A North Sea in situ evaluation of the Fitch Wind Farm Parametrization within the Mellor–Yamada–Nakanishi–Niino and 3D Planetary Boundary Layer schemes

Nathan J. Agarwal[1], Julie K. Lundquist[1,2], Timothy W. Juliano[3], and Alex Rybchuk[2]

[1]Johns Hopkins University, Baltimore, MD, United States
[2]National Renewable Energy Laboratory, Golden, CO, United States
[3]U.S. National Science Foundation National Center for Atmospheric Research, Boulder, CO, United States

Correspondence: Nathan J. Agarwal (nagarw22@jh.edu)

Abstract. Wind resource assessments and wind power forecasts that account for wind farm wakes are sensitive to the choice of planetary boundary layer (PBL) scheme. This work compares the one-dimensional Mellor–Yamada–Nakanishi–Niino (MYNN) PBL scheme with a three-dimensional PBL (3DPBL) scheme, evaluating predictions made with both schemes against two sets of North Sea in situ observations of wind farm wakes. The optimal PBL scheme varies based on the observations (FINO1 tower vs. aircraft), the quantity of interest (wind speed vs. turbulence kinetic energy [TKE]), and the error metric (bias, centered root mean square error [$cRMSE$], R^2, and earth mover's distance [EMD]). Whereas 3DPBL wind speeds outperform MYNN wind speeds with respect to the $cRMSE$ at the FINO1 site within the turbine rotor layer, 3DPBL TKE bias underperforms MYNN TKE bias when compared to aircraft observations. Wind speeds in the aircraft region are ambiguous as to which PBL scheme is optimal. Aircraft MYNN wind speeds outperform 3DPBL wind speeds with respect to R^2 and $cRMSE$ but underperform with respect to bias and EMD.

1 Introduction

More energy generation technologies are being developed and deployed as global energy demand continues to increase. This new energy generation is also becoming increasingly renewable. Offshore wind is one renewable technology that continues to grow, especially in the North Sea and Baltic areas (Backwell et al., 2024). Wind and wake forecasts are becoming increasingly

crucial in project planning. These forecasts are sensitive to the underlying wind resource (Optis and Perr-Sauer, 2019) and meteorological wind turbine wake models are continually in development (Fischereit et al., 2022a).

Wind turbine impacts on the weather are often expressed through wind farm parameterizations (WFPs) (Fischereit et al., 2022a) within numerical weather prediction (NWP) models like the Weather Research and Forecasting (WRF) model (Skamarock et al., 2021). Most WFPs treat wind farms as elevated sources of turbulence and sinks of momentum (Baidya Roy et al., 2004; Fitch et al., 2012; Adams and Keith, 2013), especially after the importance of treating the wind farm as an elevated source of drag (rather than proximate to the surface) was demonstrated (Fitch et al., 2013). WFPs may include (Fitch et al., 2012) or exclude (Volker et al., 2015) an explicit source of turbulence kinetic energy (TKE), and this decision drives differences in both wind speeds and TKE (Shepherd et al., 2020; Pryor et al., 2020; Larsén and Fischereit, 2021; García-Santiago et al., 2024; Quint et al., 2024). WFPs also differ based on if and how they represent subgrid-scale processes (Abkar and Porté-Agel, 2015; Volker et al., 2015; Pan and Archer, 2018; Ma et al., 2022), which is important when multiple turbines need to be represented within one grid cell.

Wind fields predicted from NWP simulations are also sensitive to modeling choices within the WFP. Simulated wind fields depend on horizontal and vertical grid cell spacing (Lee and Lundquist, 2017; Mangara et al., 2019; Tomaszewski and Lundquist, 2020; Pryor et al., 2020), the strength of the explicit TKE source (Fitch et al., 2012; Vanderwende et al., 2016; Rajewski et al., 2016; Mangara et al., 2019; Tomaszewski and Lundquist, 2020; Siedersleben et al., 2020), the advection option (Siedersleben et al., 2020; Archer et al., 2020; Larsén and Fischereit, 2021), and the planetary boundary layer (PBL) scheme choice (Peña et al., 2023). Within the PBL scheme, parameterizations of physical quantities can further affect results. For example, the turbulence dissipation rate, ϵ, affects modeled wind fields within the Mellor–Yamada–Nakanishi–Niino (MYNN) scheme (Yang et al., 2017; Bodini et al., 2020).

Given that WFPs are impacted by multiple uncertainties and these uncertainties have significant implications for power predictions, recent efforts have focused on WFP intercomparison and validation. Winds from WFP simulations have been validated against meteorological tower observations, aircraft observations, synethic aperture radar (SAR), and lidar measurements (Fischereit et al., 2022a). However, WFP intercomparison and validation efforts have experienced challenges. Observations are generally staged at a distance from the wind farms that allows for validation of only the background meteorology as opposed to the wake behavior. Conclusions drawn from these validation studies may also be influenced by site-specific or meteorological conditions. Further, many previous WRF intercomparison studies contained a TKE advection bug in the Fitch et al. (2012) scheme, as identified in Archer et al. (2020).

