Response to the comments of reviewer 1

We thank the reviewer Xin Zhou for this positive and thoughtful review. We provide a detailed response to all comments below with our responses shown in blue. Changes to the manuscript text are included in italic fonts.

This manuscript presents a rigorous, multi-season evaluation of the NOAA HRRR model v4 in the U.S. Northeast coastal marine boundary layer using unique, high-quality remote sensing observations from WFIP3. The work addresses a critical knowledge gap and delivers clear, physically consistent results relevant to offshore wind energy, aviation, and marine forecasting. The paper is well-structured, technically sound, and I have only four major comments:

Major Comments:

1. The manuscript suggests that overestimated SST contributes to warm biases near the bottom of the stability layer and therefore to underestimated static stability. This is plausible and supported by the weak negative correlation between SST error and stability error at NANT and BLOC. However, the correlations are modest, and the relationship is not evident at RHOD. Therefore, the causal wording should be softened in places. For example, instead of implying that SST errors are the primary cause of stability errors, the authors could state that SST errors are "a likely contributing factor," particularly for island sites and southwesterly flow regimes. Other possible contributors, such as boundary-layer mixing, surface-layer parameterization, vertical resolution, land–sea mask representation, and horizontal advection errors, should also be acknowledged. This would make the interpretation more balanced and avoid over-attribution.

We modified our wording, now saying in Sect. 3.3: *This indicates that many of the cases with too-weak static stability in the model were linked to an overestimation of SST and consequently warm biased low-level temperature, suggesting that the specification of the lower boundary conditions are likely a contributing factor to the near-surface model errors.*
In the summary, we highlight now that the correlation between SST and stability errors at the island sites was weak and we make clear that better SSTs could lead to reduced errors of wind and temperature at the island sites, since this is where the weak correlation between SST and static stability errors was found.
We added the following sentence to the summary mentioning other possible factors contributing to model errors: *Other possible contributing factors could be the boundary-layer and surface-layer parameterization, land-sea mask representation, initial and boundary conditions, and errors in horizontal advection.*

2. The authors use a modified Bonner-type LLJ criterion with a wind speed threshold of 8 m s^{-1}, which is lower than in some previous studies. The justification is that many profiles with clear LLJ structure would otherwise be missed. This is reasonable, but the choice of threshold can strongly influence LLJ frequency, CSI, frequency bias, and the classification of weak LLJs. Since about 35–40% of observed LLJs are classified as the weakest class 0, the results may be sensitive to this threshold. I suggest adding a short sensitivity test, perhaps in the supplement,

showing how LLJ frequency and model skill change if a 10 m s^{-1} threshold is used. Even a brief comparison would help readers assess whether the main conclusions are robust to the LLJ definition.

We tested the sensitivity of the model errors and our conclusions to the LLJ criteria by repeating the analysis for LLJ classes 1 through 4 only and produced the respective plots (Figs. 1 to 4 here and Figs. 10 to 13 in the manuscript) using a minimum nose wind speed threshold of 10 m/s. The new plots are included in this response (Fig. 1 to 4). As expected the frequency of observed and simulated LLJ decreased, now ranging between 10 and 13 % in the observations and 8-10 % in the model (Fig. 1), compared to 18-21 % and 14-16 % when class 0 is included. The monthly cycle of average nose height and LLJ frequency as well as the diurnal cycle of LLJ frequency did not change much, except for having slightly lower relative frequencies, as expected. CSI and FB also remained very similar. The HRRR model errors with respect to nose height, wind speed at nose height, shear below and above nose, and stability below and above nose were also very similar (Fig. 2 here and Fig. 11 in the manuscript). The same is true when looking at fixed layer errors (Fig. 3 here and Fig. 12 in the manuscript) or the Bulk Richardson number distribution (Fig. 4 here and Fig. 13 in the manuscript). Because of the overall insensitivity of the errors and conclusions to the chosen nose wind speed threshold, we decided not to include the additional plots using a different nose wind speed threshold in the supplement. Instead we included this additional information in Sect. 2.4: *We tested the sensitivity of our results to the speed threshold by performing the same analysis using a 10 m/s threshold for nose wind speed, that is, only considering LLJ that fall into classes 1 through 4 . While this reduced the number of LLJs in the data set, the model errors were largely the same when removing the weakest LLJ class.*

Response to the comments of reviewer 2

We thank the reviewer Dr. Jeffrey Freedman for this thoughtful review and the useful comments and suggestions. We provide a detailed response to all comments below with our responses shown in blue. Major changes to the manuscript text are included in italic fonts.

