUR25	0.21	0.83	0.68	0.28
	OSTIA	-0.14	0.61	0.83	0.27
	GOES-16	-0.02	0.64	0.81	0.2

which is significantly coarser than either of the other two SST data sets (Chin et al., 2017). We evaluate SST performance by interpolating the satellite-based products to 10-minute intervals (the in situ data output resolution) and making a brief qualitative assessment followed by the calculation of each of the validation metrics selected for this study calculating and comparing at each buoy validation metrics for the different products.

205　Taking a qualitative look at each product's SST—specifically at Atlantic Shores—as well as the in situ measurements at the buoy, we see that The time series plots in Fig. 2 show that that while GOES-16 tracks the dirunal diurnal cycle seen by observations, the other two products do not (Fig. 2 the highlighted time frames in the figure are wind events that are discussed in Section 3.3). Despite this feature, however, there are exist a number of times during periods each month when GOES-16 does not accurately capture observed dips in temperature. Moreover And, during many of these periods times, the daily data

210　sets, though missing the nuance of GOES-16, better predict represent the colder SSTs. Despite this, at Atlantic Shores,

A comparison of all three products show markedly poorer validation against the floating lidar buoys, for at least two metrics at each buoy, by MUR25 as compared with GOES-16 does validate better than both OSTIA and MUR for both June and July at this location and OSTIA during July 2020 (Table 5). Because July 2020 is one half of our study period, we remove MUR25 from further analysis.

215　Time series of MUR, OSTIA, GOES-16, and in situ measurements of SST at the Atlantic Shores buoy for June (top) and July (bottom).
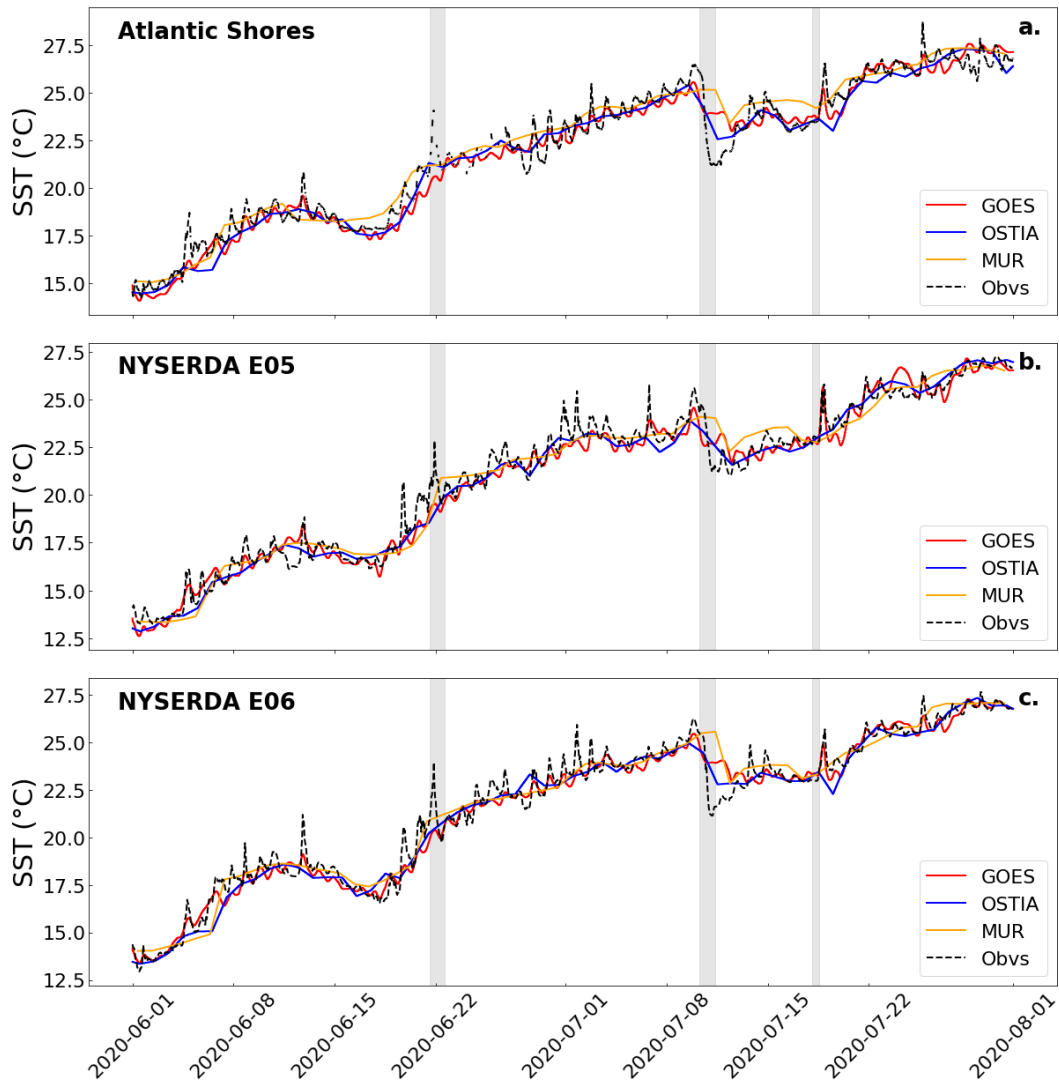
Figure 2. Time series of MUR25, OSTIA, GOES-16, and in situ measurements of SST at the Atlantic Shores buoy for June and July at Atlantic Shores (a.), NYSERDA E05 (b.), and NYSERDA E06 (c.). Specific wind events that are evaluated in Section 3.3 are highlighted in gray.

~~To broaden the scope of SST validation, we take an even more nuanced look at SST product performance by calculating validation metrics at all the buoys listed in Table 1, instead of just Atlantic Shores. Because MUR markedly underperforms compared with the other two data sources, it is not included in this (and further) analysis.~~

220   Over the course of June and July combined, looking across the entire buoy array, OSTIA overall outperforms GOES-16, as shown in Fig. 3. Both products have a relatively strong cold bias (between -0.1°C and -0.25°C) compared with observations at the three lidar locations. GOES-16 ~~has~~ presents a negative bias at five additional buoys, and OSTIA has a negative bias at

11
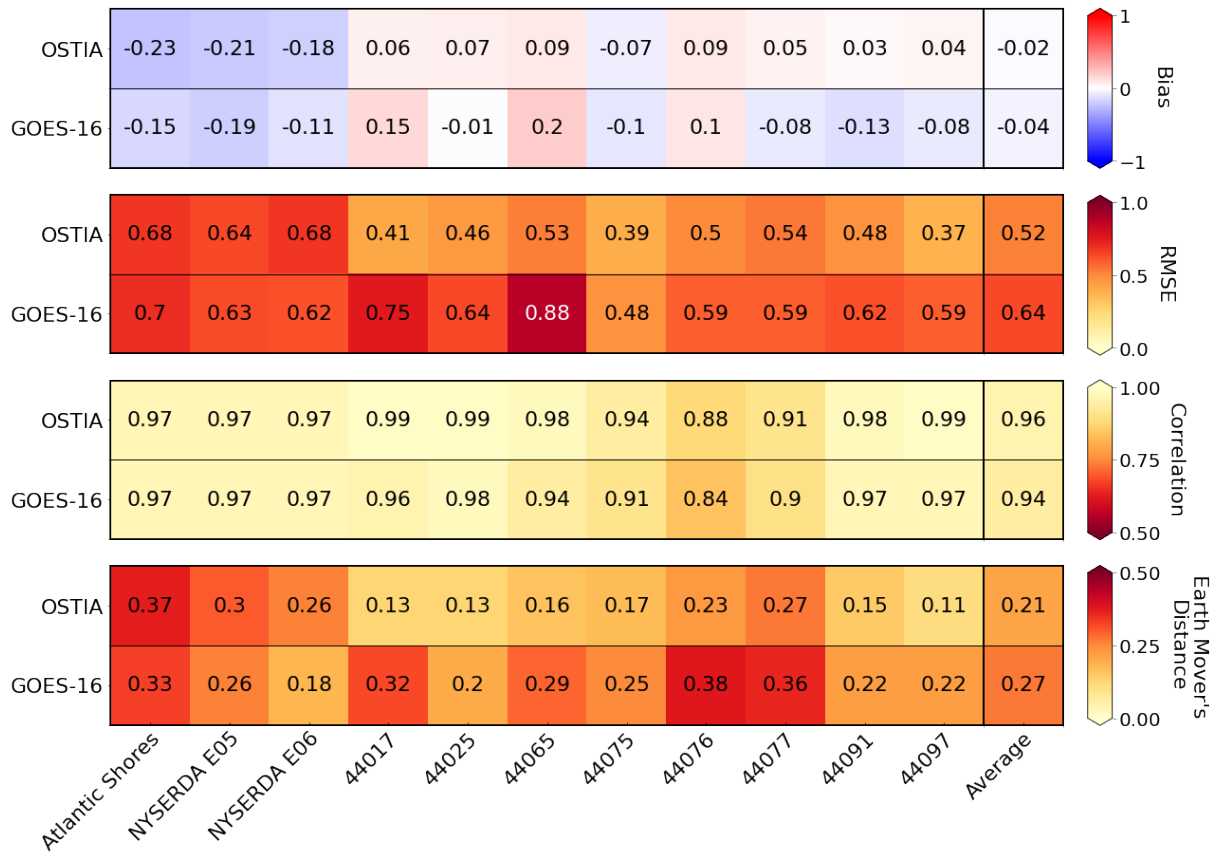
Figure 3. Mean bias, RMSE, correlation, and EMD for GOES-16 and OSTIA SSTs at each buoy location shown, combined for June and July, along with the average metrics over all sites for each product.

one other buoy. Both SST products display warm biases at buoys 44017 (off the northeast coast of Long Island) and 44076 (the farthest offshore location considered, southeast of Cape Cod). Although GOES-16 follows the diurnal cycle rather than representing only the daily average SSTs, ~~it still does not correlate with observations as well as OSTIA.~~ as is the case with OSTIA, on average the correlation across all sites is comparable between the two SST products. RMSE is higher for GOES-16 at every buoy (with an average 0.64°C, compared with 0.52°C for OSTIA), and EMD for GOES-16 is higher than that for OSTIA, at every location except the floating lidars (an average of 0.27, compared with 0.21 for OSTIA).

~~Specifically within the leasing areas (NYSERDA E05 & E06, Atlantic Shores, and buoys 44025, 44017, and 44065), OSTIA still delivers as the strong product, although~~ Although OSTIA on average, across all buoys surveyed, better represents measured SSTs, GOES-16 ~~at these sites alone delivers more promising average metrics than the average across all buoys~~ performs better on average across only the floating lidar sites—Atlantic Shores, NYSERDA E05, and NYSERDA E06 (Table 6).

12

Table 6. Average performance metrics for OSTIA and GOES-16 SSTs across the three floating lidar sites for June and July 2020.

| Metric	OSTIA	GOES-16
Bias	-0.21	-0.15
RMSE	0.67	0.65
Correlation	0.97	0.97
EMD	0.31	0.27

3.2 Monthly Wind Speeds

The probability distribution functions (PDFs) of hub-height wind speeds at each lidar show that WRF, on a monthly time scale and with a 10-minute output resolution, generally captures the shape of the observed wind speed distribution at each lidar (Fig. 4), which ~~indicates~~ suggests that the model itself is performing as it should. The wind speed distributions for each simulation maintain an even closer similarity in shape to one another, which ~~helps highlight~~ highlights the biases directly related to the particular SST data set being used. A box plot of wind speeds across the entire domain ~~,~~ for each simulation and for both months ~~, shows~~ (not shown) indicates that although GOES-16 and OSTIA present near-identical average hub-height wind speeds, GOES-16 ~~has~~ winds show a greater spread than OSTIA. Additionally, in both simulations, whole-domain winds in ~~July~~ June tend to be significantly faster than ~~June winds .~~ July winds (Fig. 4).

A comparison between each of the two simulations of monthly hub-height wind speed bias and correlation shows that ~~although they perform overall similarly , GOES-16 tends to forecast 100-m wind speeds at each lidar with slightly more accuracy than OSTIA (e.g.,~~ they perform quite similarly during both June and July (bias and correlation shown in Fig. 5). ~~Performance at other heights varies, but, in general, GOES-16 continues to outperform OSTIA moving higher above sea level, which will be relevant for wind resource forecasting for larger offshore wind turbines. At each location, both products tend to~~ Both products over-predict ~~winds at hub height and below~~ wind speeds, except at the two NYSERDA lidars during June. The correlation of each product's forecasted winds ~~at hub height with observations~~ with observations, from 100 m ASL and higher, is above 0.65 ~~for all times and locations~~ at all three lidars during both June and July.

A map of the domain ~~showing~~ displaying the June and July average wind speed differences between the GOES-16 and OSTIA-forced simulations, overlaid by wind barbs ~~depicting~~ indicating the average GOES-16 wind speeds, is shown in Fig. 6. In general, wind speeds deviate from each other only by small amounts on a monthly time scale. The results show maximum average wind speed differences between the two simulations of up to 0.25 m s^{-1} for each month.

3.3 Event-Scale Wind Speeds

Using the event selection algorithm detailed in Section 7, we locate three periods—one in June and two in July—during which ~~100-m~~ 140-m wind speeds in the OSTIA and GOES-16 simulations differ from one another by statistically significant amounts
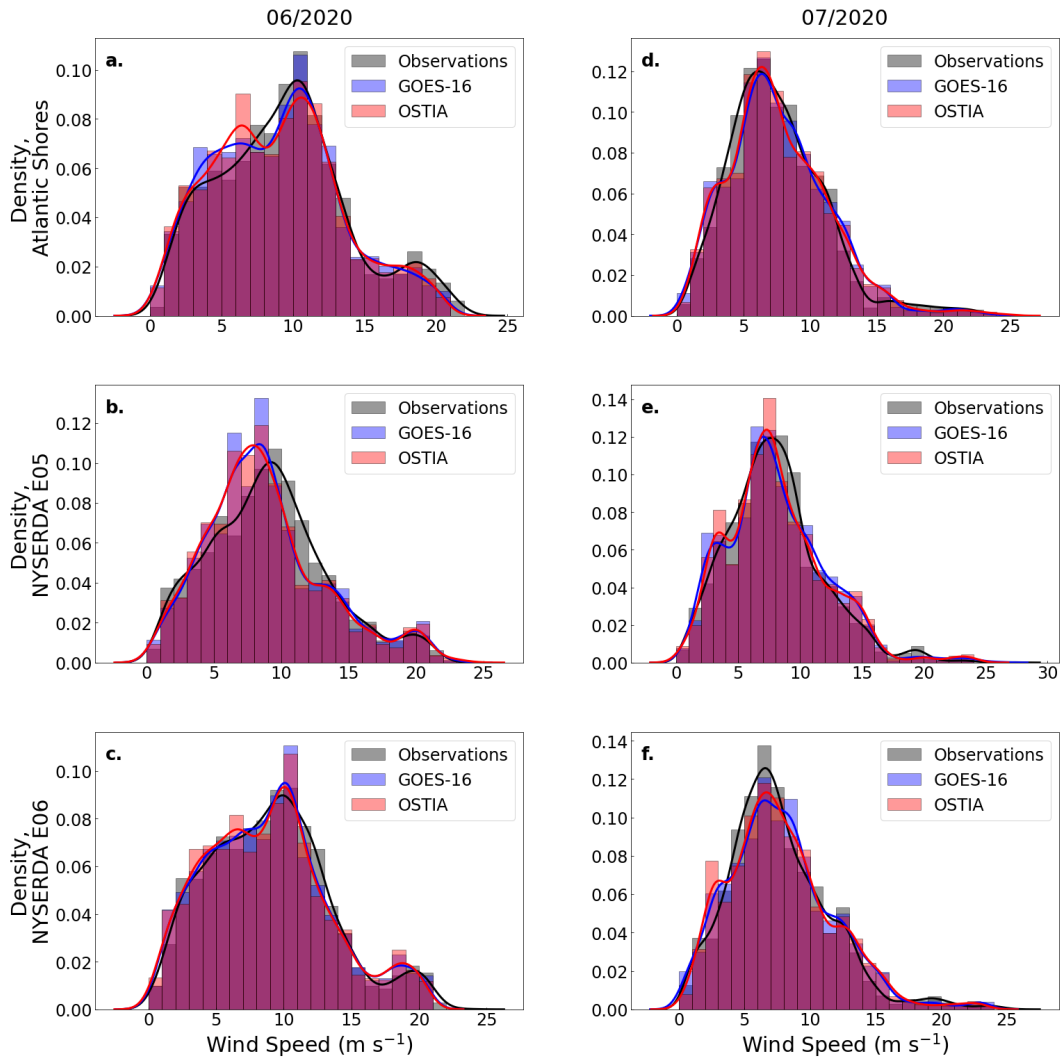
Figure 4. PDFs of ~~100-m~~ 140-m wind speeds at each lidar location, for June (left) and July (right), taken via observations (gray), GOES-16-forced model output (blue), and OSTIA-forced model output (red).

and still both validate relatively well against observations (r^2 > ~~0.5~~ 0.6 at 2 or more lidars), per the criteria outlined in Section 7. These events are listed in Table 7. Lidars at which there ~~is little observational data for~~ are relatively large observational data gaps during these time periods (> 20%) are not considered. Therefore, data from only two locations for the June event and the second July event are ~~considered~~ used. For all three cases, at each lidar being considered, GOES-16 delivers overall ~~better predictions of 100-m~~ stronger characterization of 140-m wind speeds than OSTIA. Validation metrics vary at different heights, so vertical profiles of bias, r^2, RMSE, and EMD for each event can be found in Appendix A.