A recent North Sea measurement campaign has stimulated interest in WFP intercomparison and validation efforts. The Wind Park Far Field (WIPAFF) project was an aircraft expedition to understand offshore wind wake behavior in the German Bight. This expedition took place in a location with multiple wind farms, with 41 total flights between 6 September 2016 and 15 October 2017 (Platis et al., 2018; Bärfuss et al., 2021). Siedersleben et al. (2018b) modeled a case study that included one day of these aircraft observations and found that, because of challenges in characterizing air-sea interactions with a mesoscale model, improving background flow characterization contributed greater model improvement than any wind farm-specific parameter configuration. Siedersleben et al. (2018a) then extended the analysis of this one case study to demonstrate the presence of hub-

height potential temperature and water vapor wakes during stably-stratified conditions. Siedersleben et al. (2020) then leveraged 3 days of these aircraft observations that occurred during stable conditions to explore the sensitivity of grid cell spacing and TKE advection within the Fitch parameterization. Siedersleben et al. (2020) found that refined horizontal and vertical grid cell spacing improved model agreement and that the Fitch et al. (2012) scheme improved performance with a non-zero wind farm TKE factor. Siedersleben et al. (2020) also found (with the bug-advected Fitch et al. (2012) scheme) that model performance improved when the advection option was turned off. Larsén and Fischereit (2021) extended the work of Siedersleben et al. (2020) by comparing the explicit wake parameterization (EWP) (Volker et al., 2015) and Fitch (Fitch et al., 2012) schemes and exploring model performance during wind farm interactions with low-level jets (LLJs). Larsén and Fischereit (2021) found that the EWP significantly underestimated TKE compared to the Fitch scheme and that the wind farm TKE factor within the Fitch et al. (2012) scheme was the most sensitive modeling parameter. Larsén and Fischereit (2021) also noted that introducing wind farms into the domain altered the LLJ profile. Larsén et al. (2024) then extended the work in Larsén and Fischereit (2021) by introducing ocean and wave coupling. Larsén et al. (2024) found that wind farm wakes reduced surface wind speeds and wave heights, except for when the wind farm TKE factor was 1. When the wind farm TKE factor was 1, excessive TKE generation instead led to increased surface wind speeds and wave heights. Ali et al. (2023) also used data from one day of the Siedersleben et al. (2020) case study to validate five (Fitch et al., 2012; Volker et al., 2015; Abkar and Porté-Agel, 2015; Pan and Archer, 2018; Redfern et al., 2019) common WFPs with aircraft measurements as well as nearby meteorological tower and synthetic aperture radar observations. Ali et al. (2023) found that that the EWP (Volker et al., 2015) underestimated TKE as compared to the Fitch et al. (2012) scheme and that while the Redfern et al. (2019) parameterization did not noticeably deviate from the Fitch et al. (2012) scheme, the Pan and Archer (2018) parameterization showed a significant drop in power generation and the Abkar and Porté-Agel (2015) parameterization predicted lower levels of TKE.

The influence of the PBL scheme choice is one parameter that has not yet been considered for this case study and has generally been absent in the literature evaluating WFPs (Fischereit et al., 2022a). Although the influence of the PBL scheme on the wind resource has been an active field of research in turbine-free NWP simulations (Zhang and Zheng, 2004; Jankov et al., 2005; Li and Pu, 2008; Nolan et al., 2009; Shin and Hong, 2011; Draxl et al., 2014), research on the impacts of the PBL scheme on turbine simulations has been limited because the default Fitch WFP has, until recently, been integrated with only the MYNN PBL scheme (Nakanishi and Niino, 2009).

However, the recent development (Kosović et al., 2020; Juliano et al., 2022; Eghdami et al., 2022) and evaluation (Arthur et al., 2022; Peña et al., 2023; Arthur et al., 2024) of the U.S. National Science Foundation National Center for Atmospheric Research (NCAR) three-dimensional PBL (3DPBL) scheme, followed by its integration with the Fitch scheme (Rybchuk et al., 2022), offers an opportunity to better understand the sensitivity of wind farm behavior to PBL scheme choice. The 3DPBL scheme is based on the Mellor and Yamada (1982) scheme and accounts for the 3D effects of turbulence by explicitly calculating the horizontal momentum, heat, and moisture flux divergences. Similar to the MYNN PBL scheme, the 3DPBL scheme is a level 2.5 model so that TKE is a prognostic variable. The 3DPBL scheme reduces errors in potential temperature, wind speed, and TKE relative to the one-dimensional (1D) MYNN scheme when compared to cold-air pool observations in

the Columbia River basin (Arthur et al., 2022). The potential value of the 3DPBL scheme compared to the 1D MYNN scheme may also depend on the grid cell resolution (Peña et al., 2023).

This work compares the MYNN PBL scheme with a 3DPBL scheme by validating against both tower and aircraft observations for a North Sea case study. These results address the research gap regarding the sensitivity of wake behavior to the PBL scheme and offer guidance to the offshore forecasting community. This manuscript is organized as follows. In section 2, we describe the North Sea case study, detail the observational datasets, and outline the WRF simulation setup. In section 3, we present the results from our WRF simulations and compare these results to meteorological tower and aircraft observations associated with the case study. In section 4, we offer potential implications of the differing performance between the PBL schemes for wind resource assessments and wind power forecasting.

2 Methods

2.1 PBL scheme comparison

The analysis in this work compares the MYNN 2.5 (Mellor and Yamada, 1982; Olson et al., 2019) and NCAR 3DPBL (Juliano et al., 2022) schemes. A brief description of the differences between these two schemes is provided here, while a more complete description of the differences between the 3DPBL and MYNN schemes can be found in Kosović et al. (2020), Juliano et al. (2022), and Rybchuk et al. (2022). (Note the 3DPBL scheme described in this work is not to be confused with 3DTKE (Zhang et al., 2018). 3DTKE is a scale-adaptive model that relies on a level 3 closure. More information about 3DTKE can be found in Zhang et al. (2018).)

The MYNN and 3DPBL schemes have a common origin with the Mellor and Yamada (1982) governing equations:

$$\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j} = -\epsilon_{ijk} f_j U_k - \frac{1}{\rho}\frac{\partial P}{\partial x_i} - \frac{\partial < u_j u_i >}{\partial x_j} - g_i \tag{1}$$

$$\frac{\partial \Theta}{\partial t} + U_j \frac{\partial \Theta}{\partial x_j} = -\frac{\partial < u_j \Theta >}{\partial x_j} \tag{2}$$

where the ensemble mean momentum per unit mass is U_i, potential temperature is θ, P describes the mean kinematic pressure, g_i captures the influence of gravity, and f_k is the Coriolis vector.