This paper examines the National Oceanic and Atmospheric Administration's (NOAA) High Resolution Rapid Refresh (HRRR) model in the context of its performance within the marine atmospheric boundary layer (MABL) over the coastal and offshore waters south of New England. The investigators leverage data provided by the deployment of a variety of land-based and offshore instruments during the 3rd Wind Forecast Improvement Project (WFIP3) to quantify the HRRR's forecast performance over a period of one year. The paper provides a thorough examination of HRRR's forecasts on diurnal and monthly timescales, focusing on how well the model predicts the wind and thermodynamic profiles at several sites within the WFIP3 domain. The paper provides critical information regarding the performance of HRRR over offshore waters, historically a region suffering from a dearth of observational data.

Thank you for this positive assessment.

However, the narrative suffers from some lapses in analyses, including an incomplete literature review, use of adjectives (e.g., "excellent") without quantitative support, arbitrary use of certain threshold criteria (such as what constitutes a low-level jet [LLJ]), and lack of more descriptive figures confirming hypotheses put forth in the manuscript (e.g., field plots showing SST and/or wind speed distribution at critical levels). More specific comments follow below.

Thank you for catching oversights in the literature review and the valuable comments on quantitative support and threshold criteria. As we describe in the manuscript, we use a method based on maximum wind speed and wind speed decrease above the nose to detect LLJ, which was used in several previous studies in the area. We modified the text in Sect. 2.4 to make this more clear. We show that errors in the SST, which are inherited in the HRRR model, can cause errors in low-level stability and wind and hypothesize that improved SST as input to the HRRR model (such as from OSTIA which agrees well with in situ buoy measurements in the area) may lead to improved forecasts (this is included in Sect. 5). We find that the SST errors vary spatially in our area (Sect. 2.3). We now include that one possible reason for this might be related to coastal cold water upwelling and reference McCabe and Freedman (2025) who showed that the SST bias between OSTIA and HRRR varied spatially in the New York Bight just south of our investigation area during summer sea breezes possibly owing to coastal upwelling. McCabe and Freedman (2026) showed for the New York Bight that SST gradients can impact the sea breeze and associated LLJ structure and strength using sensitivity studies. We now reference these studies throughout the manuscript. Instead of repeating the analysis of spatial SST fields from OSTIA and HRRR conducted by McCabe and Freedman (2025) for a slightly different area, we focus in our study on the evaluation of atmospheric conditions using point observations and link those to local SST errors.

Page 2, lines 44-45: "Using 9 years of buoy-based lidar data in the New York Bight, Colle and Novak (2010)…." This is untrue. No such data was available or was used by Colle and Novak (2010). They used 10 years (April 1997 - August 2006) of data from a C-MAN station (Ambrose Light, or ALSN6), with wind observations obtained from an anemometer height of 29 m. (Note Ambrose Light was removed and in 2008 after a ship hit it, replaced by buoy 44065.)

We replaced 'buoy-based lidar data' with 'anemometer data collected on an offshore light station and a buoy'. In the manuscript by Colle and Novak (2010) it is described that data were available from 1997 to 2006, but since buoy 44025, which was also used for the analysis, was not available in 2005, the overall period for which the analysis was conducted spanned 9 years.

Page 2-3, lines 47 - 59: There is no reference to recent papers by McCabe and Freedman (2023, 2025a) which describe detailed observational analysis and modeling sensitivity studies focused on the Low Level Jet (LLJ) characteristic of the New York Bight, coastal and offshore waters that overlap with the WFIP3 study region. This is quite surprising given the topic and geographic location of this paper. Note also that the New York State Energy Research and Development Authority (NYSERDA) deployed two buoys over a three-year period (2019 - 2022; see https:// www.nyserda.ny.gov/All-Programs/Offshore-Wind/Focus-Areas/Regional-Collaboration/Siting Offshore-Wind) in the coastal waters south of Long Island and east of New Jersey.