14
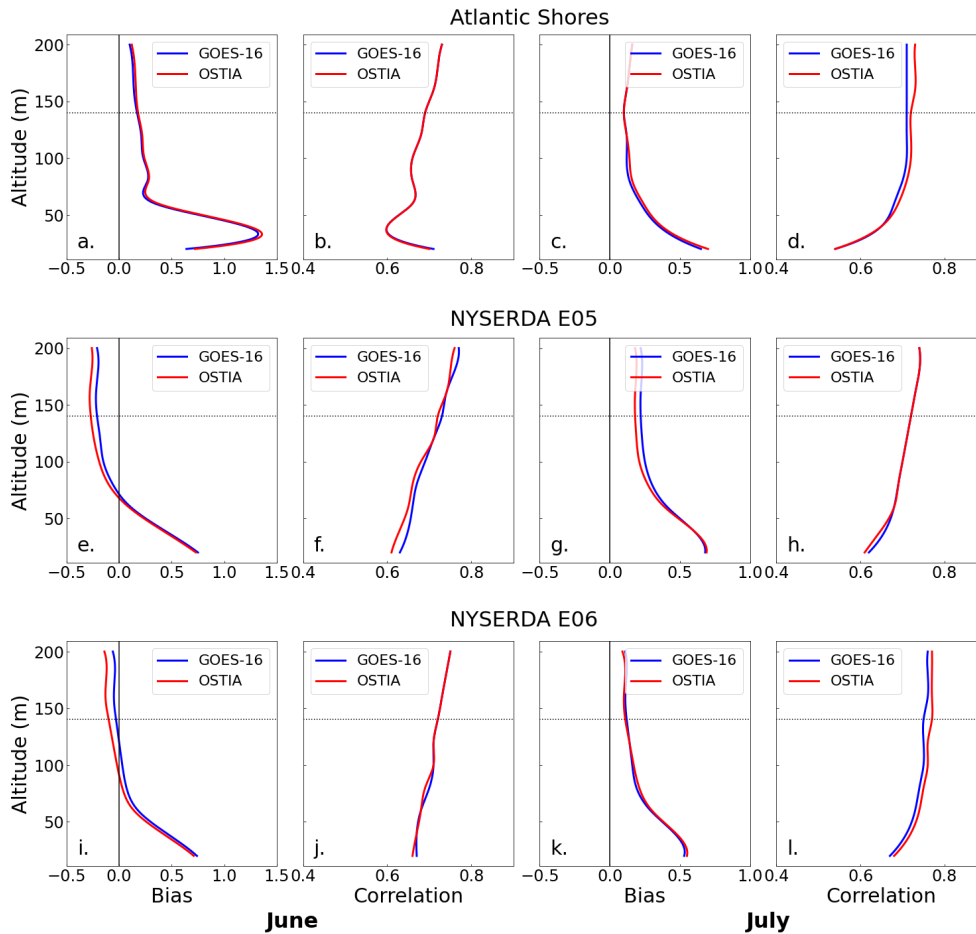
Figure 5. Modeled hub-height wind speed bias (left) and correlation (right) for simulations forced with GOES-16 (blue) and OSTIA (red) SSTs during June and July 2020 at the Atlantic Shores lidar (a, b, c, d), the NYSERDA E05 lidar (ee, df, g, h), and the NYSERDA E06 lidar (ei, fj, k, l).

3.3.1 June 21-22, 2020 Event

We have identified an event that meets our criteria beginning on 06-21-2020 at 13:40:00 and ending on 06-22-2020 15:20:00. During this time period, offshore winds near the coast are southerly, with a tendency to follow the coastline as they rotate around a high-pressure system southeast of New Jersey.

Differences in SSTs between the two simulations do not correspond with significant 140-m wind speed differences at the three lidars; however, in other areas of the domain, including a region planned for development just south of Rhode Island (Fig. 1), these differences are larger, with magnitudes reaching over 2 m s^{-1} (Fig. 7). There are not significant SST differences noted within that the region, although land surface temperatures may contribute to a stronger temperature gradient.

15
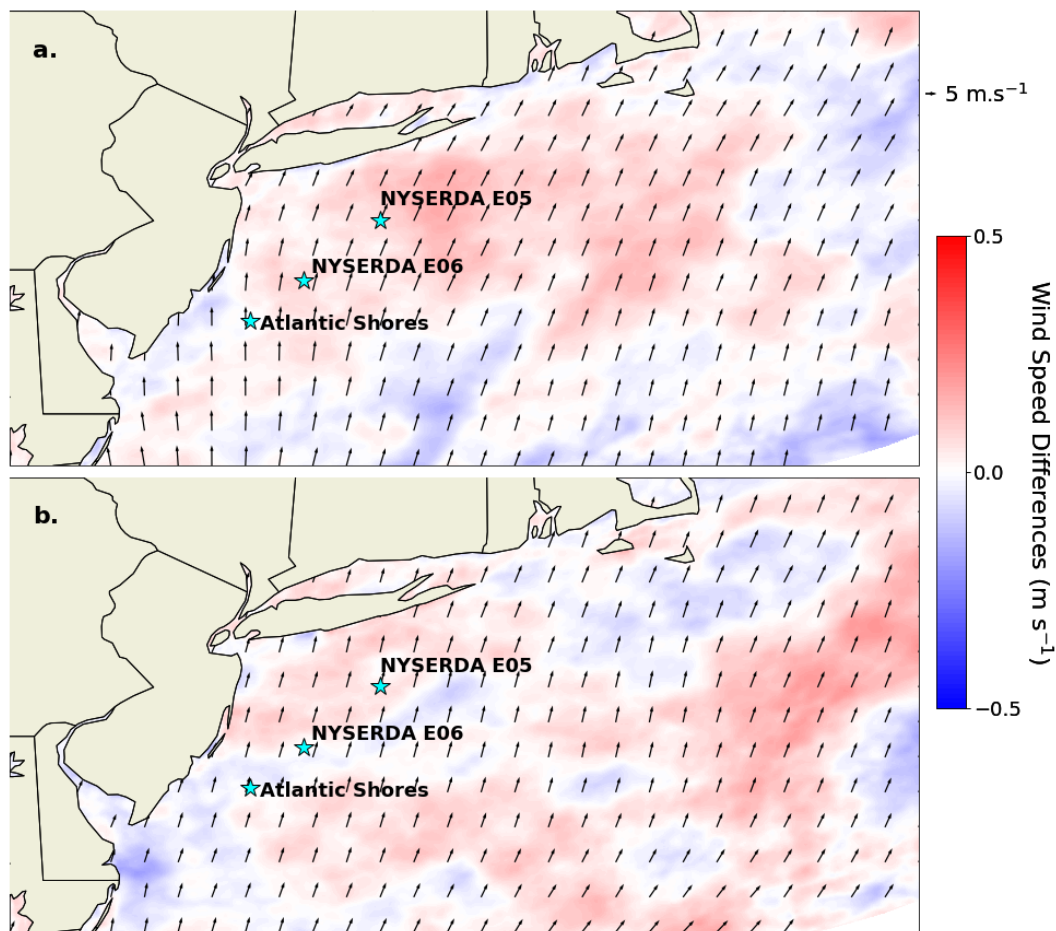
Figure 6. Modeled hub-height (~~100 m~~140-m) wind speed differences, GOES-16 - OSTIA, for June (a.) and July (b.).

Table 7. Weather Events Selected for Evaluation

Start Date	End Date	Event Length
06-21-2020 13:40:00	06-22-2020 15:20:00	25h 40m
07-10-2020 06:25:00	07-11-2020 09:00:00	26h 25m
07-18-2020 02:05:00	07-18-2020 14:10:00	12h 5m

A planar depiction of 140-m wind speeds across the entire domain shows that, although the mean of the magnitude of the differences is roughly 0.25 m s^{-1}, they vary by location across the region. The maximum difference in average wind speeds
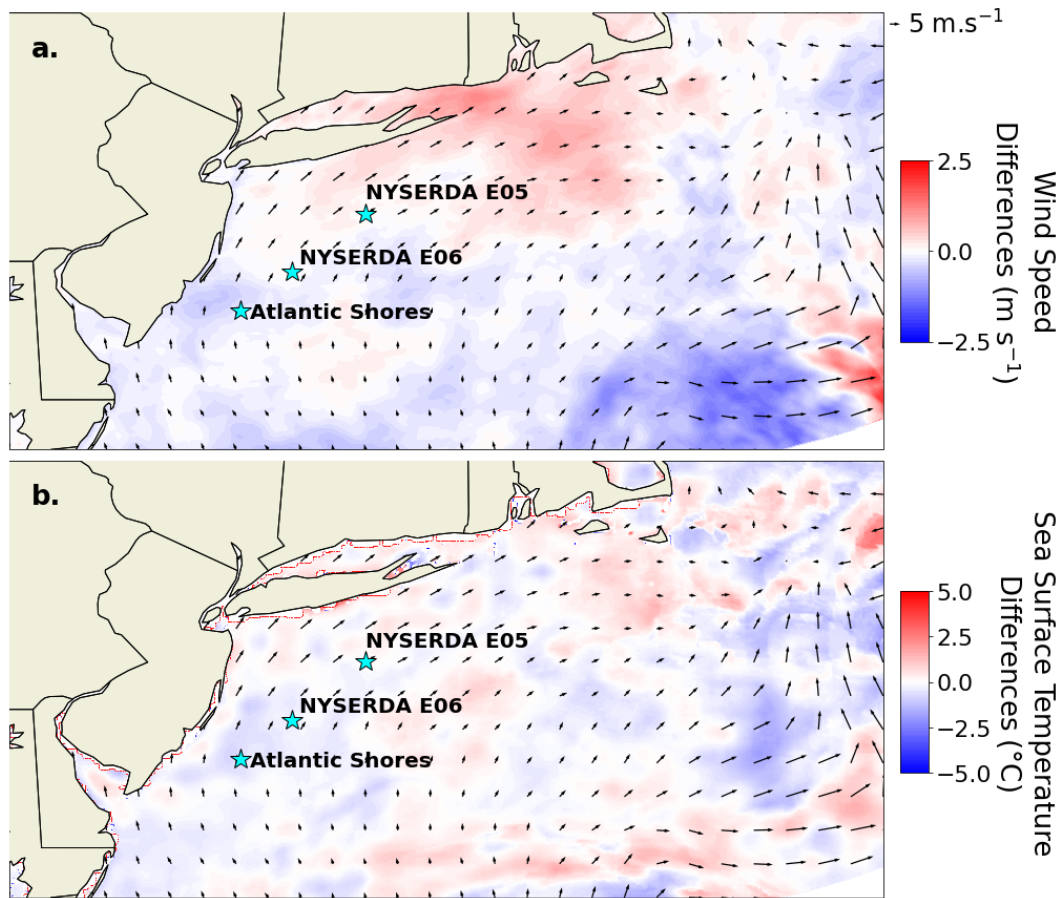
Figure 7. Differences (GOES-16 - OSTIA) in average 140-m wind speeds (a.) and SSTs (b.) over the entire domain for the June 21-22 event.

during this event is actually 2.25 m s^{-1}. This is noteworthy, as wind speed differences of this size have significant implications regarding power generation.

275   Model output from the two simulations ~~captured~~ captures the observed trend of a wind speed increase—but model ~~output wind speeds deviated significantly from one~~ outputted wind speeds deviate significantly from each another leading into a first ramping event, during a second ramp event, and once wind speeds began to stabilize (Fig. 8). Observational data are only available at the two NYSERDA lidars during these periods. The Atlantic Shores lidar ~~exhibited~~ has significant data gaps, so validation at this location is not conducted.

280   ~~Validation metrics for the event~~ Event validation metrics at both lidars ~~indicated~~ indicate stronger performance by the GOES-16 simulation than OSTIA, particularly at vertical levels associated with typical offshore turbine hub heights. Validation metrics at hub height (~~100~~ 140 m) for the two ~~stimulations~~ simulations at NYSERDA E05 and NYSERDA E06 are shown in Fig. 9. Averaged correlation between the two sites is ~~0.75~~ 0.78 for GOES-16 and ~~0.72~~ 0.76 for OSTIA. The bias at each site is comparable between the two simulations, at -0.26 m s^{-1} and -0.24 m s^{-1} for OSTIA and GOES-16, respectively. Model

17
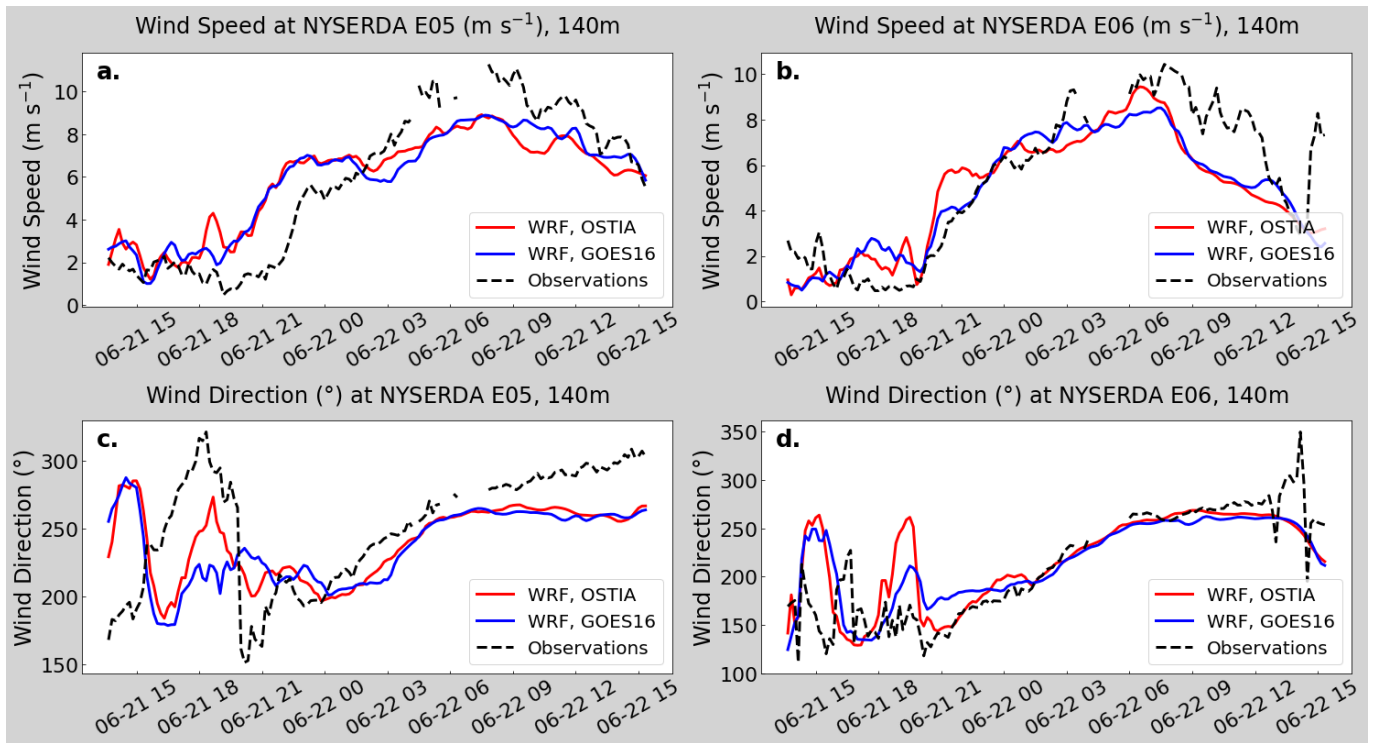
Figure 8. Hub-height (140-m) wind ~~speed~~ speeds and wind ~~direction~~ directions at the NYSERDA E05 (a, c) and NYSERDA E06 (b, d) lidars during the June 21-22 event. Atlantic Shores is not shown due to a lack of observational data.