With a level 2.5 closure, these governing equations become:

$$\frac{\partial q^2}{\partial t} + U_k \frac{\partial q^2}{\partial x_k} = \frac{\partial}{\partial x_k}[l q S_q \frac{\partial q^2}{\partial x_k}] + 2(P_s + P_b - \epsilon) \tag{3}$$

$$< u_i u_j >= \frac{\delta_{ij}}{3} q^2 - 3\frac{l_1}{q}[(< u_k u_i > - C_1 q^2 \delta_{ki})\frac{\partial U_j}{\partial x_k} + (< u_k u_j > - C_1 q^2 \delta_{kj})\frac{\partial U_i}{\partial x_k} - \frac{2}{3}\delta_{ij} < u_k u_n > \frac{\partial U_n}{\partial x_k}]$$
$$- 3\frac{l_1}{q}\beta[g_j < u_i \theta_v > + g_i < u_j \theta_v > - \frac{2}{3}\delta_{ij} g_n < u_n \theta_v >] \tag{4}$$

$$< \theta_v^2 >= -\frac{\Lambda_2}{q} < u_k \theta_v > \frac{\partial \Theta_v}{\partial x_k} \tag{5}$$

$$< u_j \theta_v >= -3\frac{l_2}{q}[< u_j u_k > \frac{\partial \Theta_v}{\partial x_k} + < u_k \theta_v > \frac{\partial U_j}{\partial x_k} + \beta g_j < \theta_v^2 >] \tag{6}$$

such that $q^2 = 2TKE =< u^2 > + < v^2 > + < w^2 >$, where the TKE budget includes shear production $P_s = - < u_i u_j > \frac{\partial U_i}{\partial x_k}$, buoyant production $P_b = -\beta g_i < u_i \theta_v >$, and dissipation $\epsilon = \frac{q^3}{\Lambda_1}$. β, S_q, and C_1 are all constants. Finally, the four length scales–l_1, l_2, Λ_1, and Λ_2–are proportional in such a way that:

$$[l_1 \ \Lambda_1 \ l_2 \ \Lambda_2] = l[A_1 \ B_1 \ A_2 \ B_2] \tag{7}$$

where l represents a master length scale and A_1, A_2, B_1, and B_2 represent another set of constants.

In this level 2.5 closure, TKE is a prognostic variable, advecting TKE generated within the scheme. This TKE treatment for both MYNN and 3DPBL differs from that for several other PBL schemes (i.e. YSU) and allows for both the MYNN and 3DPBL schemes to integrate with the Fitch WFP (described below) to analyze wind farm wake effects on TKE.

From this common origin, the MYNN and 3DPBL schemes differ. First, the two schemes differ in how they represent horizontal turbulent mixing. The MYNN scheme uses Smagorinsky mixing (Smagorinsky, 1963), which captures horizontal diffusion through eddy-viscosity:

$$\tau_{ij} = 2K_{h,m} D_{ij} \tag{8}$$

where D_{ij} is the resolved deformation tensor such that $D_{ij} = \frac{1}{2}(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_j})$ and $K_{h,m}$ is the horizontal eddy viscosity defined as $K_{h,m} = C_s^2 l_h^2 [0.25(D_{11} - D_{22})^2 + \overline{D_{12}^2}^{xy}]^{1/2}$, where C_s is a constant with a typical value of 0.25, l_h represents the horizontal gridcell spacing, and $\overline{D_{12}^2}^{xy}$ is the average horizontal shear deformation for a grid cell. In contrast, the 3DPBL scheme explicitly calculates the horizontal turbulent flux divergences. These flux divergences are calculated at each model grid cell by defining and solving the full system of algebraic equations that includes the six turbulent stresses, three turbulent sensible heat fluxes, and the potential temperature variance.

A "boundary layer" approximation of the 3DPBL scheme also exists. This approximation retains the same 3DPBL length scale formulation (and corresponding closure constants) but instead sets the horizontal gradients of the mean quantities and the vertical gradient of the vertical velocity to zero. This 3DPBL approximation, which has also been used in previous comparisons with the 3DPBL scheme (Rybchuk et al., 2022; Arthur et al., 2022; Sanchez Gomez et al., 2023; Peña et al., 2023; Arthur et al., 2024) is employed for this analysis.

The two PBL schemes also rely on differing master length scales and corresponding closure constants. For this analysis, the 3DPBL scheme uses a Blackadar (1962) master length formulation as follows:

$$l = \frac{kz}{1 + kz/l_0} \tag{9}$$

such that k is the von Kármán constant, z is the height above ground level, and l_0 is defined by:

$$l_0 = \alpha \frac{(\int_0^h qz\,dz)}{(\int_0^h q\,dz)} \tag{10}$$

with α as an empirical constant and the length scale constants determined experimentally in neutrally-stratified conditions. These constants are those used in Mellor and Yamada (1982).

In contrast, the version of the MYNN scheme used in this analysis does not use closure constants from Mellor and Yamada (1982). Instead, the version of MYNN used in this analysis is best characterized by the default (i.e. bl_pbl_opt = 5, bl_mynn_mixlength=2) configuration of the MYNN-EDMF, as described in Olson et al. (2019). As a result, the closure constants used by the MYNN scheme in this analysis reflect the update in Nakanishi and Niino (2009). The MYNN scheme for this analysis also then calculates its master length scale based on a blending of PBL depth, buoyancy, and surface layer length scales (Olson et al., 2019).