We added the additional literature to the introduction: *Sea-breeze circulations in the New York Bight are frequently associated with LLJs, especially from late spring through fall, which was investigated by McCabe and Freedman (2023, 2025) using coastal and buoy-based lidars and surface stations and modeling sensitivity studies. The inland propagation of the sea breeze and the LLJ intensity are found to be sensitive to cold water upwelling (McCabe and Freedman 2026).*

Page 3, lines 84-85: A deeper mixed layer over land (here, Block Island) is not necessarily an "internal boundary layer."

The AMS glossary (https://glossary.ametsoc.org/wiki/internal-boundary-layer/) defines an 'internal boundary layer' as: 'A layer that forms within an existing boundary layer due to horizontal advection across a discontinuity in some surface property (e.g., aerodynamic roughness length or surface heat flux). IBLs originate at the discontinuity and grow in thickness in the downstream direction until being homogenized with the existing boundary layer via ambient mixing.' Given the change in surface properties due to the islands, we think that both sites may be impacted by the formation of an internal boundary layer, BLOC potentially more than NANT due to the larger distance from the discontinuity (coast line) for southwesterly flow. We agree that a deeper mixed layer over land is not necessarily an internal boundary layer, and we do not state this in our manuscript.

Page 5, section 2.1.1: No information provided regarding the manufacturer or specifications of the radar wind profilers referenced here and used in this study.

Given the large number of different remote sensing instruments used for the retrieval of wind and temperature profiles, we decided to include a table in the Supplemental specifying the different manufacturers and specifications for the variables of interest used with WINDoe and TROPoe for the sake of brevity. The table is pasted here for convenience:

Table 1. Overview of manufacturer, variables of interest and spatial resolution and maximum range of the variables of interest from the remote sensing instruments that are used as input to WINDoe and TROPoe. For the radar wind profilers, specification for high- and low-resolution model are given with the latter in brackets. Horizontal wind components are denoted with u and v, radial velocity along the line of sight with rv, and cloud base height with CBH. Maximum range is the theoretical range that can be achieved based on the settings, the actual range depends on atmospheric conditions and is usually lower.

Site	Instrument	Manufacturer, model	Variable of interest	Spatial resolution (m)	Maximum range (m)
NANT	Radar wind profiler	Built by NOAA Physical Sciences Laboratory, 915 MHz	u, v	58 (97)	121-2,925 (164-6,097)
	Scanning Doppler lidar	Halo Photonics, XR+	rv	30	15-12,000
	Profiling Doppler lidar	Vaisala, WindCube V2.1	u, v	20	47-197
	Infrared Spectrometer	LRTech, ASSIST-II	radiance	N/A	N/A
	Ceilometer	Vaisala, CL51	CBH	5	40-13,000
BLOC	Radar wind profiler	Built by NOAA Physical Sciences Laboratory, 915 MHz	u, v	58 (97)	121-2,925 (164-6,097)
	Scanning Doppler lidar	Halo Photonics, XR	rv	30	15-6,000
	Profiling Doppler lidar	Vaisala, WindCube V1	u, v	20	40-220
	Infrared Spectrometer	LRTech, ASSIST-II	radiance	N/A	N/A
	Ceilometer	Vaisala, CL51	CBH	5	40-13,000
RHOD	Radar wind profiler	Vaisala, LAP-3000 915 MHz	u, v	62 (97)	103-2,788 (195-6,031)
	Scanning Doppler lidar	SgurrEnergy, Galion	rv	30	15-6,000
	Profiling Doppler lidar	ZX Lidars, ZephIR-300	u, v	20	10-280
	Infrared Spectrometer	LRTech, ASSIST-II	radiance	N/A	N/A

Page 5, section 2.1.2: Same comment as above.

See above.

Page 6, section 2.1.5: Same comment as above.

See above.

Page 7, Figure 2: This figure is very difficult to read, especially as the legends in (a) and (b) overlap with the data availability (a) and uncertainty profiles (b).

We modified that figure to prevent the legends from overlapping with the data availability and uncertainty profiles. The updated figure is pasted below.