285 performance across the vertical should be taken into consideration if the rotor-equivalent wind speed is used to calculate power generation during extreme shear events as well as when looking at turbines of varying hub heights. Of note is that both RMSE and EMD are lower for GOES-16 at all heights (Appendix A).

~~The mean average difference in wind speeds throughout the domain is nominally zero. However, we can see via a planar depiction of wind speedsacross the entire domain that, although the average difference is roughly 0.01~~

290 **3.3.2 July 10-11, 2020 Event**

We have next flagged an event that occurred between 07-10-2020 06:25:00 and 07-11-2020 09:00:00. Synoptically, Tropical Storm Fay was observed to be moving in a southerly direction through the region during this time. High wind speeds, which peaked at 07-10-2020 18:00:00, may be attributed to this storm. Average wind directions also reflect the storm path (Fig. 10). Differences in average wind speed between the two simulations peak at only 1.35 m s^{-1}—which is around 1 m s^{-1} ~~,~~

295 ~~the differences vary by location across the region. The maximum difference in~~ less than the maximum difference during the June event. The mean difference in the magnitudes of average wind speeds ~~during this event is actually 2.27 m s^{-1}. This is~~

18
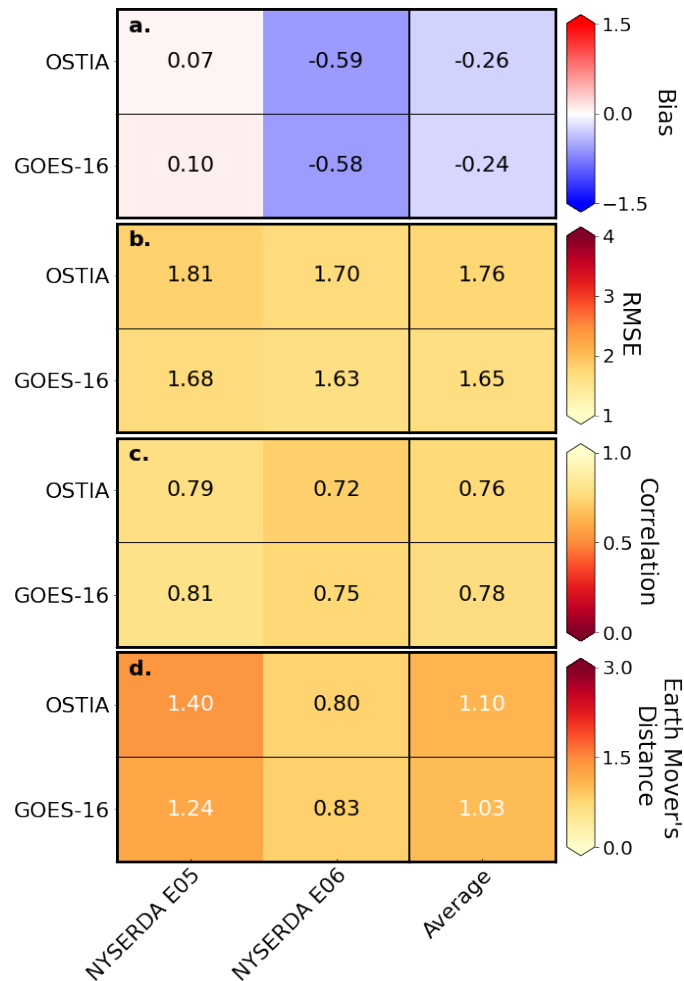
Figure 9. Hub-height wind speed validation metrics at the NYSERDA E05 and NYSERDA E06 lidars during the June 21-22 event: bias~~(a)~~, ~~correlation (b),~~ RMSE~~(c)~~, correlation, and EMD~~(d)~~.

~~noteworthy, as wind speed differences of this magnitude have significant implications regarding power generation.~~ across the domain is 0.2 m s^{-1}.

~~During this event, offshore winds near the coast are southerly, with a tendency to follow the coastline as they rotate~~ ~~around a high-pressure system southeast of New Jersey. The differences in SST do not effect significant 100-m wind speed~~ ~~differences between the two simulations at the three lidars; however, in other areas of the domain, including a region planned~~ ~~for development just south of Rhode Island~~ With the winds blowing in a south-southeasterly direction across the leasing area, and the differences in GOES-16 showing a relatively stronger temperature gradient across that area in the direction of the wind, the GOES-16 simulation outputs stronger wind speeds compared with OSTIA (Fig. ~~1~~), ~~these differences are quite a bit larger~~

300

305 10). This is reflected in the overall bias seen at the three lidars, although both simulations overall underestimate wind speeds during this time (Fig. ~~7~~11).
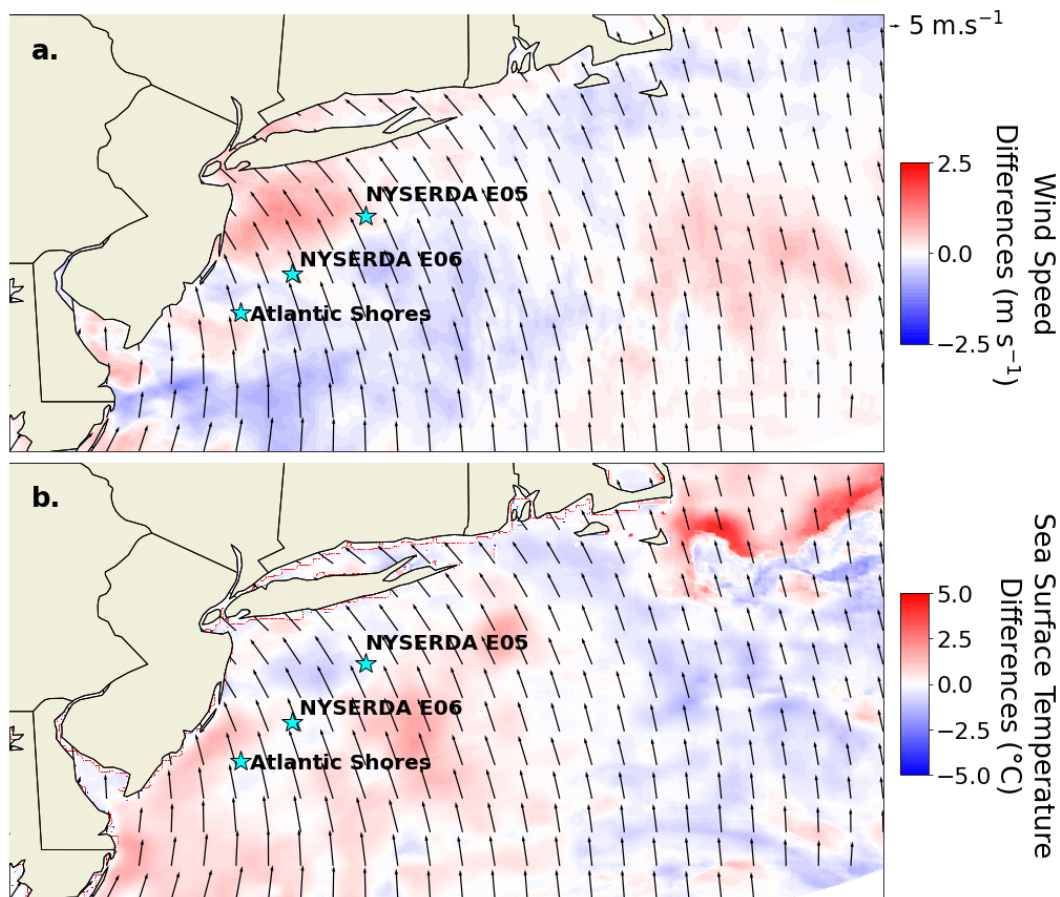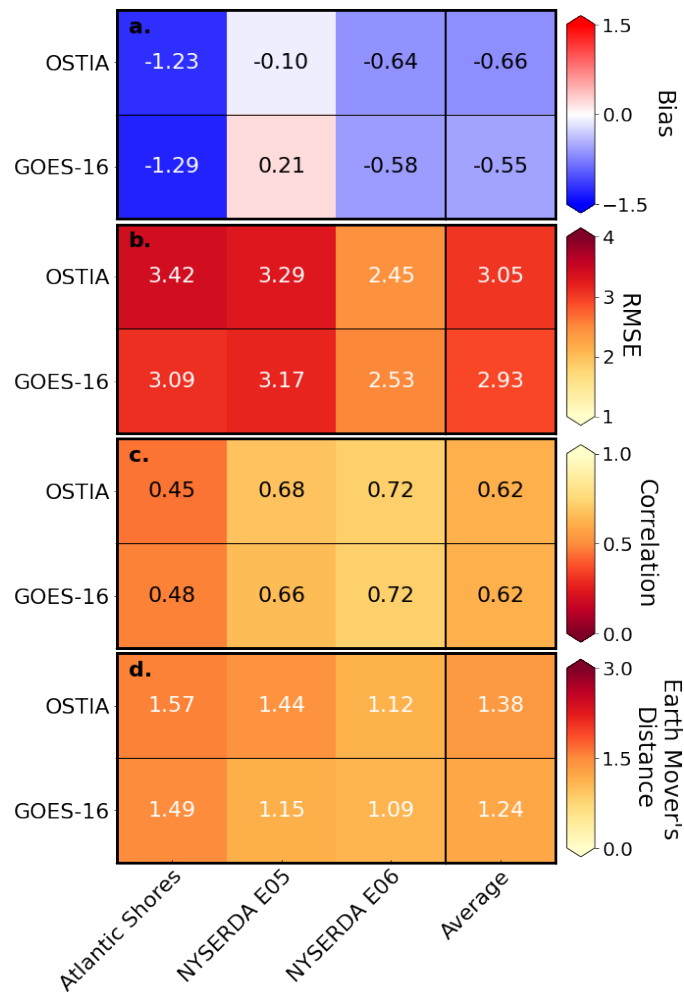

Figure 10. Differences (GOES-16 - OSTIA) in average ~~100-m~~ 140-m wind ~~speeds~~ speed (a.) and ~~SSTs~~ average SST (b.) ~~over~~ during the ~~entire domain for the June 21-22~~ July 10-11 event.

3.3.3 ~~July 10-11, 2020 Event~~

~~We have next flagged an event that occurred between 07-10-2020 06:25:00 and 07-11-2020 09:00:00.~~ Near-complete wind data sets exist at all three lidars for the duration of this event. Validation metrics show a correlation of 0.45 or greater between
310 model output and observations, for both simulations, at 140 m at each lidar site. The models perform the best, across all four metrics, at NYSERDA E06. Average r^2 between the two products is comparable, at 0.62. Bias strength peaks at -1.29 m s^{-1} at Atlantic Shores in the GOES-16 simulation. The bias is strongly negative at Atlantic Shores, but is positive at NYSERDA E05, the northernmost lidar, in the GOES-16 simulation. On average, bias in the OSTIA simulation is 0.11 m s^{-1} stronger (more

Figure 11. Wind speed validation metrics at an 140-m hub height at the Atlantic Shores, NYSERDA E05, and NYSERDA E06 lidars during the July 10-11 event: bias, RMSE, correlation, and EMD.

The GOES-16 and OSTIA simulations output statistically significantly different ~~100-m~~ 140-m wind speeds during this period. Both generally ~~tracked~~ track the major observed wind speed patterns at each observation site, although the timing and magnitude of some longer-scale changes ~~were~~ is missed by the models. For example, although the models captured the wind speed increase beginning at 07-10-2020 09:00:00 at Atlantic Shores, both erroneously ~~forecasted~~ forecast a subsequent

320 down-~~and then up-ramp~~, up-, and then down-ramp event (Fig. 12(a.~~))~~). At NYSERDA E05, the initial spike in observed speeds occurred several hours earlier than the models forecasted (Fig. 12(c))~~))~~. And, at NYSERDA ~~E05, the timing of a wind speed increase lagged observations by~~ E06, the models miss the timing and magnitude of a dip in wind speeds that was observed

between 07-10-2020 15:00:00 and 07-10-2020 20:00:00; they forecast a much larger drop in speed, almost 10 m s^{-1} compared with the original 6 ~~hours, and a down-ramp event was missed altogether (Fig. 12(b)). These faults~~ m s^{-1}, beginning just as the observed dip recovers. These faults all contribute to the relatively lower correlation of model output from both simulations with observations for this event.

~~Validation metrics for this event show a correlation of 0.5 or greater between model output and observations, for both simulations, at 100 m at each lidar site. Average r^2 is slightly stronger with GOES-16, which exceeded that of OSTIA by 0.02. OSTIA only outperforms GOES-16 in terms of bias, although both present values very close to 0 m s^{-1}. Hub-height (100 m) validation metrics for each simulation at each lidar are shown in Fig. 11.~~

~~Wind speed validation metrics at an 100-m hub height at the Atlantic Shores, NYSERDA E05, and NYSERDA E06 lidars during the July 10-11 event: bias (a), RMSE (b), correlation (c), and EMD (d).~~

3.3.3 July 18, 2020 Event

~~Synoptically, Tropical Storm Fay was observed to be moving in a southerly direction through the region~~during this time. The ~~high wind speeds , which peaked at 07-10-2020 18~~The last flagged event occurred between 07-18-2020 02:~~00~~05:~~00, may be attributed to the storm. Average wind directions also reflect the storm path (Fig. 10). Differences in average wind speed~~00 and 07-18-2020 14:10:00. During this time, a high pressure system was moving westerly into the region, causing a fall in wind speeds throughout the morning. This front spanned eastward into the Atlantic and aligned perpendicular to the New Jersey coast. Hub-height wind speed differences between the two simulations ~~peak at only 1.3~~are varied throughout the domain, with a maximum difference in the magnitudes of average event wind speeds of 1.7 m s^{-1} ~~which is around 1~~, and a mean difference of 0.29 m s^{-1} ~~less than the maximum difference during the June event~~$^{-1}$. (Fig. 13). Larger differences, in particular, are found near the coast, and positive/negative differences vary spatially. ~~However, as with the June event, the average differences in average wind speeds across the domain are nominally zero.~~

3.4 ~~July 18, 2020 Event~~

~~The last flagged event occurred between 07-18-2020 02:05:00 and 07-18-2020 14:10:00.~~ Sufficient data for this event are present at Atlantic Shores and NYSERDA E06, but not at NYSERDA E05. ~~Over this time period, wind speeds drop at both locations from over 14~~Validation analysis for the event at the two lidars indicates that overall the products perform comparably, although OSTIA shows an almost 0.5 m s^{-1} ~~to 3 m s^{-1} or lower. As with the previous two cases, the models capture the general wind speed trend of slowing throughout the duration of the event, although they both present errors in the magnitude of wind speeds and the timing of wind profile changes. Both simulations exhibit greater spread in wind speeds at each lidar than observations, and both have slower average velocities~~$^{-1}$ stronger bias at NYSERDA E06. At lower heights (Fig. ~~15).~~

~~Hub-height wind speed and wind direction at the Atlantic Shores lidar (a, c) and the NYSERDA E06 lidar (b, d) during the July 18 event.~~

~~Validation analysis for this event generally indicates that~~A3), there there is clear evidence of GOES-16 ~~is the stronger product during this time, as it outperforms OSTIA in every metric at hub height except for correlation~~outperforming OSTIA