The two PBL schemes are also located within different sections of the WRF codebase. While the MYNN scheme is located within the physics (/phys/) directory, the 3DPBL scheme resides in the dynamics (/dyn_em/) directory. As a result, the 3DPBL scheme introduces an additional subroutine that ensures both that velocity and TKE tendencies in the 3DPBL are scaled in the same manner as those for the MYNN scheme and that the Fitch scheme (described in Sect. 2.3) appropriately modifies the 3DPBL TKE tendency. Further details on the implementation of the 3DPBL Fitch integration can be found in Rybchuk et al. (2022).

2.2 North Sea case study

The WIPAFF aircraft expedition explored the impact of several North Sea offshore wind farms (Table 1) on the atmosphere (Platis et al., 2018; Bärfuss et al., 2021) (Fig. 1). The expedition included 41 flights spanning September 2016 to October 2017, where a subset of both transect and profile flights during stably stratified conditions on 14 October 2017 has been identified as one common research case study (Siedersleben et al., 2018a, b, 2020; Larsén and Fischereit, 2021; Ali et al., 2023; Larsén et al., 2024).

Aircraft measurements from the six transect flights (Table 2) for this 14 October 2017 case study (Platis et al., 2018; Bärfuss et al., 2021) were collected in the vicinity of the Gode wind farm at an altitude of roughly 250 m (Fig. 1b and Fig. 2b). The flight paths across the six transects were roughly symmetrical, with transects 1, 3, and 5 traveling towards the northwest and transects 2, 4, and 6 traveling towards the southeast (Platis et al., 2018; Bärfuss et al., 2021). While observations of wind speed, wind direction, temperature, pressure, and relative humidity were collected (Corsmeier et al., 2001), here, we focus on 100 Hz wind speed observations and derived TKE calculations for our evaluation (Table 4).

The paths for the six profile flights–described more fully in Platis et al. (2018), Bärfuss et al. (2021), and Larsén and Fischereit (2021)–were taken at the same frequency as that of the transect flights according to the duration outlined in Table 3 and paths described in Fig. 1b. Because the transect flights occurred at roughly the same altitude and therefore cannot provide profile assessments of atmospheric stability, we employ the temperature, potential temperature, wind speed, wind direction,

and TKE observations from these profile flights to assess the atmospheric stability of the region for comparison to the model simulations.

Table 1. Selected wind farm characteristics. Wind turbine performance curves are as in Ali et al. (2023). The wind farms in bold are present in immediate environs of the FINO1 and aircraft measurement regions.

Wind farm	Hub height (m)	Diameter (m)	Turbine rating	Capacity (MW)	Number of turbines	Rated wind speed (m s^{-1})
Alpha Ventus ["A.V."]	**90**	**116**	**M5000-116**	**60**	**12**	**12.5**
Nordsee One	**90**	**126**	**6.2M126**	**332**	**54**	**14**
Gode	**110**	**154**	**SWT-6.0-154**	**582**	**97**	**12**
Bard	90	116	M5000-116	400	80	12.5
Global Tech	90	116	M5000-116	400	80	12.5
Borkum Riffgrund	**90**	**120**	**SWT-4.0-120**	**312**	**78**	**16**
Meerwind	88	120	SWT-3.6-120	288	80	14
Amrumbank West ["A.W."]	88	120	SWT-3.6-120	288	80	14
Veja Mate ["V.M."]	106	154	SWT-6.0-154	402	67	12
Gemini	95	130	SWT-4.0-130	600	150	14
Riffgat	88	120	SWT-3.6-120	108	30	14
Nordsee Ost	95	126	6.2M126	295.2	48	14

Table 2. Transect timings for 14 October 2017 WIPAFF case

Transect Number	Start Time	End Time	WRF Comparison Timestep	Start Latitude/Longitude	End Latitude/Longitude
1	14:20:50.860	14:30:12.370	14:30	(53.90, 7.06)	(54.25, 6.96)
2	14:34:41.180	14:44:37.520	14:40	(54.25, 6.95)	(53.90, 7.06)
3	14:48:27.970	14:57:43.640	14:50	(53.90, 7.07)	(54.25, 6.96)
4	15:01:38.120	15:11:34.970	15:00	(54.25, 6.96)	(53.90, 7.06)
5	15:45:01.130	15:54:05.160	15:50	(53.90, 7.06)	(54.25, 6.97)
6	15:58:29.630	16:08:34.810	16:00	(54.25, 6.95)	(53.90, 7.06)

Table 3. Profile timings for 14 October 2017 WIPAFF case used for stability characterization

Profile Number	Start Time	End Time	WRF Comparison Timestep	WRF Comparison Cell (West East, South North)	Map Color
1	13:22:59.700	13:25:08.410	13:20	(193, 296)	Orange
2	14:14:41.600	14:17:10.470	14:10	(180, 280)	Pink
3	15:13:31.260	15:15:53.240	15:10	(175, 278)	Turquoise
4	16:10:21.570	16:12:49.110	16:10	(181, 282)	Grey
5	16:16:35.240	16:19:55.230	16:10	(188, 297)	Purple
6	16:23:22.060	16:25:05.120	16:20	(196, 294)	Blue

This 14 October 2017 North Sea case study also includes observations from a meteorological tower (Fig. 1b and Fig. 2a). The FINO1 tower is located immediately west of the Alpha Ventus wind farm and provides observations of wind speed, wind direction, pressure, and temperature averaged over 10 minutes. This analysis focused on wind speed and included all the available heights: 34 m, 41 m, 51 m, 61 m, 81 m, 91 m, and 102 m. TKE calculations at the FINO1 site were not available due to coarse temporal resolution of the wind observations at the FINO1 site (Table 4).