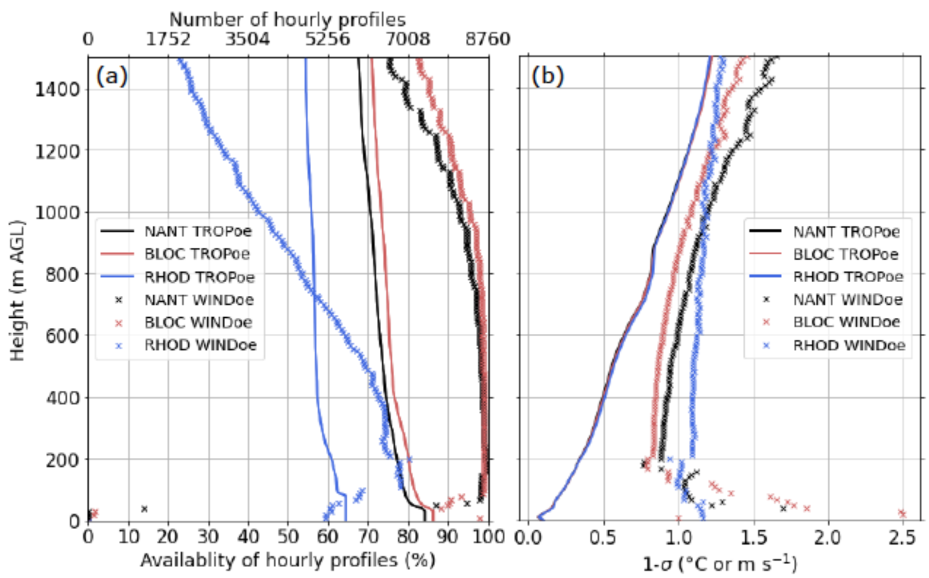

Figure 2. Profiles of (a) availability and (b) mean 1-σ uncertainty of hourly temperature and wind observations at NANT, BLOC, and RHOD retrieved with TROPoe and WINDoe. The 1-σ uncertainties for the three TROPoe retrievals in (b) are essentially identical and thus overlap. In (a), missing data are related to power outages and hardware and software issues. In addition, the availability of temperature profiles is limited by clouds that are opaque in the infrared, and the availability of wind profiles is limited by the range of the individual instruments used in WINDoe.

Page 8, line 162: The use of the term "excellent accuracy", lacks quantitative context. (The authors in Letizia et al. (2025[b]) appear to use the term "satisfactory.")

We rewrote this sentence and 'excellent accuracy' does not occur any more. The sentence now is:
These results were also corroborated by comparisons with a 135-m met tower from Letizia et al. (2026), who demonstrated biases less than 0.3 K at multiple heights.

Page 8, line 168: Why the LWP threshold of > 8 g m_{-2}?

This threshold is based on findings described by Turner (2007), Turner and Löhnert (2014), and Adler et al. (2024). The IRS is very sensitive to LWP and the sensitivity of the IRS is much smaller than the IRS uncertainty for LWP values of less than approximately 50 g/m2 (Fig. 1 in Turner, 2007). However, because scattering is not included in TROPoe for computational reasons, clouds with a lower LWP (< 10 g/m2) already impact the retrieved temperature and humidity profiles (Fig. 9 in Turner and Löhnert, 2014). Hence, we chose a conservative threshold of 8 g/m2 to filter for clouds that we empirically found to be suitable for a wide range of locations and conditions (Adler et al., 2024). We added the reference Turner (2007) as well as an explanation on the chosen threshold:

IRS-based retrievals have little to no information content above cloud base depending on the optical depth of the cloud (Turner, 2007). … The threshold value was determined empirically and found to work well for a variety of locations and conditions (Adler et al. 2024).

Page 8, lines 189-191: "…the number of operational radiosondes at Upton, New York, was not sufficient…." This is confusing. Given the high-resolution radiosonde data (~ 5 m in the vertical) available over the last several years, at ~365 days x twice daily launchings x ~10 years, that works out to approximately 7300 soundings. HRRR profiles at such coarse vertical resolution (10 levels below 1500 m) would seem to be a much inferior data set. Perhaps I am missing something here?