22
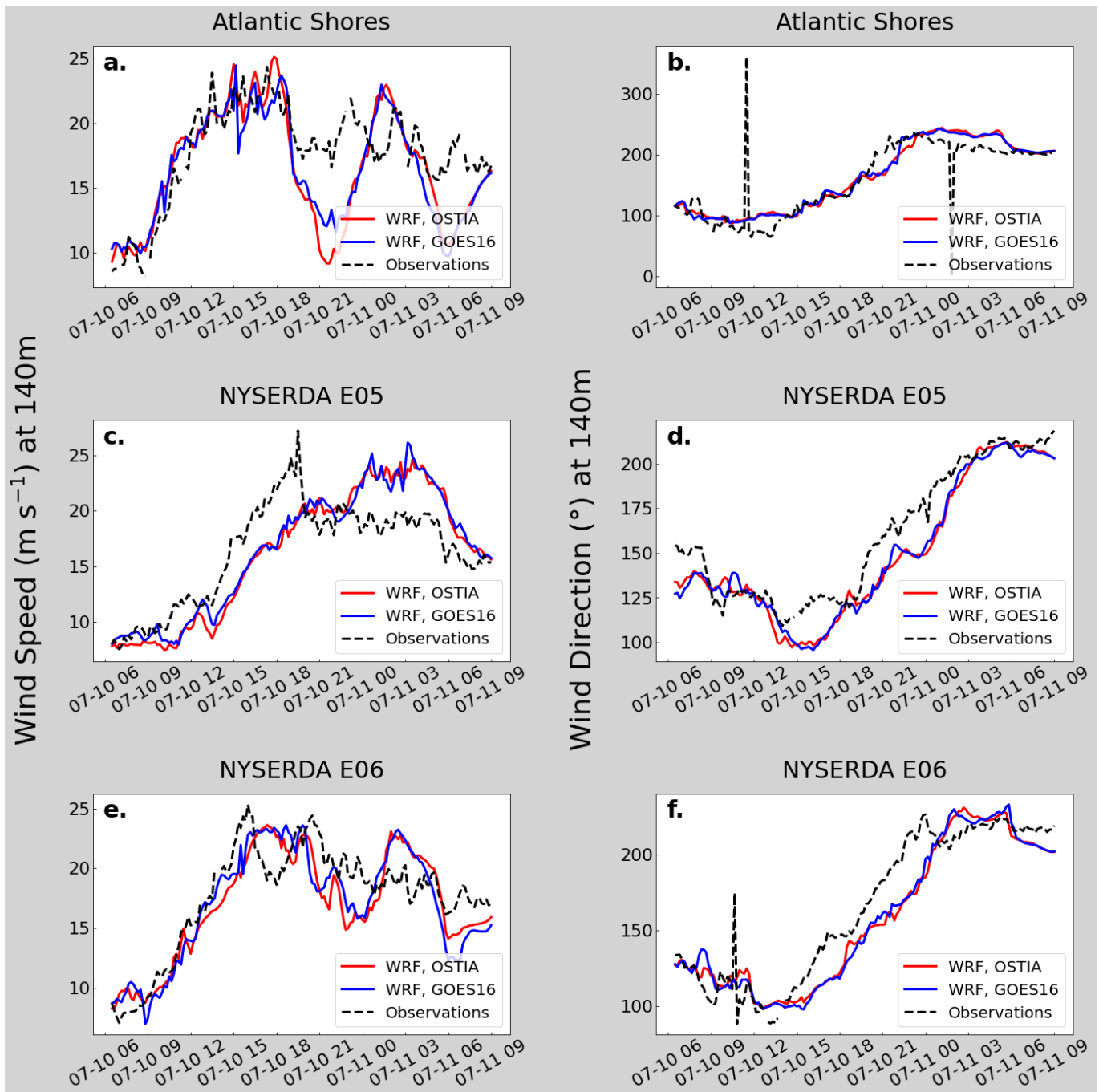
Figure 12. Hub-height (140-m) wind speed and wind direction at the Atlantic Shores lidar (a, b) the NYSERDA E05 lidar (c, d), and the NYSERDA E06 lidar (e, f) during the July 10-11 event.
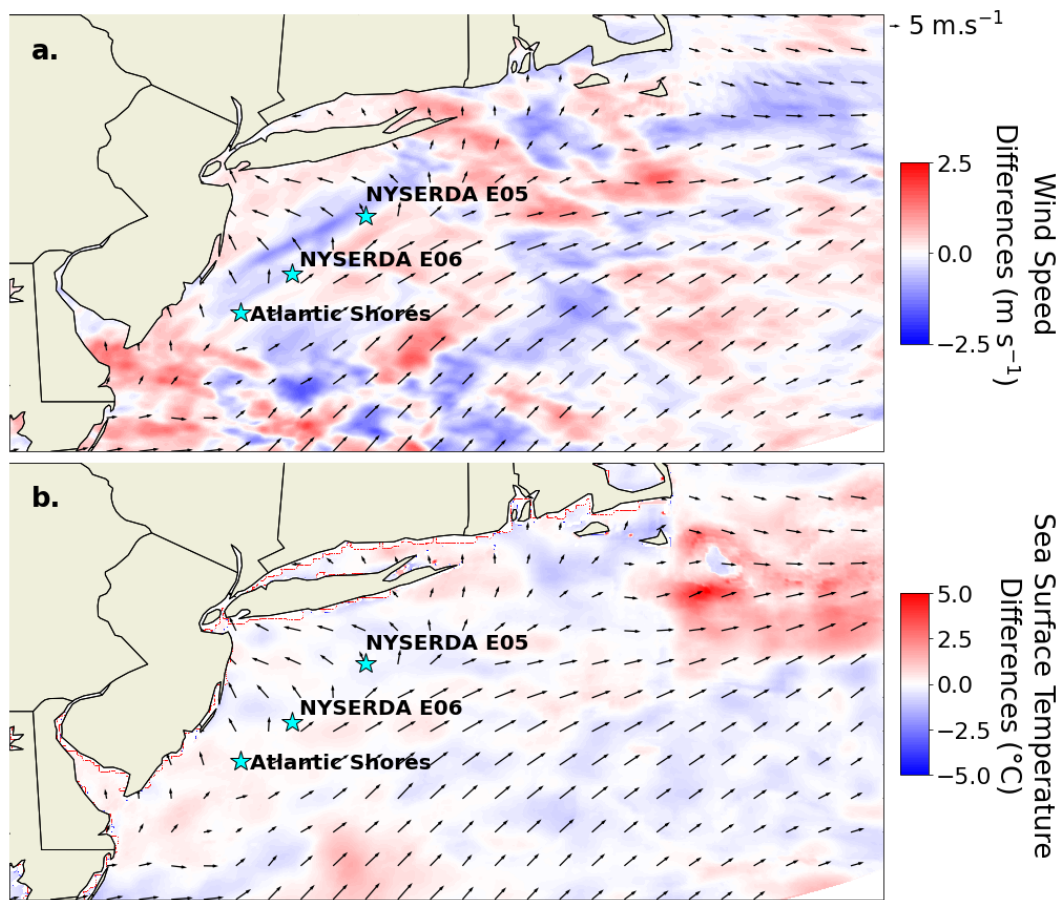
Figure 13. Differences (GOES-16 - OSTIA) in average ~~100-m~~ 140-m wind ~~speed~~ speeds (a.) and average ~~SST~~ SSTs (b.) ~~and~~ during the July ~~10-11~~ 18 event.

at Atlantic Shores. Farther north, at NYSERDA E06~~. This superior performance holds for heights up to 100 m, at which point both of the products deliver similar results (bias, Atlantic Shores), or OSTIA begins to outperform GOES-16 (RMSE and EMD,~~, the relative performance of each simulation is variable. Validation metrics at Atlantic Shores and NYSERDA E06 ~~).~~ are shown in Fig. 14. Of note is that PDFs of 140-m wind speeds during this time (not shown) indicate bimodality in the

360   observed wind speed distribution. Although OSTIA ~~somewhat~~ captures this feature, GOES-16 misses it almost entirely—at both lidars. ~~Validation metrics at Atlantic Shores and NYSERDA E06 are shown in Fig. 14.~~

    ~~Synoptically, this period was characterized by a cold front and rain having just moved northwesterly through the area, with a stationary front forming by the~~ Over this event's ~~end. This front spanned eastward into the Atlantic and aligned perpendicular to the New Jersey coast. Hub-height wind speed differences are varied throughout the entire domain, with a maximum difference~~

365   ~~in average wind speeds of 1.7~~ time period, 140-m wind speeds drop at both locations from over 14 m s^{-1} ~~. Larger wind speed differences, in particular, reside near the coast, and positive/negative differences vary spatially. Again, the average difference~~
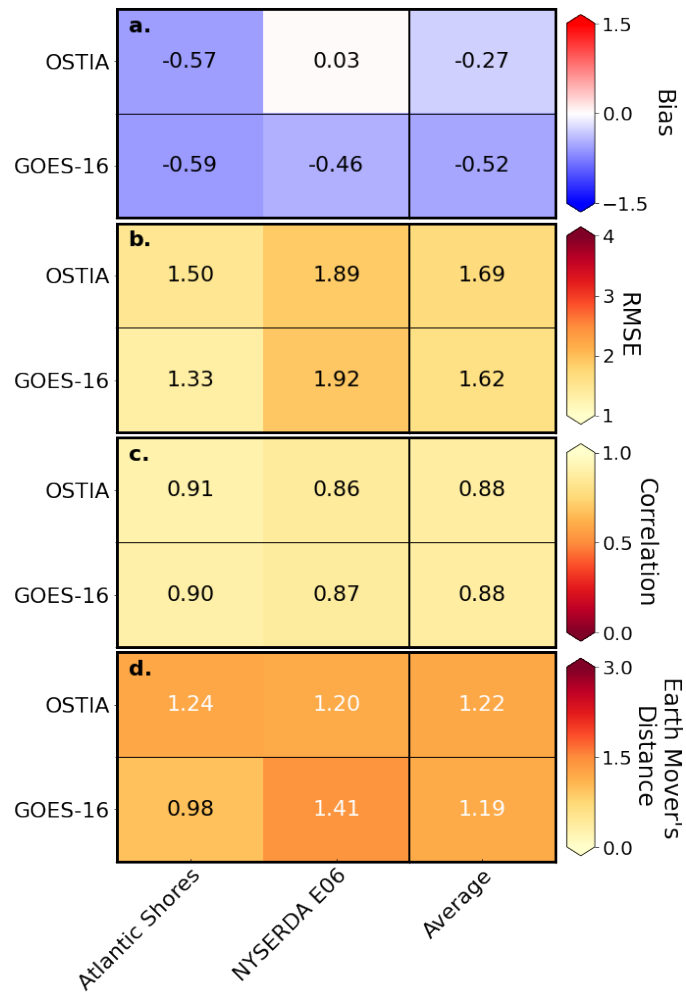
Figure 14. Hub-height (140-m) wind speed validation metrics at the Atlantic Shores and NYSERDA E06 lidars during the July 18 event: bias(a), RMSE(b), correlation(c), and EMD(d).

~~in average event wind speeds throughout the domain is nominally zero.~~ to less than 3 m s^{-1}, before beginning to increase again. As with the previous two cases, the models capture the general wind speed trend of slowing throughout the duration of the event, although they both present errors in the magnitude of wind speeds and the timing of wind profile changes. In this case, at NYSERDA E06 they predict a shift in wind direction before one was observed (there is insufficient wind direction data at NYSERDA E05 to make a determination about if this holds true at both lidars). Both simulations exhibit greater spread in wind speeds at each lidar than observations, and both have slower average velocities (Fig. ~~13~~15).
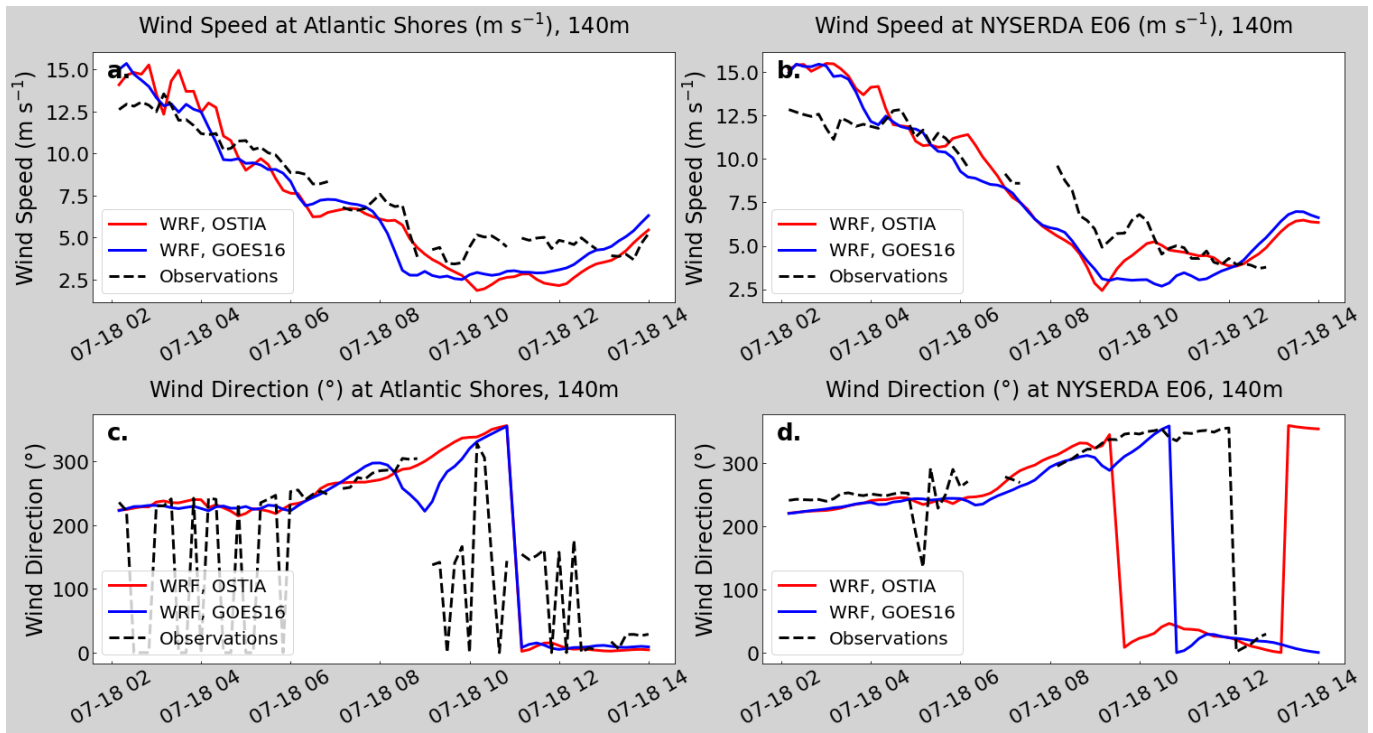
Figure 15. ~~Differences~~ Hub-height (~~GOES-16 - OSTIA~~140-m) ~~in average 100-m~~ wind ~~speeds~~ speed and wind direction at the Atlantic Shores lidar (a, c) and ~~average SSTs~~ the NYSERDA E06 lidar (b, d) ~~and~~ during the July 18 event.

4 Discussion

In Section 3, we saw varied performance between the two SST products, depending on the variable in consideration (SST
375   or winds) and the time scale (monthly or event-scale). SST in OSTIA validated better than GOES-16 across the buoy array.
However, hub-height wind speeds at each lidar location, for each month, point to a similar performance by both data sets. More
nuanced events that occurred over hourly-to-daily time scales, which were selected based on differences between simulation
output and overall WRF performance, also indicate ~~a~~ comparable performance between the two simulations, with, in general, a
slightly stronger performance using GOES-16 SSTs. Events such as these are of importance to wind energy forecasts because
380   we found that they often correlate with wind ramps, during which times wind speeds fall ~~within the nonrated~~ below the rated
power output section of turbine power curves. In this region, power output is very sensitive to fluctuations in wind speed. Therefore,
obtaining the most accurate wind forecast within this regime is ~~very~~ important.

~~Planar maps of SST and~~ Although OSTIA SSTs at the lidars validate better than GOES-16, modeled hub-height ~~winds
indicate a localized relationship between wind speed, wind direction, and SSTs (Fig. 16). As mentioned in Section 3.3, despite~~
385   ~~the fact that wind speed differences in the domain tend to average out, there can still be significant deviations between the
two in different regions of the domain, and this must be considered.~~ 140-m) winds do not demonstrate the same trend.

26

Compared with the 140-m winds heat maps for the events (Fig. 9, 11, 14), good wind characterization performance does not necessarily correspond to good SST validation (Appendix B), which implies that SSTs across the full study area, not just localized temperatures, may have an impact on winds in the leasing area.