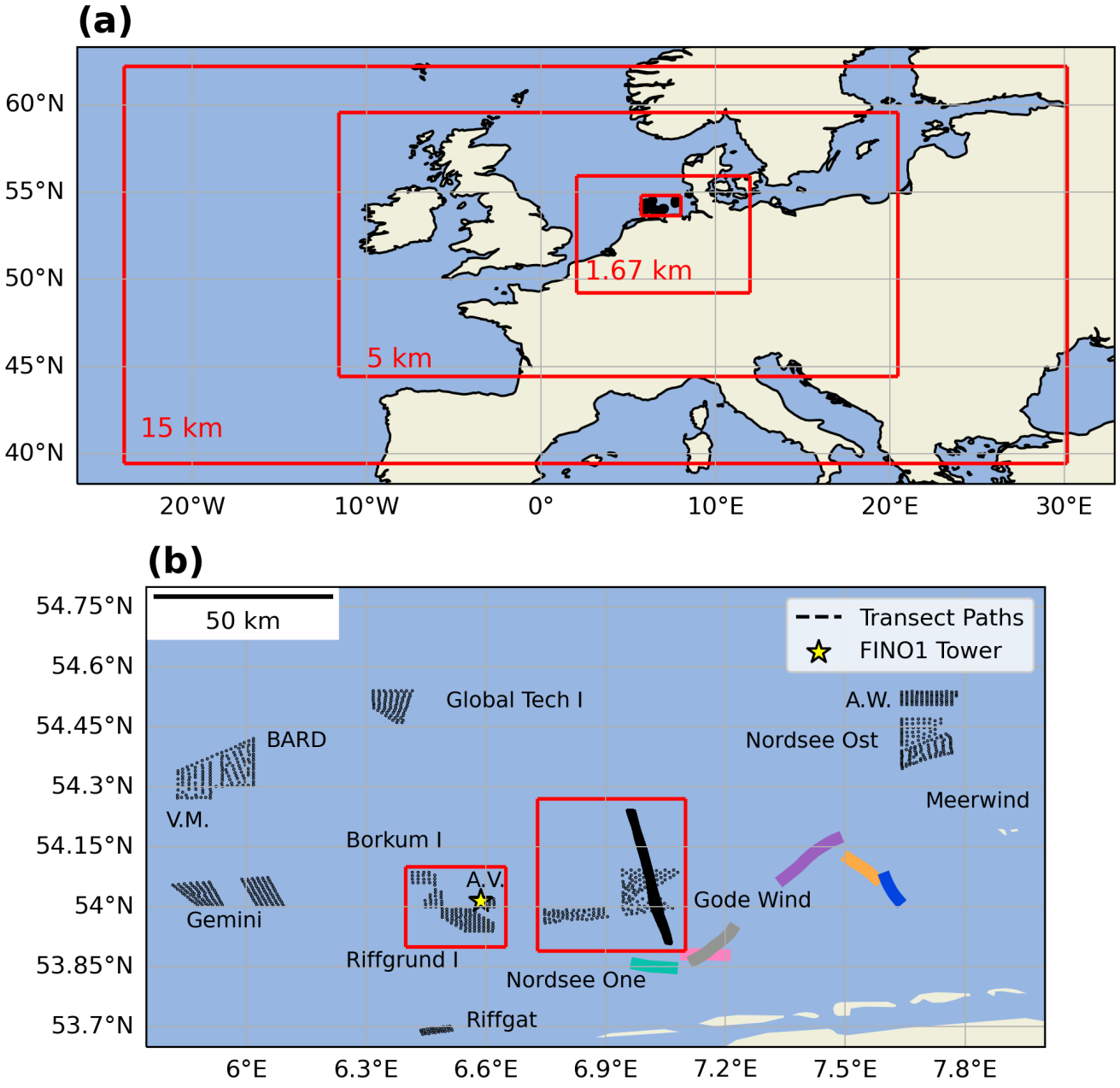

Figure 1. WRF simulation domain for the North Sea 14 October 2017 case study. (a) Three nested domains and a measurement region within the inner domain are outlined in red. (b) Measurement region with wind farms outlined in black, the FINO1 tower marked with a star, the aircraft transect paths traced in black, and the other, colored lines indicate the aircraft profile paths outlined in Table 3.