WINDoe requires monthly priors to account for the seasonal variability in atmospheric dynamics. This means that for each month less than 1000 radiosonde profiles would be available for the prior computation which is not sufficient since it would be the same order of magnitude as the state vector in our case. This is why we chose to use HRRR profiles instead. We added the following explanation to the text to make this clearer:
As a rule of thumb, the number of profiles should be at least 1 order of magnitude larger than the state vector (752 in our case). This means that even a decade of twice-daily radiosonde profiles would only provide the same order of magnitude for each monthly prior.

Page 9, equation 1: Based upon Archer (2025), isn't this still an approximation (there, should use the ≈ _symbol)?

Changed.

Page 9, line 225: Given the HRRR's limited number of levels below 1500 m, doesn't this make an accurate determination of the LLJ nose and structure problematic?

We agree that the limited vertical resolution of the HRRR model is one of the challenges for the prediction of the LLJs and may impact the ability of the HRRR to resolve sharp gradients in wind and temperature and potentially to place the jet nose too high. Despite the limited number of levels, the HRRR was able to correctly produce the overall structure and evolution of low-level wind and temperature (Figs. 4 and 5) clearly showing the jet shape during the warmer months and to correctly predict the occurrence of a large number of LLJs (Table 2 and Fig. 10). In Sect. 4.2, we discuss the impact of the coarse vertical grid spacing on the overestimation of nose height and underestimation of nose wind speed in the HRRR model.

Page 10, lines 230 - 235: See McCabe and Freedman (2025a) regarding OSTIA.

In Sect. 2.1.6 we now reference McCabe and Freedman (2026), who compared OSTIA SST to SST on floating lidar buoys in the New York Bight for a 3-year period and found a good agreement between in-situ observations and OSTIA, which is consistent with our results when comparing OSTIA to buoy observations in the WFIP3 region:
This is consistent with McCabe and Freedman (2026) who reported good agreement between daily OSTIA SST and buoy-based measurements in the New York Bight.

In Sect. 2.3, when describing the HRRR model we added the following sentence:
McCabe and Freedman (2026) found that the bias in SST between HRRR and OSTIA varied spatially in the New York Bight during sea-breeze events with a warm bias near the coasts and a cold bias further offshore possibly owing to coastal upwelling.

Page 10, section 2.4: See McCabe and Freedman (2023, 2025a).

We apply a criteria for LLJ identification that was similar to what was used in other studies on coastal LLJs in the area e.g. by Quint et al. 2025 and Bodini et al. 2026. As pointed out by Hallgren et al. (2023), there is no consensus regarding how LLJs should be defined and used criteria include falloff (the decrease in wind speed above and below core) as well as shear (wind speed change between two vertical levels). The LLJ characteristics (frequency, nose height, strength) not only depend on the applied criteria but also on the vertical window from which the LLJ is extracted. The criteria used in our study is more restrictive than some other methods applied to coastal LLJ in this region, e.g. by McCabe and Freedman (2023, 2025) who defined the LLJ as a wind speed maximum between 150 and 300 m.
We added the following sentences to Sect 2.4 to include this information:
To detect LLJs in the observed and simulated hourly wind profiles, we used a criterion based on maximum wind speed and wind speed decrease above the nose, which was first developed by Bonner (1968) and which has also been applied in previous studies on coastal LLJs in the area (e.g., Quint et al. 2025, Bodini et al. 2026). Other studies used a less restrictive method to detect LLJs in the area, e.g., McCabe and Freedman (2023, 2025) who defined the LLJ as a wind speed maximum between 150 and 300 m. As pointed out by Hallgren et al. (2023), there is no consensus regarding how LLJs should be defined and the choice of criteria as well as the vertical window from which the LLJs are extracted impact

the LLJ frequency and morphology. Hence care should be taken when making cross-study comparisons.

Page 11, lines 244 - 245 and Table 1: Why not use wind speed shear values to constrain LLJ identification? The above-nose wind speed decrease constraint seems rather arbitrary (although any filter would have some subjective characteristics). A more thorough explanation of the classes regarding LLJ identification would seem prudent.