390 ~~Differences (GOES-16 - OSTIA) in average SST for June (a) and July (c), and in modeled 100-m wind speeds for June (b) and July (d).~~

~~Although OSTIA SSTs at the lidars~~ To further examine the disjointedness between SST and wind speed point validations, we consider model performance during each event in conjunction with how well their corresponding SSTs validate with observations across the buoy array (where data are available). During the June 21-22 event, 140-m wind speed bias and EMD

395 in both simulations are the lowest of all the cases considered (Fig. 9), but there is an overall cold SST bias (-0.31°C for OSTIA, -0.84° for GOES-16) across the buoy arrays, and SST correlation is very poor (Fig. B1). During the July 10-11 event, OSTIA SST's validate better than GOES-16 ~~,~~ (Fig. B2), but the two simulation's hub-height ~~winds do not. This implies a more widespread impact on winds in the leasing area by SSTs across the full study area. More~~ wind speeds validate similarly, with relatively poor performance across all validation metrics compared with the other flagged events. And, during the July 18 event,

400 while SSTs from both products validated the best out of the three events (Fig. B3) and the models captured the overall trend in wind speeds with correlations of 0.88 for both simulations, EMD and the magnitude of the bias were both greater than the June 21-22 event.

Across all three events there is no clear relationship between wind forecasting ability and SST correctness. However, as with the montly analysis, the relatively larger differences between SST and wind speed validations in the GOES-16 event simulations

405 compared with those of OSTIA suggests that GOES-16's finer temporal and spatial resolutions may more accurately influence larger-scale dynamics throughout the region, leading to greater forecasting skill. Though beyond the scope of this study, more widespread in situ observations and validation of the two data sets against observations would help address how large of an effect this is. ~~This~~ Such a study would additionally allow for a better understanding of how much ~~the capture~~ a nuanced representation of diurnal SST cycling (as in the GOES-16 product, Fig. 2) influences the accuracy of wind forecasting in the

410 region.

Rapid fluctuations in wind speed, on the order of minutes to hours, are overall not well captured by either simulation. This is most evident during the July 10-11 event, where the magnitude and timing of wind speed changes are most misaligned. Correlation in particular is low (0.62), and RMSE is high (3 m s^{-1}). In comparison, during the July 18 event, when wind speeds slow without any ramping, both simulations perform better; in particular, correlation is relatively strong for both (0.88).

415 Similarly, during the June 21-22 event, which lacks any sharp ramping, correlation is also above 0.75 for both GOES-16 and OSTIA. Shorter, weaker fluctuations in wind speed during this event are not well captured by either simulation (Fig. 8), although GOES-16 validates with nominally less error (Fig. 9). It is possible that GOES-16's finer resolution may contribute to its slightly stronger performance in this case, as the product does capture hourly fluctuations in SSTs. In contrast, during the July 10-11 storm, winds are so strong and variable that SST resolution does not appear to make a significant difference, and

420 the two simulations both show significant error.

Also of note are the differences in weather between June and July in the Mid-Atlantic region. June average wind speeds are faster than those in July for both simulations. ~~June also had overall~~ Overall cooler SSTs throughout the region are also present in June, with a weaker temperature gradient in the cold bubble offshore of Cape Cod than in July. The difference in temperature in this region between GOES-16 and OSTIA is more pronounced in July, with GOES-16 ~~presenting~~ showing warmer monthly

425 average SSTs ~~.~~ (Fig. 16). The corresponding average wind speed and surface pressure differences between the products are also more distinct in July. While the two simulations validated similarly for each month, in the more turbulent July environment, OSTIA ~~performs better~~ outperforms GOES-16. However, due to the greater number of missing pixels that required gap-filling in July as compared with June (36.86% more over the course of each month, resulting from the increased cloud cover—and the possible stronger coastal cold pool), this indication of product superiority may be questionable.

430 ~~GOES-16 SSTs without bias correction indicate a very clear warm bias throughout the entire domain, as compared with OSTIA (Fig. 16a,c). Without the added SSES field, the product clearly underperforms compared with OSTIA, so it should be included whenever using GOES-16 as a WRF input.~~ The addition of more refined post-processing to account for incorrectly cloud-filtered pixels, as shown in Murphy et al. (2021), has been demonstrated to improve the accuracy GOES-16 SSTs~~even further~~. A comparison of simulations using OSTIA against those using the sophisticated GOES-16 product should be conducted

435 in the future to evaluate the level of improvement delivered using the latter's more computationally expensive data set.
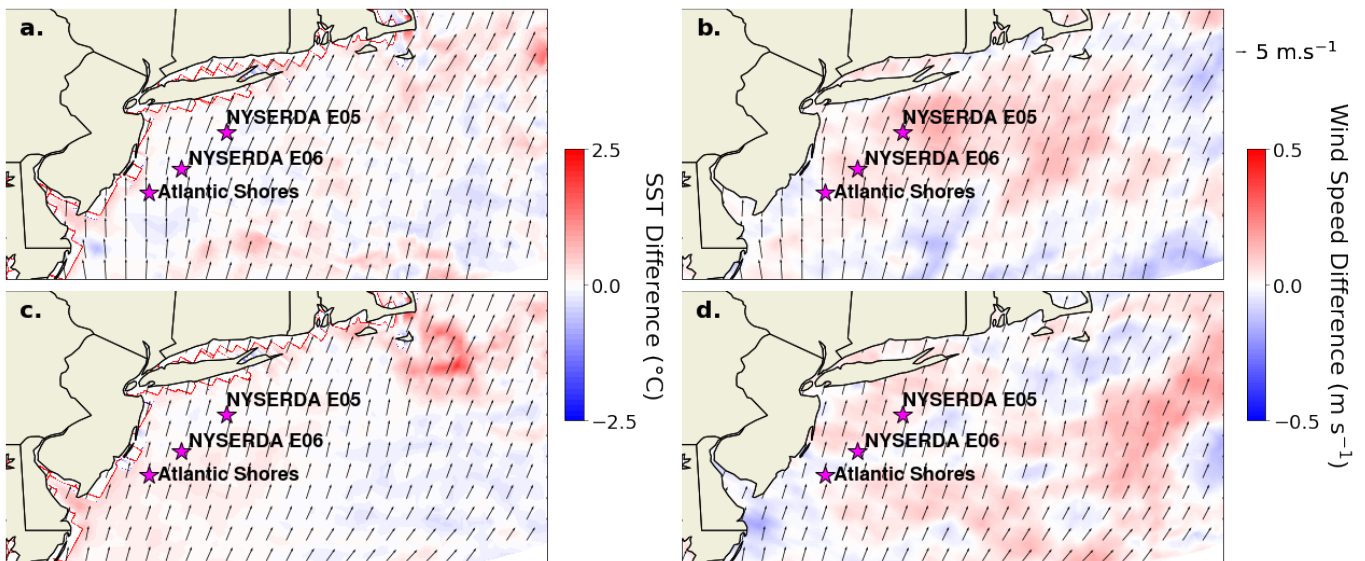
Figure 16. Differences (GOES-16 - OSTIA) in average SST for June (a.) and July (c.), and in modeled 140-m wind speeds for June (b.) and July (d.).

5 Conclusions

In this study, we evaluated ~~the performance of WRF to forecast~~ how SST inputs affect how well WRF forecasts winds off the Mid-Atlantic coast ~~using two different input~~. We initially considered three remotely-sensed SST data sets and validated them against in-situ observations taken at various buoys in the region. The first data set, OSTIA, has a 0.05° resolution, is output at a daily interval ~~. This product has been post-processed to fill any gaps in the data and assimilate it~~, and is assimilated with in situ observations. The second product, GOES-16, has a finer resolution of 0.02° and is generated on an hourly interval. ~~GOES-16 is gridded data.~~, but does not have the same level of post-processing as OSTIA. The data set contains gaps ~~where abnormally cold pixels have been flagged as clouds and removed, and there is no~~ and lacks assimilation with in situ observations. To account for ~~the missing pixels~~ missing data, we ran an EOF process to statistically analyze trends in each monthly data set and fill its gaps. ~~Our findings in comparing model performance using these two data sets indicate that, while OSTIA SSTs validate better against buoy measurements, GOES-16 tends to forecast hub-height (100 m) winds with greater accuracy.~~ The third SST data set, MUR25 has a 0.2° spatial resolution and a daily temporal resolution, and similar to OSTIA it is assimilated with in situ observations. MUR25 was removed from further analysis after SST validation, as its performance compared with the other two data sets was notably poor.

~~To compare the performance of the two products~~ Following SST validation, WRF was run for June and July of 2020 in the Mid-Atlantic domain depicted in Fig. 1. The simulation setups ~~are~~ were identical with the exception of the input SST field ~~. We consider how SST values at specified buoys validate against observational data taken at these locations for each month. Specifically, we evaluate~~ (key model parameters presented in Table 2). We compared the characterized 140-m (hub-height) winds against lidar observations by calculating correlation, bias, RMSE, and EMD ~~. Following the SST analysis, we examine hub-height wind speeds and validate model output against observational data taken at three separate lidars in the domain. We run the validation~~ on both monthly and ~~hourly-to-daily time scales. The shorter periods were identified using a flagging algorthim aimed to select periods~~ event-length (hourly) timescales. Times during which there ~~is~~ was significant difference in wind speed output from the two simulations and the general WRF performance ~~is satisfactory~~.

~~The results show a better agreement between OSTIA and in situ SSTs than between GOES-16 and in situ SSTs. Despite this, modeled winds from the two simulations validate similarly on a monthly time scale~~ was satisfactory were flagged as events. Our findings indicated that while OSTIA SSTs validated better against buoy measurements, the model validated comparably when forced with each SST data set, with GOES-16 ~~performing slightly better in June and OSTIA performing slightly better in July. Three events were flagged using the aforementioned algorithm, and on those time scales, GOES-16 outperforms OSTIA in forecasting hub-height wind speeds~~ having output nominally more accurate wind speeds. These findings suggest that the differences in temporal and spatial resolution between the two SST products influence wind speed characterization, with finer resolutions leading to better forecasting skill.

~~This~~ Overall, this study shows that SST inputs to WRF do affect forecasted winds and how well they validate against observations in future leasing areas in the Mid-Atlantic. ~~There is indication that~~ Although GOES-16 ~~, being of finer spatial and temporal resolution, is the better input product for forecasting offshore winds on the Mid-Atlantic coast. However,~~

470 ~~this~~ SSTs validate worse against buoy observations than OSTIA SSTs, on both monthly and event timescales, the wind characterization produced using GOES-16 inputs validates as well as, or in some cases better than, the OSTIA characterization, indicating that finer temporal and spatial SST resolution may contribute to improved wind forecasts in this region. To further explore this, a higher-quality post-processed GOES-16 data set assimilated with in situ observations, such as that created by Murphy et al. (2021), could be evaluated in future studies; generating a comparable data set was beyond the scope of this study.

475

This step of the research is in its early stages and more cases need to be considered—on both monthly and event scales. Future work should focus on identifying and analyzing more promising events~~—before a more definitive evaluation can be made. In future work , the authors plan to continue identifying promising events~~in particular, wind ramps and low pressure systems, ~~specifically including wind ramps ,~~ as they have shown to produce larger differences between GOES-16 and OSTIA-forced

480 model output. ~~Additionally, an advanced post-processing mechanism, as in Murphy et al. (2021), may improve GOES-16 performance even further, and a study including this method should be conducted~~Wind direction validation, as well, should be considered in the future, as it was beyond the scope of this study.