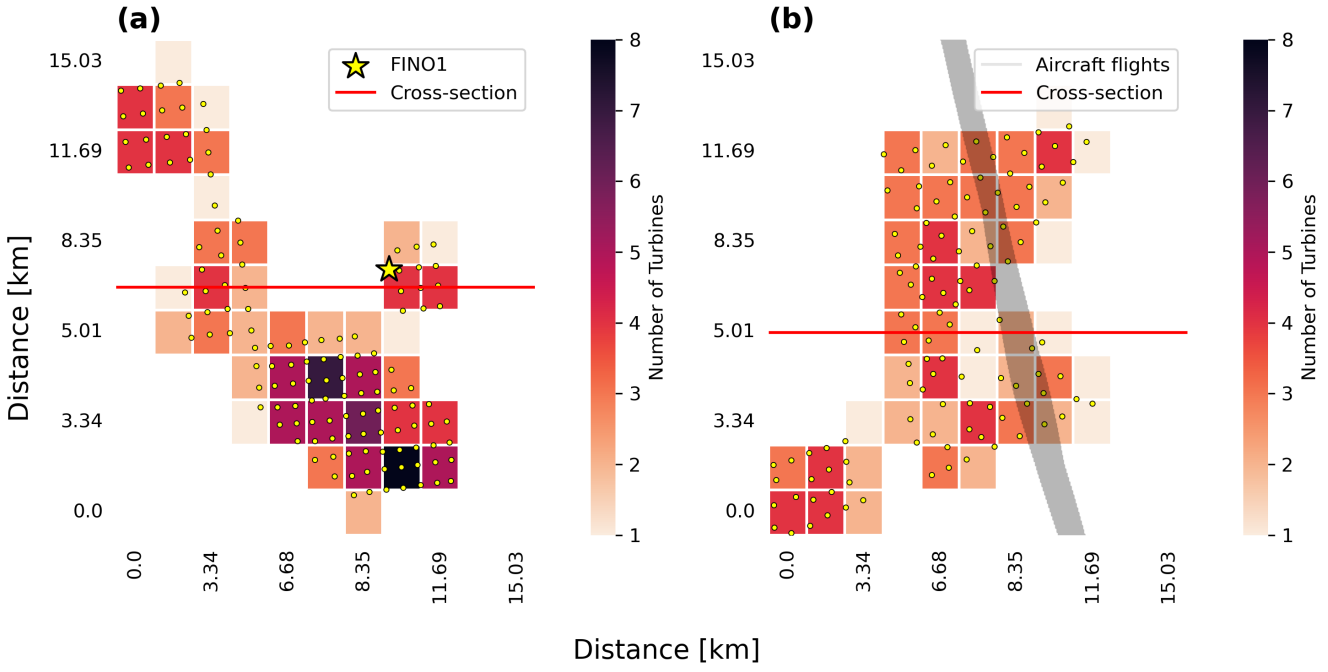

Figure 2. Map of the number of turbines per WRF grid cell in the innermost domain of the 1.67 km resolution for the two regions of interest within the inner region. The two axes represent the WRF grid system. (a) FINO1 site, with the tower marked with a star. (b) Aircraft site, with the six transect paths traced in black.

Table 4. In situ observations for validation

Site	Variable	Altitude	Temporal resolution
Aircraft	Wind Speed	250 m	100 Hz
Aircraft	TKE	250 m	100 Hz
FINO1	Wind Speed	34, 41, 51, 61, 81, 91, 102 m	10 minutes

Table 5. Baseline WRF modeling configuration as an extension of Ali et al. (2023).

Parameter	Value
Version	4.4.2
Date	14 October 2017
Timestep [s]	30
Output resolution [min]	10
dx/dy [km]	15, 5, 1.67
Vertical levels	80
Land-Sea Model	Noah
Radiation Scheme	RRTMG
Boundary Conditions	ERA5
Cloud parameterization	Kain-Fritsch

2.3 Model setup

Simulations were performed with the WRF model (Skamarock et al., 2021) with the Fitch WFP (Fitch et al., 2012), modified to incorporate the 3DPBL scheme. The WRF simulations here generally followed the setup of Ali et al. (2023) (Table 5). Simulations represented the single day of 14 October 2017 with a 30 s timestep in the outer domain and a 10 min output, starting at 00:00:00 UTC with a 12 hour spin-up period so that the analysis period starts at 12:00:00 UTC. The region was simulated using three nested domains with an outer horizontal grid size of 15 km and a nesting ratio of 3 so that the innermost domain has a grid size of 1.67 km. Eighty vertical levels were employed to ensure sufficient vertical resolution at and below rotor height per Tomaszewski and Lundquist (2020). Specifically, 17 levels were lower than 200 m, and, depending on the turbine's diameter and hub height, between 8 and 12 levels intersected the turbine's rotor. Our model configuration also replicated many of the boundary conditions and physics options from Ali et al. (2023). The initial and boundary conditions for both analyses were provided by ERA5 reanalysis (Hersbach et al., 2020). We also followed Ali et al. (2023) by using the WRF double-moment six-class microphysics scheme (Hong et al., 2010), the RRTMG shortwave and longwave radiation scheme (Mlawer et al., 1997), and the Noah land-surface model (Niu et al., 2011) and by including the Kain–Fritsch cumulus parameterization scheme (Kain, 2004) in the outer domain only.

Differences also exist between the WRF setup presented in this work and that used in Ali et al. (2023). Here, we varied the PBL scheme to explicitly consider the influence of the PBL scheme on wake behavior for the Fitch WFP. Two PBL schemes were considered: level 2.5 MYNN ("MYNN") and the NCAR 3DPBL scheme with the PBL approximation ("3DPBL") as described in Rybchuk et al. (2022). The MYNN scheme is activated in all outer domains for all simulations. Further, Ali et al. (2023) performed their analysis on a modification of WRF v4.5.1, whereas the analysis presented in this work relied on an earlier version of WRF (V4.4.2) in which the 3DPBL scheme is integrated.

Wind farm effects were represented with the Fitch WFP. Accordingly, the drag force for a given turbine was:

$$\mathbf{F_{drag}} = \frac{1}{2} C_T \rho |\mathbf{V}| \mathbf{V} A \tag{11}$$

where $\mathbf{V}$ is the horizontal wind velocity, C_T is the turbine thrust coefficient, which varies with wind speed, ρ is the air density, and A is the cross-sectional rotor area.

Further, the fraction of mean kinetic energy converted into TKE was governed by the turbine's thrust coefficient, C_T, the turbine's power coefficient, C_P, and a wind farm TKE factor, α, by:

$$C_{TKE} = \alpha (C_T - C_P) \tag{12}$$

and the turbine-induced TKE tendency was:

$$\frac{\partial TKE_{ijk}}{\partial t} = \frac{0.5 N_{ij} C_{TKE} |\mathbf{V_{ijk}}|^3 A_{ijk}}{z_{k+1} - z_k} \tag{13}$$

where i, j, and k are the zonal, meridional, and vertical grid cell indices, respectively; N_{ij} is the turbine number density for a given cell [m^{-2}]; $|\mathbf{V}|_{\mathbf{ijk}}$ are the wind speed components [m s^{-1}]; A is the turbine rotor area [m^2]; C_{TKE} is the TKE coefficient [ ]; and z is the model level height [m].

Our simulations varied the wind farm option, wind farm TKE factor, and the advection option for each PBL scheme (Table 6). We highlight the results from the two simulations with the Fitch scheme, the advection option on, and the wind farm TKE factor of 0.25 to focus the scope on the effects of the PBL scheme (Table 6). The results from the other fourteen runs are analyzed throughout the Appendix.