The criteria we used for LLJ identification are similar to what was used in other studies on coastal LLJs in the area (e.g. Quint et al. 2025; Bodini et al. 2026) that goes back to Bonner (1968) and is based on maximum wind speed and wind speed decrease above nose height. Other studies frequently call the wind speed decrease above nose height 'shear', but take the difference between the maximum and minimum wind speed above and report it in m/s. Since this is not 'shear', which would require the division by the layer depth, we use the term 'Above-nose wind speed decrease (m/s)' instead, which we think is more appropriate. We added the information that our LLJ classification followed that of Quint et al. 2025 and Bodini et al. 2026, but with an additional lowest class added to account for the weaker LLJs. For operational model development, the weaker LLJ are also of interest, which is why we decided to include them. We tested the sensitivity of our results to the nose wind speed threshold by performing the same analysis using a 10 m/s threshold for nose wind speed, that is, only considering LLJs that fall into classes 1 through 4. While this reduced the number of LLJs in the data set, the model errors were largely the same when removing the weakest LLJ class. This information is now added to Sect. 2.4 following a suggestion from reviewer 1.

Page 11, lines 257 - 259: Using two different layers for static stability (100 m - 300 m) and wind shear (50 m - 200 m) is inconsistent. The explanation is also arbitrary, and more hand-waving than science. And wouldn't this make comparison with the HRRR problematic (e.g., Bulk Richardson analysis in section 4.3)?

During our analysis we tried several different combinations of layers, including starting near the surface for both variables or going up to 300 m for wind. Each combination was associated with some challenges. When choosing the lower boundary close to the surface model errors were present due to the island impact. Since our goal was not to evaluate the representation of the island in the model but to investigate model errors in the marine boundary layer, we chose a higher lower boundary. When choosing the upper boundary for wind at 300 m, the LLJ nose height was sometimes lower than that height which impacted the shear error. As we explain in the manuscript, we choose the different boundaries for temperature and wind based on the typical wind and temperature profile shape and data availability. However, to satisfy the reviewer's comment we modified the layer to compute static stability to be consistent with the shear layer. While the main results and overall conclusions do not change, the specific values were sensitive to the choice of layers as can be expected.

Page 13, Figure 4: It would be interesting to see the wind shear *exponent* (common in wind energy applications) as a variable in the analysis, although I understand this would be redundant.

We agree with the reviewer that adding the wind shear exponent would be a valuable addition for wind energy applications. However since we are not focusing on wind energy applications, but rather model evaluation, and as it would be redundant as the reviewer stated, we prefer to not add an additional variable to the analysis for the sake of brevity.

Page 16, lines 316 -318: "Another reason [for the overestimation of 50 m wind speeds] could be the vicinity of Block Island to Long Island…." This assertion could be tested by examining HRRR model field plots (x-y) to see if any channeling is in fact present. Although eastern Long Island and Block Island are quite flat (maximum elevations are ~ 50 m), frictional effects could possibly induce some "channeling."

Testing the potential channeling effects of Long Island could make for an interesting sensitivity study, e.g. by changing the roughness of Long Island or even removing the land. However, we think that this is beyond the scope of the current manuscript. Instead, we decided to remove the speculative impact of channeling from the text.

Page 21, lines 366 - 367: Same comment as above.

We removed the speculative impact of a channeling due to Long Island.

Page 21, lines 363- 371: SST overestimation may be a result of cold water upwelling given the wind trajectory (southwesterly). See McCabe and Freedman (2025b). Model grids that extend simultaneously over land and water can be problematic.

We fully agree that model grids that extend over land and water are problematic. This is one of the reasons why we are evaluating the HRRR model in this region. We added the following sentence:
The relationship between stability errors, SST errors, and southwesterly flow indicates that those could be linked to coastal cold water upwelling that may occur under persistent southwesterly flow as shown by e.g. McCabe and Freedman (2026) for the New York Bight.

Page 22, lines 389 - 390: Wouldn't a secondary peak in LLJ frequency (late summer) be inconsistent with the LLJ mechanism being driven by SST-land temperature differences (given SSTs peak in late August early September in this region)? Southwesterly LLJs can also be associated with pre-frontal systems in the northeast US.