Appendix A: Average validation metrics through 200 m for flagged cases

Contents of this Appendix include plots of horizontally averaged bias, r²correlation, RMSE, and EMD from sea level through
a vertical height of 200 m for each flagged case, at each relevant lidar. Fig. A1 shows metrics at the two NYSERDA lidars for
the June 21-22 event, Fig. A2 shows metrics at all three lidars for the July 10-11 event, and Fig. A3 shows metrics at Atlantic
Shores and NYSERDA E06 for the July 18 event. The GOES-16 simulations are indicated by the blue lines, while the OSTIA
simulations are shown in red. The labels AS, E05, and E06 refer to Atlantic Shores, NYSERDA E05, and NYSERDA E06,
respectively.
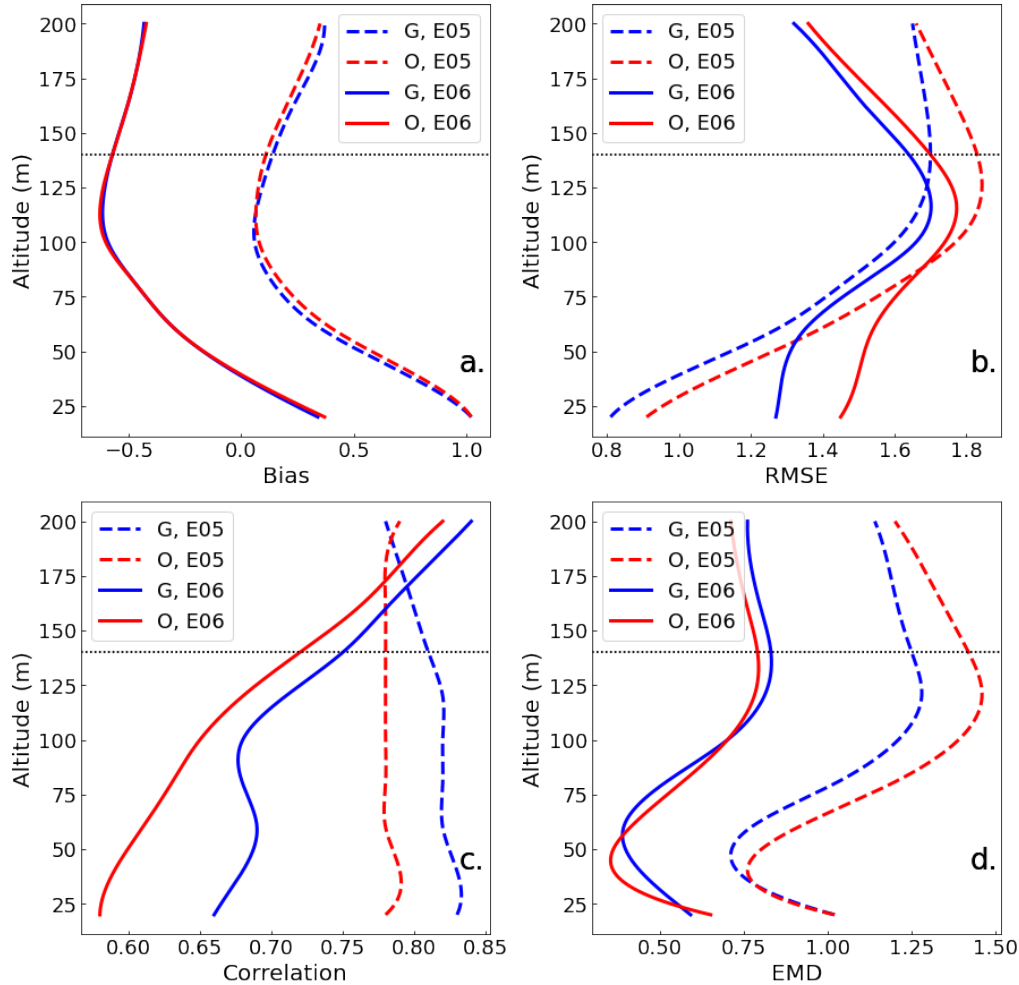
Figure A1. Validation metrics at NYSERDA E05 and NYSERDA E06 during the June 21-22 event: bias (a.), RMSE (b.), correlation (c.),
and EMD (d.). Hub height is marked by a horizontal line at ~~100~~ 140 m.
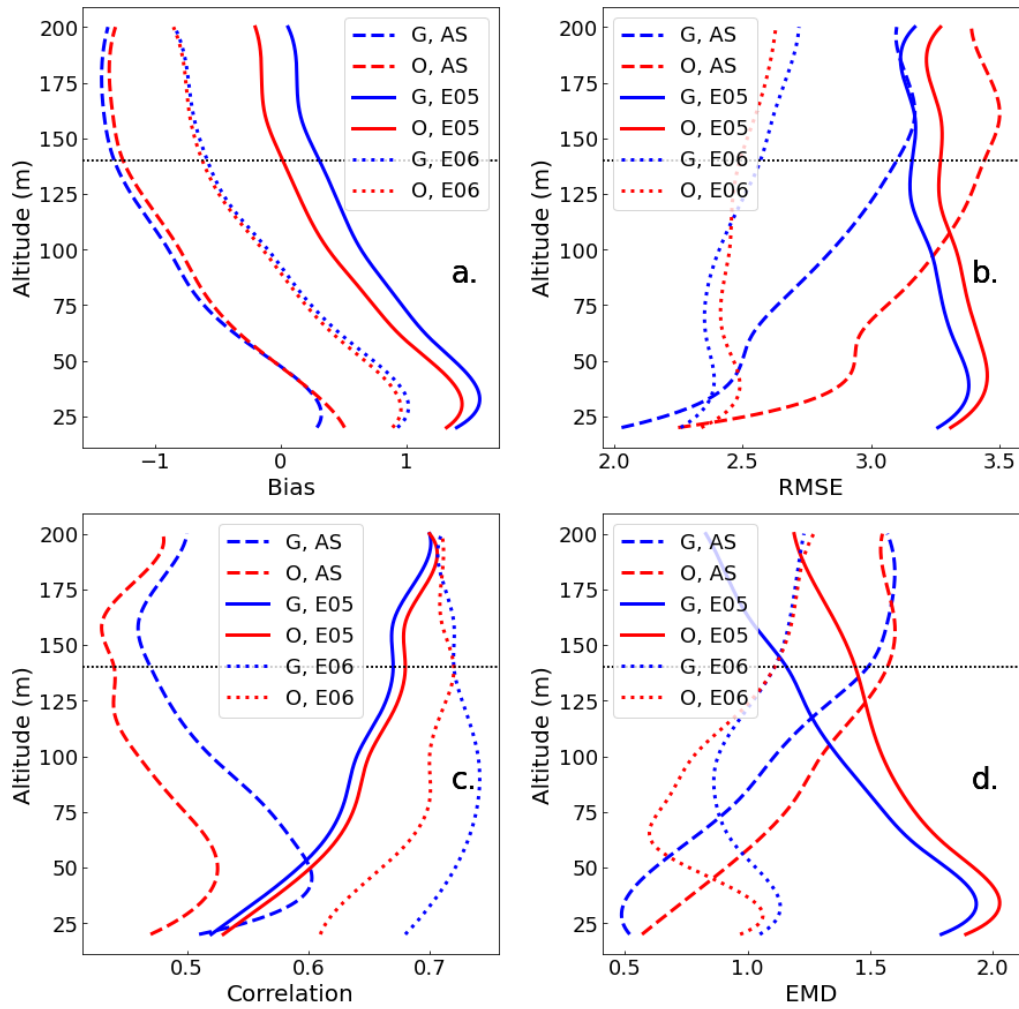
Figure A2. Validation metrics at Atlantic Shores, NYSERDA E05, and NYSERDA E06 during the July 10-11 event: bias (a.), RMSE (b.), correlation (c.), and EMD (d.). Hub height is marked by a horizontal line at ~~100~~ 140 m.
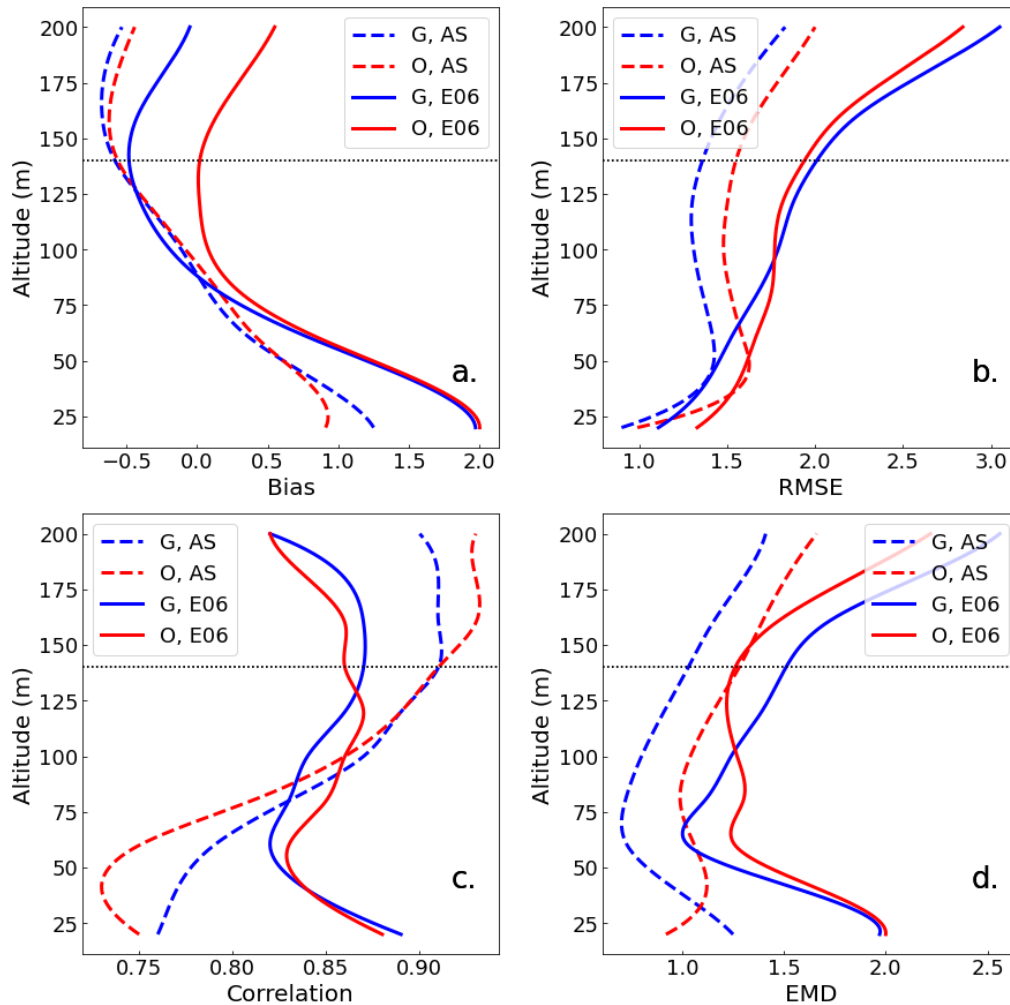
Figure A3. Validation metrics at Atlantic Shores and NYSERDA E06 during the July 18 event: bias (a.), RMSE (b.), correlation (c.), and EMD (d.). Hub height is marked by a horizontal line at ~~100~~ 140 m.

Appendix B: Sea Surface Temperature Validation Metrics for the Flagged Events

Contents of this Appendix include heat maps of SST validation metrics for each of the three events flagged in this study. Each plot includes a different subset of the buoy array; locations with missing data were removed from analysis on a case by case basis. The average value for each metric across all buoys considered is located on the far right side of each row. The validation metrics considered are bias, RMSE, correlation, and EMD. Fig. B1 shows SST validation metrics for the June 21-22 event, Fig. B2 shows SST validation metrics for the July 10-11 event, and Fig. B3 shows SST validation metrics for the July 18 event.
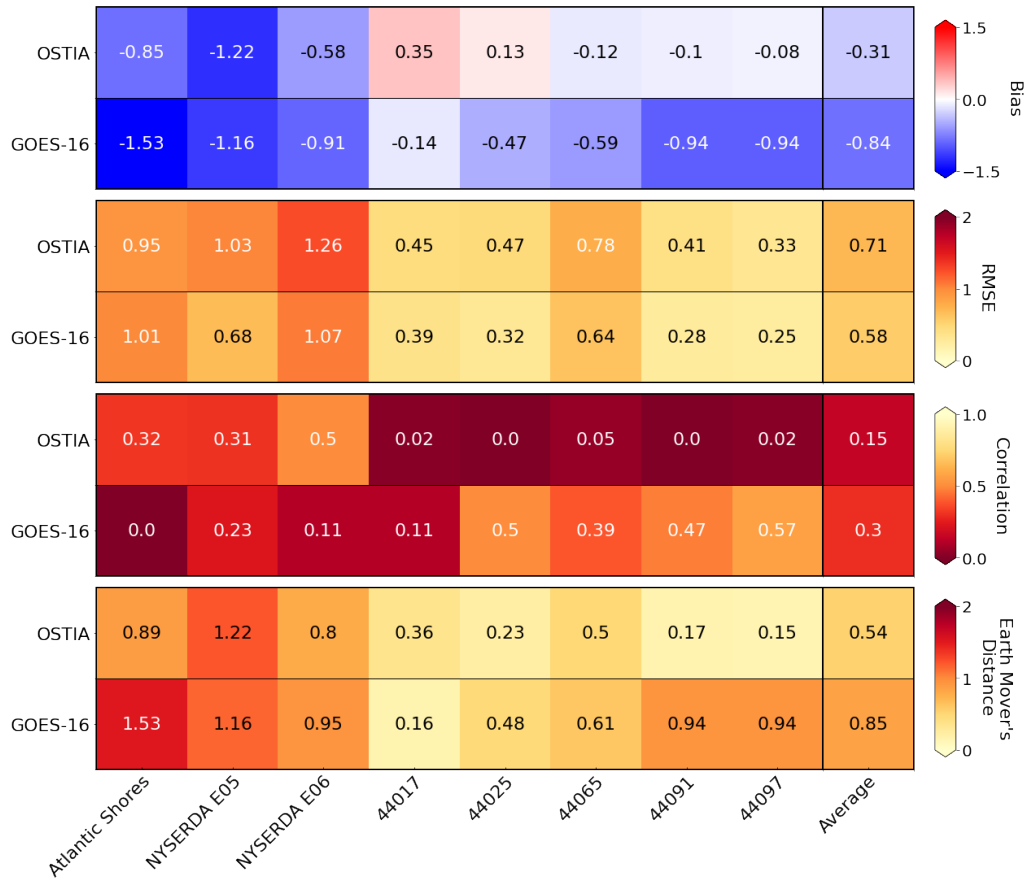
Figure B1. SST validation metrics (bias, RMSE, correlation, EMD) during the June 21-22 event, calculated across a subset of the buoy array. Omitted buoys are absent due to missing observational data at those locations.
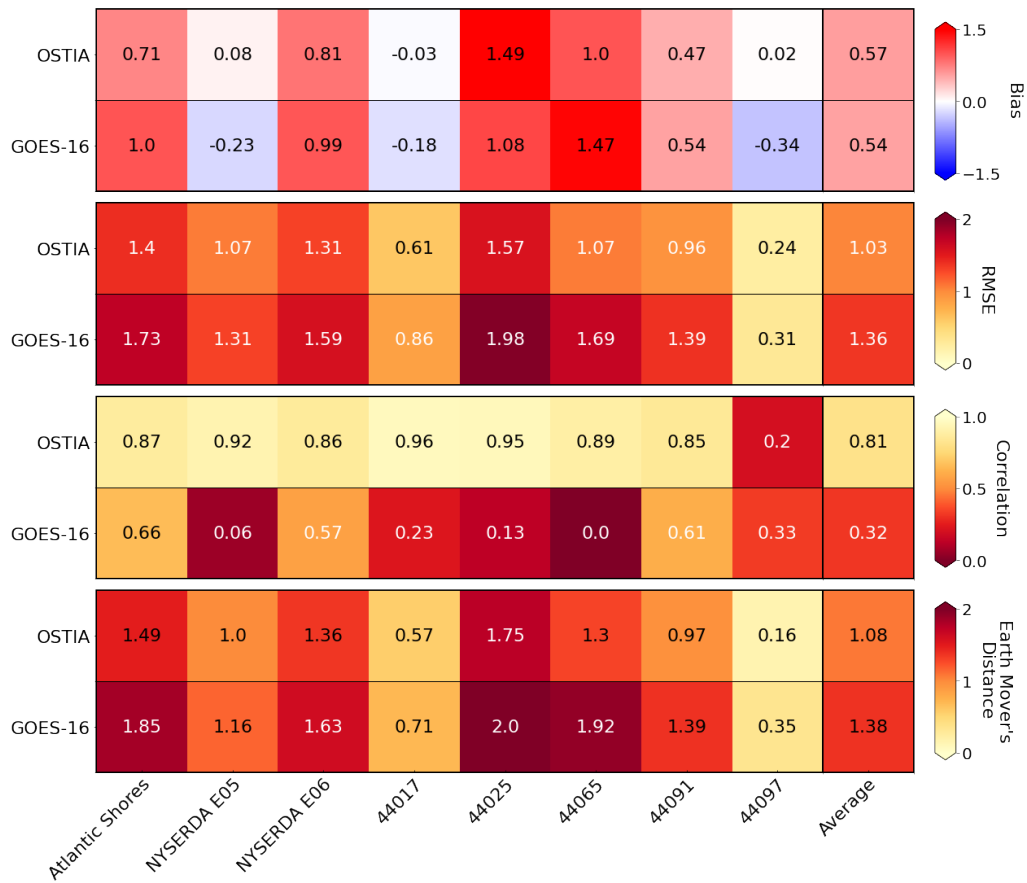
Figure B2. SST validation metrics (bias, RMSE, correlation, EMD) during the July 10-11 event, calculated across a subset of the buoy array. Omitted buoys are absent due to missing observational data at those locations.
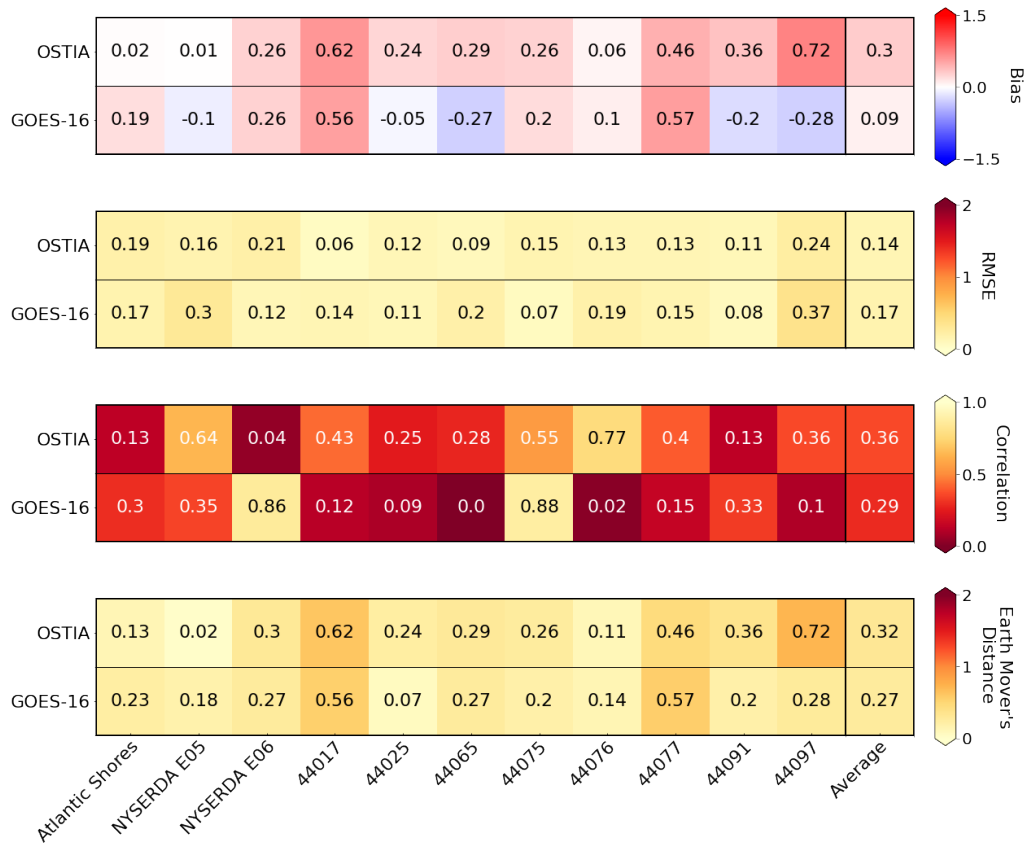
Figure B3. SST validation metrics (bias, RMSE, correlation, EMD) during the July 18 event, calculated across a subset of the buoy array. Omitted buoys are absent due to missing observational data at those locations.

Appendix C: Wind Speed Validation Heat Maps at 100 m and 120 m

Although this study considered a hub height of 140 m and therefore analyzed wind speeds primarily at that height, we include here validation metrics for each flagged event, at each lidar, for 100-m and 120-m hub heights. 140-m hub heights are expected to be close to the standard for offshore wind development (Lantz et al., 2019), but shorter turbines may also be constructed; additionally, having a better understanding of model performance across the rotor-swept area and not just at hub height can better inform wind power potential characterization.