We first compared "no wind farm" (NWF) simulations to Fitch simulations. These simulations were performed to distinguish turbine effects from the underlying meteorology. Results from these runs are presented in Sect. A1.

We then varied the wind farm TKE factor for each of the Fitch simulations. Recent scientific discussion (Mangara et al., 2019; Archer et al., 2020; Siedersleben et al., 2020; Sanchez Gomez et al., 2023; Ali et al., 2023; García-Santiago et al., 2024; Optis, 2024) has focused on the determination of the optimal value of the wind farm TKE factor, C_{TKE}, with varying conclusions. Several works corroborate improved performance with increased C_{TKE}. Vanderwende et al. (2016) found that Fitch simulations with a wind farm TKE factor of 0 underpredicted TKE. García-Santiago et al. (2024) similarly showed that the Fitch scheme could improve TKE underpredictions. Siedersleben et al. (2020), Ali et al. (2023), and Sanchez Gomez et al. (2023) also identified improved predictions with larger C_{TKE}. However, the overestimation of C_{TKE} is also evident in the literature. Vanderwende et al. (2016) likewise found that a wind farm TKE factor of 1 contributed to TKE overpredictions. Similarly, García-Santiago et al. (2024) found that the Fitch scheme with a wind farm TKE factor of 1 overpredicted TKE. Mangara et al. (2019) and Larsén et al. (2024) likewise identified excess TKE predictions with a wind farm TKE factor of 1. One challenge in comparing our results with those in Vanderwende et al. (2016), Mangara et al. (2019), and Siedersleben et al. (2020) is that these works were completed prior to the advection bug identification. However, the results presented in Siedersleben et al. (2020) are also corroborated in the post-bug-identification work of Larsén and Fischereit (2021). We considered TKE factors of 0, 0.25, and 1. We focused on these three wind farm TKE factors both to cover the full range of

variability and to consider the 0.25 factor suggested by Archer et al. (2020). We present the results with a 0.25 wind farm TKE factor here in the main text and discuss the results for the other wind farm TKE factors in Sect. A2.

Finally, we varied the advection option. The appropriate treatment of the TKE advection option for simulations of wind farm impacts – and specifically this case study – has also received much discussion in the literature. Most recently, Ali et al. (2023) argue that the advection option should be turned off. Ali et al. (2023) make this judgment based on the Siedersleben et al. (2020) finding that performance in the aircraft region improved by turning the advection option off. However, Siedersleben et al. (2020) performed their analysis with a version of the Fitch scheme that included the advection bug. As such, the results presented in Siedersleben et al. (2020) are qualitatively different from those presented in Larsén and Fischereit (2021). Although Larsén and Fischereit (2021) performed their analysis after the advection bug was addressed and performed simulations with advection both on and off, Larsén and Fischereit (2021) argue that further analysis would be necessary to make a formal recommendation.

We recommend using TKE advection for this case study, in contrast to prior recommendations. We argue that the advection option should be included on two bases. First, we note that TKE advection was omitted only from early versions of MYNN for reasons of numerical stability, not on a physical basis (Olson et al., 2019). Second, we note the guidance from Wadler et al. (2023), who show that introducing TKE advection allows for a more realistic distribution of TKE and argue that TKE advection should be modeled in a closure scheme unless there is an explicit reason to exclude this process. In addition to this physical argument, we quantify the (statistically-insignificant) performance differences depending on the advection option in Sect. A3.

The wind-speed- and turbine-model-dependent thrust and power coefficients were integrated into the WRF model through turbine specification files (Fig. 3). Individual turbines were also integrated into the WRF grid with a file from Ali et al. (2023) that contains a given turbine's latitude, longitude, and turbine type. We used files from the Ali et al. (2023) repository and extracted the key information to fit the standard Fitch WFP format.

2.4 Model validation

2.4.1 Other diagnostic variables

We now describe other calculated quantities used to understand site performance and physical mechanisms.

The wind speed deficit, λ, characterizes wake strength:

$$\lambda_{ijk} = U_{NWF_{ijk}} - U_{Fitch_{ijk}} \tag{14}$$

where $U_{NWF_{ijk}}$ is the horizontal wind speed at a specific grid cell ij and specific height index k with no wind farms and $U_{Fitch_{ijk}}$ is the horizontal wind speed with the corresponding wind farm simulation.

The "drag proxy", D, is directly proportional to the turbine drag force (Eq. (11)) for a constant density and rotor area:

$$D = C_T V^2 \tag{15}$$

where C_T and V are the corresponding manufacturer-specified proper thrust coefficients and wind speeds, respectively (Fig. 3a, Fig. 3b).