We agree that LLJs along the northeast coast may be linked to frontal passage and state this in the introduction. Motivated by the comment, we investigated the wind direction distribution at nose height for September and found that LLJs in September were predominantly from a

northeasterly direction. This indicates that those were not driven by a land-sea temperature difference. Further inspection of synoptic maps indicated that multiple frontal systems affected the investigation area in September, especially in late September, which could explain the frequent northeasterly LLJ. We added this information to the text in Sect. 4.1:
LLJ frequency showed a secondary peak in September, especially at NANT, which was also present in the model, although underestimated. LLJs in September were predominantly associated with northeasterly flow (not shown) and likely linked to frontal passages. This distinguishes them from the LLJs that occur in late spring and early summer, which are predominantly from the southwest and likely driven by the land-sea temperature difference between the warm land and still cool ocean surface.

Page 27, line 455: Use of “Unprecedented” here seems overly hyped, given previous deployments by NYSERDA and offshore wind developers in the region.

We rephrased this sentence to more clearly highlight what is new about the used data set:
We used unique continuous observations of temperature and wind profiles throughout the depth of the marine boundary layer at three sites, two of which were deployed on islands (NANT and BLOC) and one was deployed at the coast (RHOD) for the Third Wind Forecast Improvement Project. These data extended beyond the wind-only, buoy-based lidar measurements previously available in the region, which were limited to the lowest few hundred meters.

Page 27, lines 464-465: “The HRRR model did an excellent job in capturing these characteristics on average.” What does this mean?

This statement refers to the 24-h monthly composites which show that the HRRR was able to capture the overall temporal evolution and vertical structure, e.g. the LLJ and low-level inversion as well as the weak diurnal cycle. We rewrote the sentence to:

While the HRRR model well captured the overall seasonal and diurnal evolution of temperature and wind in the marine boundary layer when computing monthly composites, some differences were evident, especially during the warmer months when the modeled low-level shear and static stability were too weak.

Page 28, line 488: “…overall great performance….” Hyperbole?

We rephrased this sentence to:
We conclude that the HRRR's ability to forecast wind and temperature in the marine boundary layer is sufficient for most applications, but we identified LLJs as one atmospheric phenomenon with clear systematic errors.

References:

Adler, B., Turner, D. D., Bianco, L., Djalalova, I. V., Myers, T., andWilczak, J. M.: Improving solution availability and temporal consistency of an optimal-estimation physical retrieval for ground-based thermodynamic boundary layer profiling, Atmos. Meas. Tech., 17, 6603–6624, https://doi.org/10.5194/amt-17-6603-2024, 2024.

Colle, B. A. and Novak, D. R., 2010: The New York Bight jet: climatology and dynamical evolution, *Mon. Wea. Rev*., 138, 2385–2404, https://doi.org/10.1175/2009MWR3231.1.

Hallgren, C., Aird, J. A., Ivanell, S., Körnich, H., Barthelmie, R. J., Pryor, S. C., and Sahlée, E.: Brief communication: On the definition of the low-level jet, Wind Energ. Sci., 8, 1651–1658, https://doi.org/10.5194/wes-8-1651-2023, 2023.

McCabe, E., and J. M. Freedman, 2023: Development of an Objective Methodology for Identifying the Sea Breeze Circulation and Associated Low-Level Jet in the New York Bight. *Wea. Forecasting, 38, 571–589, https://doi.org/10.1175/WAF-D-22-0119.1*.

McCabe, E. J., and J. M. Freedman, 2025a: Quantifying the Uncertainty in the Weather Research and Forecasting Model under Sea Breeze and Low-Level Jet Conditions in the New York Bight: Importance to Offshore Wind Energy. *Wea. Forecasting*, 40, 425–450, https://doi.org/10.1175/WAF-D-24-0086.1.

McCabe, E., and J. M. Freedman, 2025b: Assessing the Impact of Cold-Water Coastal Upwelling along the New Jersey Coastline: Amplification of the Sea Breeze and Low-Level Jet. *J. of Atmos. and Oceanic Tech.*; DOI 10.1175/JAMC-D-25-0069.1.

Turner, D.: Improved ground-based liquid water path retrievals using a combined infrared and microwave approach, J. Geophys. Res., 112, https://doi.org/10.1029/2007JD008530, 2007.