Below, Fig. C1 presents 100-m model performance at the two NYSERDA lidars for the June 21-22 event, and Fig. C2 shows performance metrics for the same event and lidars at 120 m. At both heights, GOES-16 performs slightly better than OSTIA.
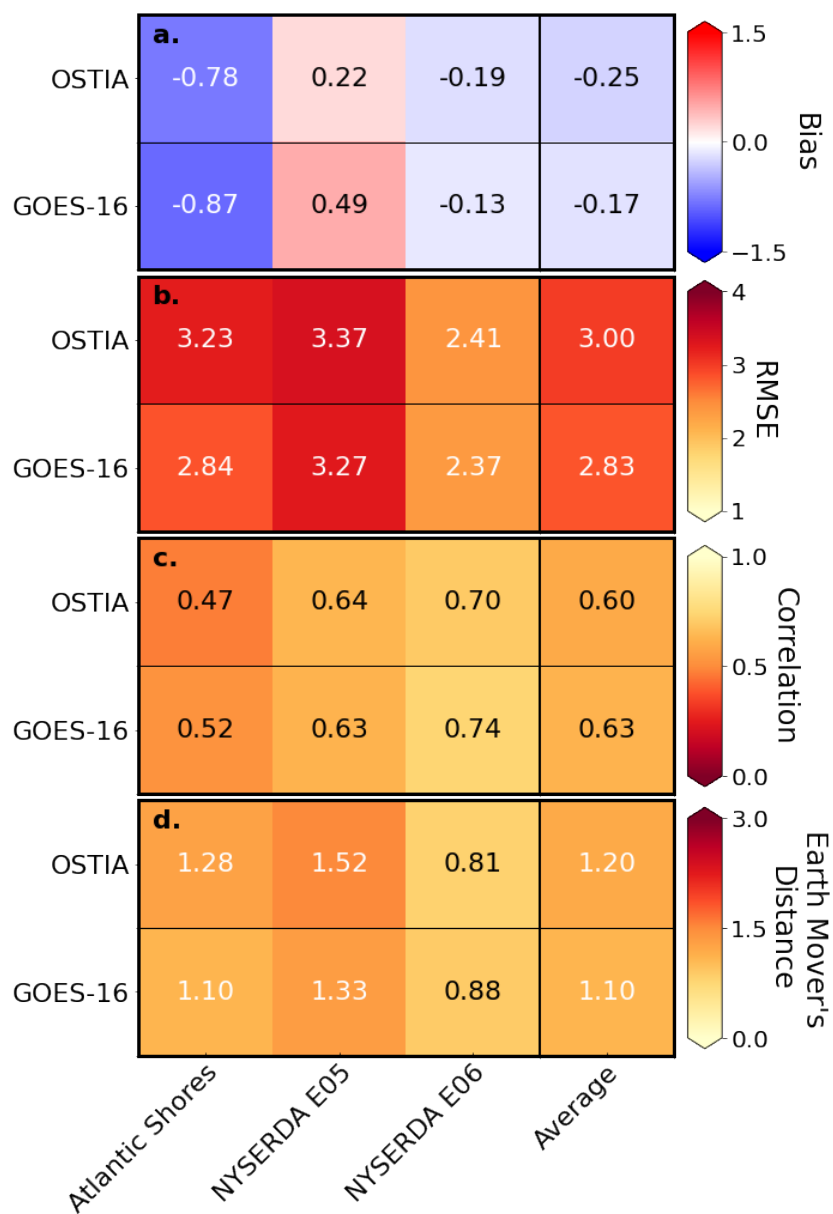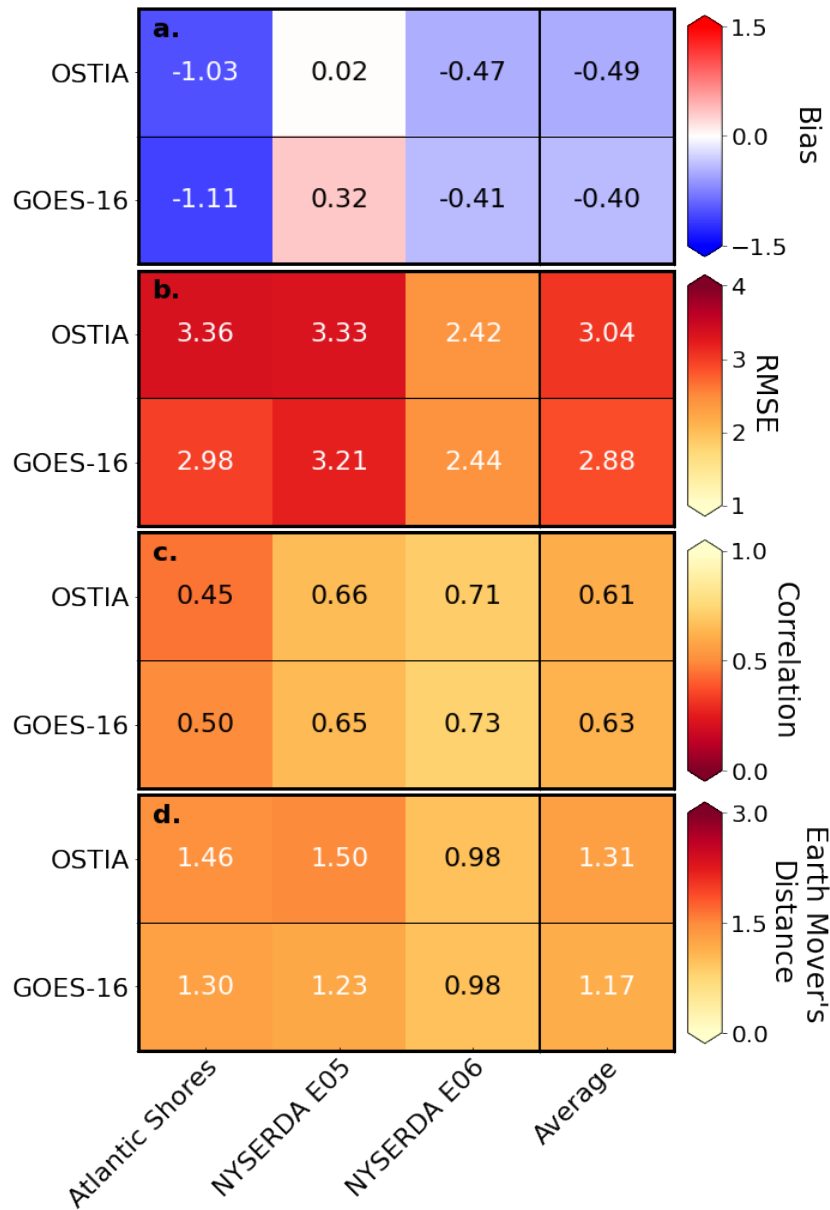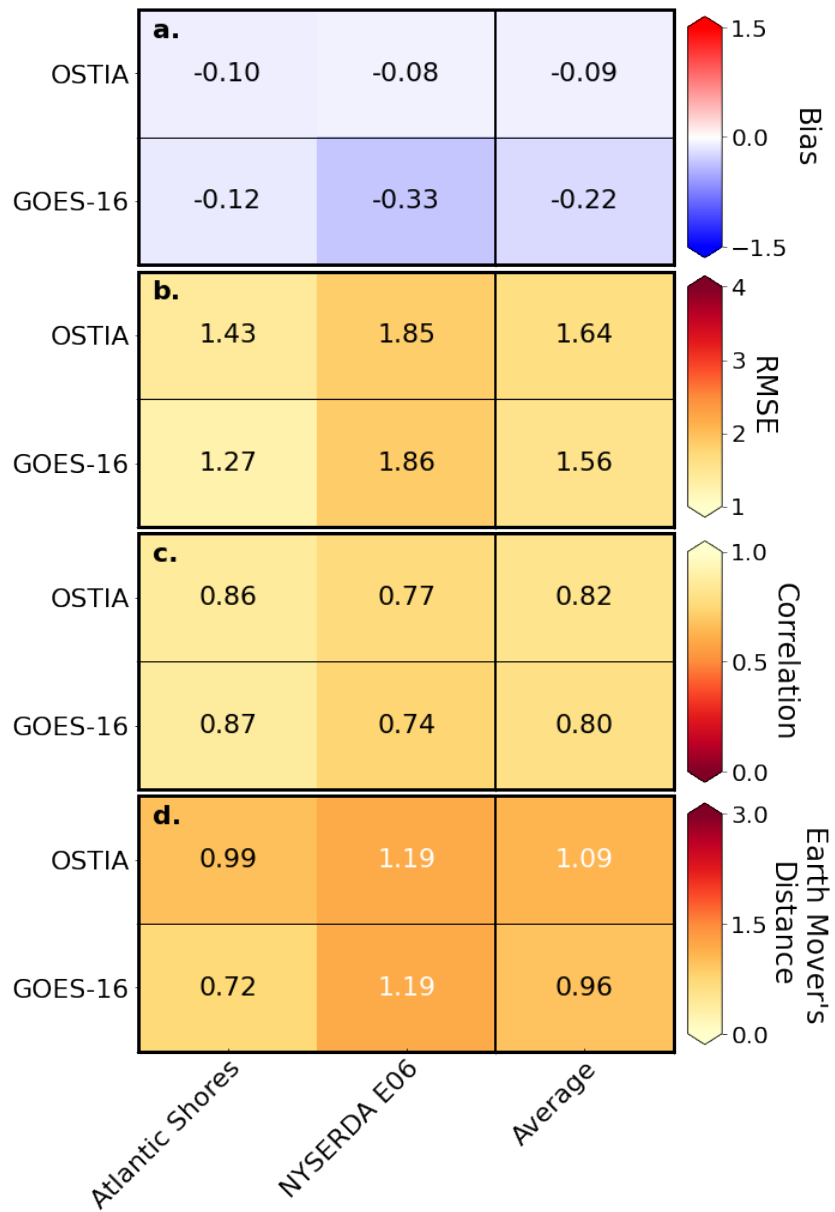
500
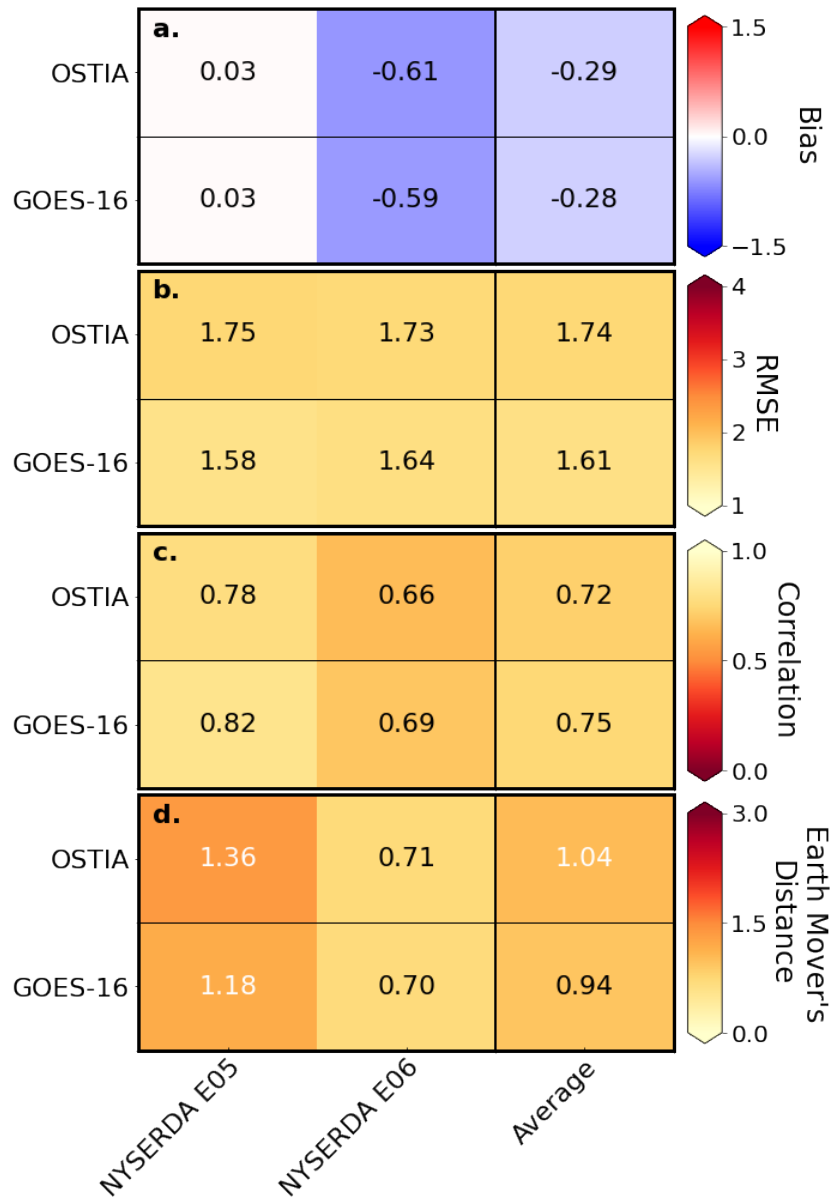
Figure C1. 100-m hub-height validation metrics at NYSERDA E05 and NYSERDA E06 during the June 21-22 event: bias, RMSE, correlation, and EMD.
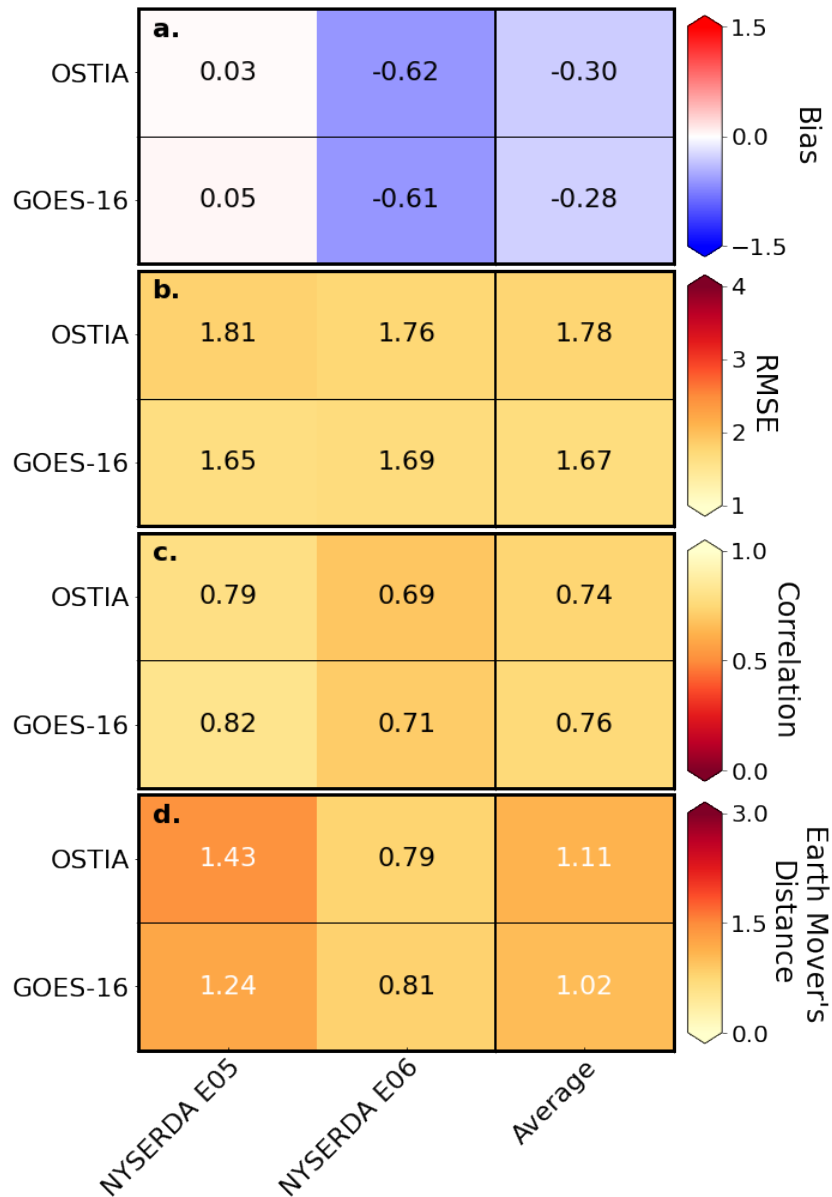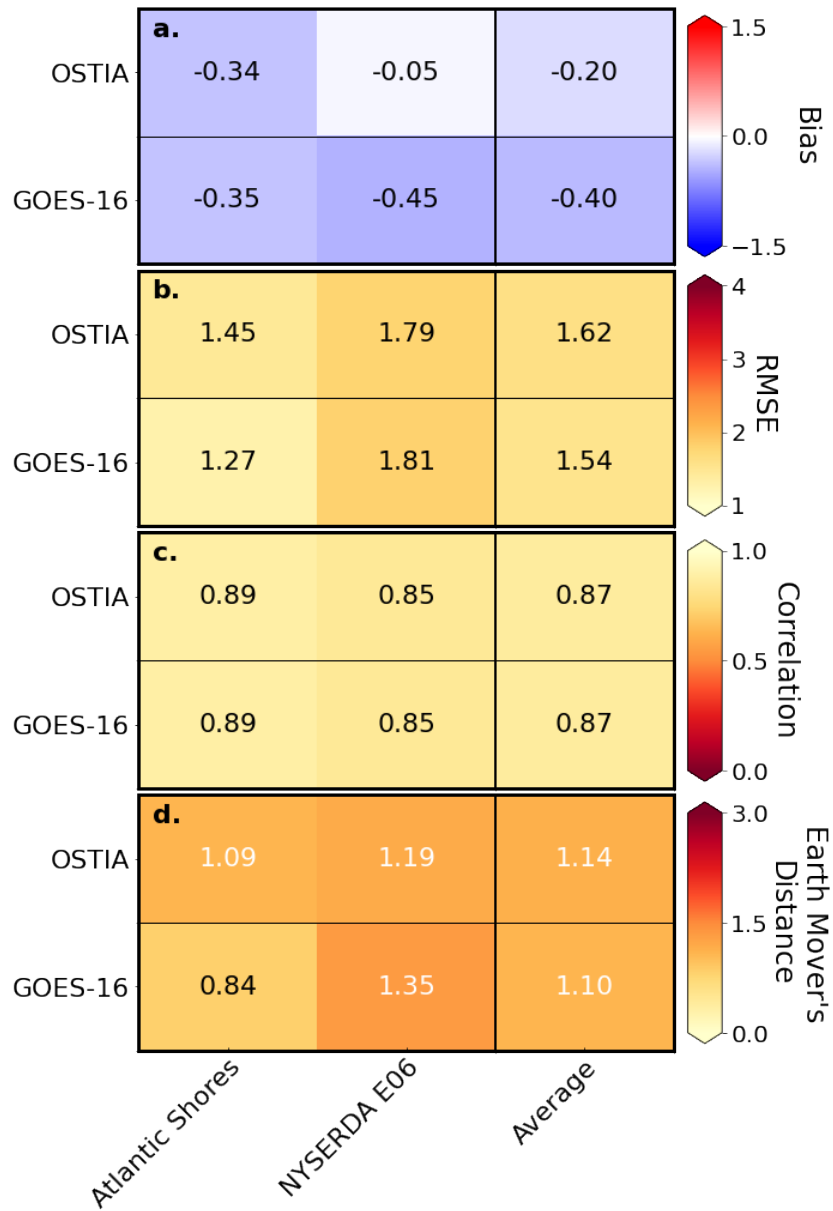
Figure C2. 120-m hub-height validation metrics at NYSERDA E05 and NYSERDA E06 during the June 21-22 event: bias, RMSE, correlation, and EMD.

Fig. C3 shows 100-m model performance at all three lidars for the July 10-11 event, and Fig. C4 presents performance
505 metrics at all three lidars for the July 10-11 event at 120 m. While GOES-16 shows a higher average wind speed bias than
OSTIA, the results from GOES-16 have a lower average RMSE and a higher average correlation than OSTIA. The wind speed

bias at all three lidars is notably strong, with a negative bias at Atlantic Shores (over 1 m s^{-1} slower than observed) and a positive bias at both NYSERDA lidars (well over 1 m s^{-1} at a 120-m hub height at NYSERDA E05).

Figure C3. 100-m hub-height validation metrics at NYSERDA E05 and NYSERDA E06 during the July 10-11 event: bias, RMSE, correlation, and EMD.

Figure C4. 120-m hub-height validation metrics at NYSERDA E05 and NYSERDA E06 during the July 10-11 event: bias, RMSE, correlation, and EMD.

Fig. C5 depicts 100-m model performance at Atlantic Shores and NYSERDA E06 for the July 18 event, and Fig. C6 shows
510 performance metrics at Atlantic Shores and NYSERDA E06 for the July 18 event at 120 m. Although both simulations'
hub-height wind speeds show comparable correlation with observations, OSTIA has an over 0.1 m s^{-1} stronger wind speed
bias than GOES-16, and GOES-16 validates with lower RMSE and EMD values.

Figure C5. 100-m hub-height validation metrics at Atlantic Shores and NYSERDA E06 during the July 18 event: bias, RMSE, correlation, and EMD.

Figure C6. 120-m hub-height validation metrics at Atlantic Shores and NYSERDA E06 during the July 18 event: bias, RMSE, correlation, and EMD.

Data availability.

Buoy data outside of the Atlantic Shores and NYSERDA floating lidars were collected via the National Oceanic and Atmo-
515   spheric Administration's NBDC ~~.~~ (https://www.ndbc.noaa.gov). Satellite data were collected via the Physical Oceanography
Distributed Active Archive Center. If interested in obtaining other information (e.g. WRF namelists), please contact the authors
directly using the listed correspondence email.

References

Aird, J. A., Barthelmie, R. J., Shepherd, T. J., and Pryor, S. C.: Occurrence of Low-Level Jets over the Eastern U.S. Coastal Zone at Heights Relevant to Wind Energy, Energies, 15, 445, https://doi.org/10.3390/en15020445, number: 2 Publisher: Multidisciplinary Digital Publishing Institute, 2022.

Alvera-Azcárate, A., Barth, A., Rixen, M., and Beckers, J. M.: Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature, Ocean Modelling, 9, 325–346, https://doi.org/10.1016/j.ocemod.2004.08.001, 2005.

Banta, R. M., Pichugina, Y. L., Brewer, W. A., James, E. P., Olson, J. B., Benjamin, S. G., Carley, J. R., Bianco, L., Djalalova, I. V., Wilczak, J. M., Hardesty, R. M., Cline, J., and Marquis, M. C.: Evaluating and Improving NWP Forecast Models for the Future: How the Needs of Offshore Wind Energy Can Point the Way, Bulletin of the American Meteorological Society, 99, 1155–1176, https://doi.org/10.1175/BAMS-D-16-0310.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2018.

Bureau of Ocean Energy Management: Outer Continental Shelf Renewable Energy Leases Map Book, 2018.

Byun, D., Kim, S., Cheng, F.-Y., Kim, H.-C., and Ngan, F.: Improved Modeling Inputs: Land Use and Sea-Surface Temperature, Final Report, Texas Commission on Environmental Quality, 2007.

Chen, F., Miao, S., Tewari, M., Bao, J.-W., and Kusaka, H.: A numerical study of interactions between surface forcing and sea breeze circulations and their effects on stagnation in the greater Houston area, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2010JD015533, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010JD015533, 2011.

Chen, Z., Curchitser, E., Chant, R., and Kang, D.: Seasonal Variability of the Cold Pool Over the Mid-Atlantic Bight Continental Shelf, Journal of Geophysical Research: Oceans, 123, 8203–8226, https://doi.org/https://doi.org/10.1029/2018JC014148, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018JC014148, 2018.

Chin, T. M., Vazquez-Cuervo, J., and Armstrong, E. M.: A multi-scale high-resolution analysis of global sea surface temperature, Remote Sensing of Environment, 200, 154–169, https://doi.org/10.1016/j.rse.2017.07.029, 2017.

Colle, B. A. and Novak, D. R.: The New York Bight Jet: Climatology and Dynamical Evolution, Monthly Weather Review, 138, 2385–2404, https://doi.org/10.1175/2009MWR3231.1, publisher: American Meteorological Society Section: Monthly Weather Review, 2010.

Debnath, M., Doubrawa, P., Optis, M., Hawbecker, P., and Bodini, N.: Extreme wind shear events in US offshore wind energy areas and the role of induced stratification, Wind Energy Science, 6, 1043–1059, 2021.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140–158, https://doi.org/10.1016/j.rse.2010.10.017, 2012.

Dragaud, I. C. D. V., Soares da Silva, M., Assad, L. P. d. F., Cataldi, M., Landau, L., Elias, R. N., and Pimentel, L. C. G.: The impact of SST on the wind and air temperature simulations: a case study for the coastal region of the Rio de Janeiro state, Meteorology and Atmospheric Physics, 131, 1083–1097, https://doi.org/10.1007/s00703-018-0622-5, 2019.

Fiedler, E. K., Mao, C., Good, S. A., Waters, J., and Martin, M. J.: Improvements to feature resolution in the OSTIA sea surface temperature analysis using the NEMOVAR assimilation scheme, Quarterly Journal of the Royal Meteorological Society, 145, 3609–3625, https://doi.org/https://doi.org/10.1002/qj.3644, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3644, 2019.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gerber, H., Chang, S., and Holt, T.: Evolution of a Marine Boundary-Layer Jet, Journal of the Atmospheric Sciences, 46, 1312–1326, https://doi.org/10.1175/1520-0469(1989)046<1312:EOAMBL>2.0.CO;2, publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, 1989.

Gutierrez, W., Araya, G., Kiliyanpilakkil, P., Ruiz-Columbie, A., Tutkun, M., and Castillo, L.: Structural impact assessment of low level jets over wind turbines, Journal of Renewable and Sustainable Energy, 8, 023 308, https://doi.org/10.1063/1.4945359, 2016.

Gutierrez, W., Ruiz-Columbie, A., Tutkun, M., and Castillo, L.: Impacts of the low-level jet's negative wind shear on the wind turbine, Wind Energy Science, 2, 533–545, https://doi.org/10.5194/wes-2-533-2017, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803, 2020.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, 2008.

Kaellstrand, B.: Low level jets in a marine boundary layer during spring, Contributions to Atmospheric Physics, 71, https://www.osti.gov/etdeweb/biblio/660911, 1998.

Kain, J. S. and Fritsch, J. M.: Convective parameterization for mesoscale models: The Kain-Fritsch scheme, in: The representation of cumulus convection in numerical models, pp. 165–170, Springer, 1993.

Kikuchi, Y., Fukushima, M., and Ishihara, T.: Assessment of a Coastal Offshore Wind Climate by Means of Mesoscale Model Simulations Considering High-Resolution Land Use and Sea Surface Temperature Data Sets, Atmosphere, 11, 379, https://doi.org/10.3390/atmos11040379, number: 4 Publisher: Multidisciplinary Digital Publishing Institute, 2020.

Lantz, E. J., Roberts, J. O., Nunemaker, J., DeMeo, E., Dykes, K. L., and Scott, G. N.: Increasing wind turbine tower heights: Opportunities and challenges, 2019.

Li, H., Claremar, B., Wu, L., Hallgren, C., Körnich, H., Ivanell, S., and Sahlée, E.: A sensitivity study of the WRF model in offshore wind modeling over the Baltic Sea, Geoscience Frontiers, 12, 101 229, 2021.

Lombardo, K., Sinsky, E., Edson, J., Whitney, M. M., and Jia, Y.: Sensitivity of Offshore Surface Fluxes and Sea Breezes to the Spatial Distribution of Sea-Surface Temperature, Boundary-Layer Meteorology, 166, 475–502, https://doi.org/10.1007/s10546-017-0313-7, 2018.

Miller, S. T. K., Keim, B. D., Talbot, R. W., and Mao, H.: Sea breeze: Structure, forecasting, and impacts, Reviews of Geophysics, 41, https://doi.org/https://doi.org/10.1029/2003RG000124, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2003RG000124, 2003.

Murphy, P., Lundquist, J. K., and Fleming, P.: How wind speed shear and directional veer affect the power production of a megawatt-scale operational wind turbine, Wind Energy Science, 5, 1169–1190, https://doi.org/10.5194/wes-5-1169-2020, 2020.

Murphy, S. C., Nazzaro, L. J., Simkins, J., Oliver, M. J., Kohut, J., Crowley, M., and Miles, T. N.: Persistent upwelling in the Mid-Atlantic Bight detected using gap-filled, high-resolution satellite SST, Remote Sensing of Environment, 262, 112 487, https://doi.org/10.1016/j.rse.2021.112487, 2021.

610 Nakanishi, M. and Niino, H.: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog, Boundary-Layer Meteorology, 119, 397–407, 2006.

Nakanishi, M. and Niino, H.: Development of an improved turbulence closure model for the atmospheric boundary layer, Journal of the Meteorological Society of Japan. Ser. II, 87, 895–912, 2009.

Nunalee, C. G. and Basu, S.: Mesoscale modeling of coastal low-level jets: implications for offshore wind resource estimation, Wind Energy,

615 17, 1199–1216, https://doi.org/https://doi.org/10.1002/we.1628, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.1628, 2014.

Optis, M., Bodini, N., Debnath, M., and Doubrawa, P.: Best Practices for the Validation of US Offshore Wind Resource Models, Tech. rep., Tech. Rep. NREL/TP-5000-XXXXX, NREL–National Renewable Energy Laboratory . . . , 2020.

Park, R. S., Cho, Y.-K., Choi, B.-J., and Song, C. H.: Implications of sea surface temperature deviations in the prediction of wind and precipitable water over the Yellow Sea, Journal of Geophysical Research: Atmospheres, 116,

620 https://doi.org/https://doi.org/10.1029/2011JD016191, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011JD016191, 2011.

Pichugina, Y. L., Brewer, W. A., Banta, R. M., Choukulkar, A., Clack, C. T. M., Marquis, M. C., McCarty, B. J., Weickmann, A. M., Sandberg, S. P., Marchbanks, R. D., and Hardesty, R. M.: Properties of the offshore low level jet and rotor layer wind shear as measured by scanning Doppler Lidar, Wind Energy, 20, 987–1002, https://doi.org/https://doi.org/10.1002/we.2075, _eprint:

625 https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.2075, 2017.

Ping, B., Su, F., and Meng, Y.: An Improved DINEOF Algorithm for Filling Missing Values in Spatio-Temporal Sea Surface Temperature Data, PLOS ONE, 11, e0155 928, https://doi.org/10.1371/journal.pone.0155928, publisher: Public Library of Science, 2016.

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., Peckham, S. E., Grell, G. A., Michalakes, J., Trahan, S., Benjamin, S. G., Alexander, C. R., Dimego, G. J., Wang, W., Schwartz, C. S., Romine,

630 G. S., Liu, Z., Snyder, C., Chen, F., Barlage, M. J., Yu, W., and Duda, M. G.: The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions, Bulletin of the American Meteorological Society, 98, 1717–1737, https://doi.org/10.1175/BAMS-D-15-00308.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2017.

Schmit, T. J., Gunshor, M. M., Menzel, W. P., Gurka, J. J., Li, J., and Bachmeier, A. S.: Introducing the next-generation Advanced Baseline Imager on GOES-R, Bulletin of the American Meteorological Society, 86, 1079–1096, 2005.

635 Schmit, T. J., Li, J., Li, J., Feltz, W. F., Gurka, J. J., Goldberg, M. D., and Schrab, K. J.: The GOES-R Advanced Baseline Imager and the Continuation of Current Sounder Products, Journal of Applied Meteorology and Climatology, 47, 2696–2711, https://doi.org/10.1175/2008JAMC1858.1, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology, 2008.

Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., and Lebair, W. J.: A closer look at the ABI on the GOES-R series,

640 Bulletin of the American Meteorological Society, 98, 681–698, 2017.

Schoenberg Ferrier, B.: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description, Journal of Atmospheric Sciences, 51, 249–280, 1994.

Shimada, S., Ohsawa, T., Kogaki, T., Steinfeld, G., and Heinemann, D.: Effects of sea surface temperature accuracy on offshore wind resource assessment using a mesoscale model, Wind Energy, 18, 1839–1854, https://doi.org/https://doi.org/10.1002/we.1796, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.1796, 2015.

Stark, J. D., Donlon, C. J., Martin, M. J., and McCulloch, M. E.: OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system, in: OCEANS 2007 - Europe, pp. 1–4, https://doi.org/10.1109/OCEANSE.2007.4302251, 2007.

Stull, R. B. et al.: Practical meteorology: an algebra-based survey of atmospheric science, 2015.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., Cuenca, R., et al.: Implementation and verification of the unified NOAH land surface model in the WRF model, in: 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, vol. 1115, pp. 2165–2170, 2004.

Xia, G., Draxl, C., Optis, M., and Redfern, S.: Detecting and Characterizing Sea Breezes Over the US Northeast Coast with Implication for Offshore Wind Energy, Wind Energy Science Discussions, pp. 1–23, 2021.