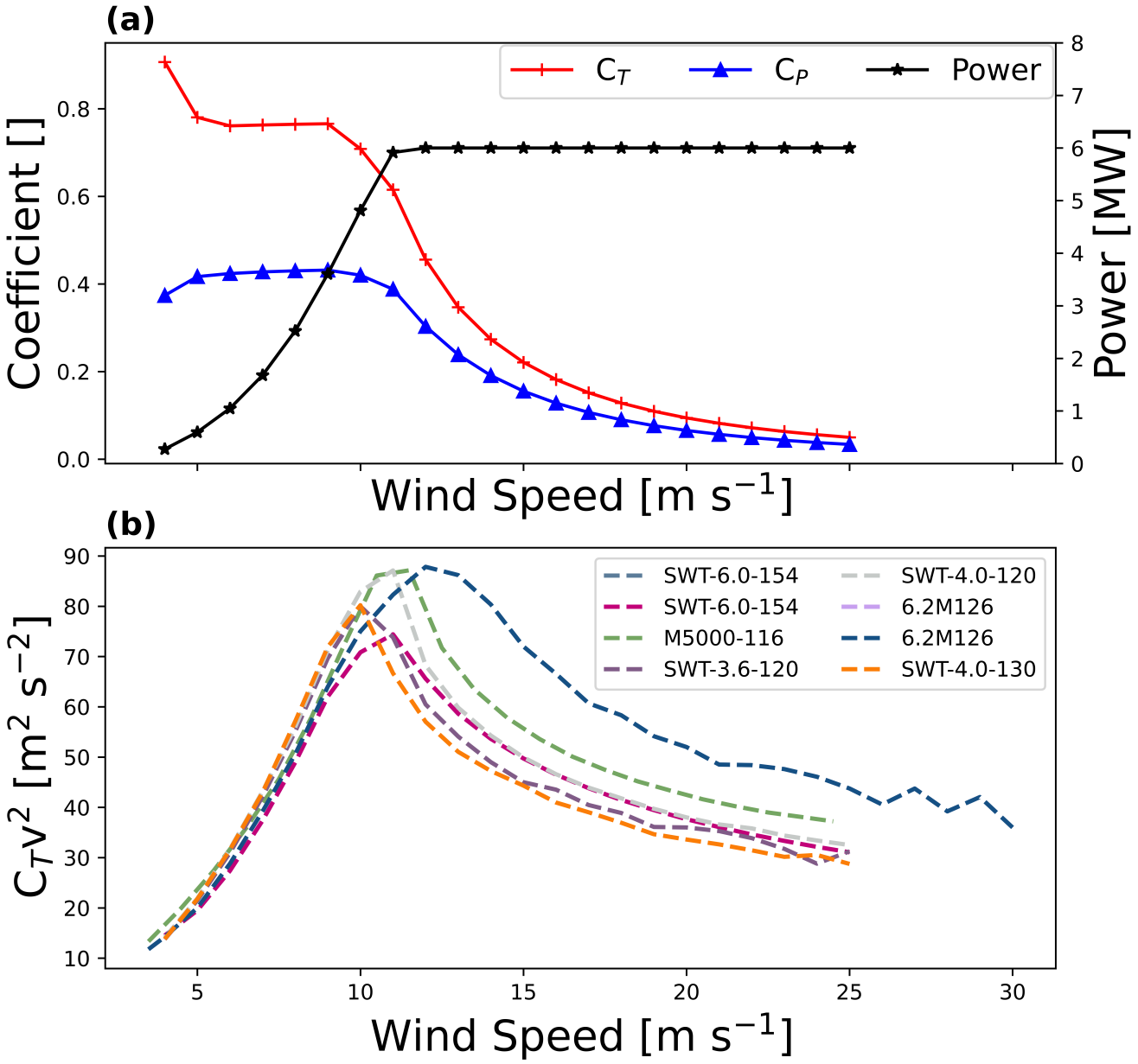

Figure 3. (a) Curve illustrating turbine C_T, C_P, and power specifications for the turbine model in the Gode wind farm. (b) Drag proxy for each of the eight turbine models present in this case study.

Table 6. Full set of WRF simulations and sensitivities. The simulations in bold are those formally evaluated for performance, and the simulations not in bold are sensitivity runs explored in the Appendix.

Simulation Name	PBL Scheme	WFP	TKE Advection	TKE Factor	Short Name
MYNN NWF Noadvect NA	MYNN	N/A	Off	N/A	mnn_NA
3DPBL NWF Noadvect NA	3DPBL	N/A	Off	N/A	3nn_NA
MYNN NWF Advect NA	MYNN	N/A	On	N/A	mna_NA
3DPBL NWF Advect NA	3DPBL	N/A	On	N/A	3na_NA
MYNN Fitch Noadvect 000	MYNN	Fitch	Off	0	mfn_000
3DPBL Fitch Noadvect 000	3DPBL	Fitch	Off	0	3fn_000
MYNN Fitch Advect 000	MYNN	Fitch	On	0	mfa_000
3DPBL Fitch Advect 000	3DPBL	Fitch	On	0	3fa_000
MYNN Fitch Noadvect 025	MYNN	Fitch	Off	0.25	mfn_025
3DPBL Fitch Noadvect 025	3DPBL	Fitch	Off	0.25	3fn_025
MYNN Fitch Advect 025	**MYNN**	**Fitch**	**On**	**0.25**	**mfa_025**
3DPBL Fitch Advect 025	**3DPBL**	**Fitch**	**On**	**0.25**	**3fa_025**
MYNN Fitch Noadvect 100	MYNN	Fitch	Off	1	mfn_100
3DPBL Fitch Noadvect 100	3DPBL	Fitch	Off	1	3fn_100
MYNN Fitch Advect 100	MYNN	Fitch	On	1	mfa_100
3DPBL Fitch Advect 100	3DPBL	Fitch	On	1	3fa_100

ΔTKE represents the time-averaged difference in TKE between the two PBL schemes at a given i, j, and k location.

$$\Delta TKE = TKE_{3DPBL} - TKE_{MYNN} \tag{16}$$

This difference field, consistent with all subsequent difference fields, is defined as 3DPBL - MYNN.

2.4.2 Spatial and temporal processing

Additional spatial and temporal averaging was applied, with care taken to reflect site differences. For the FINO1 region, 10 min WRF output data for the hours of 12:00:00–00:00:00 UTC were averaged to compare with the FINO1 data. For the FINO1 region, the 10 min model output data were subset to the closest ij grid cell and k model level to the 10 min observations. Each observational height was evaluated separately.

Additional processing was necessary to capture the temporal and spatial variability of the aircraft transect paths. Because each transect represented approximately 10 minutes of observations, each transect could be reasonably compared to a single 10 min model output. At the same time, observations within a single transect spanned multiple model grid cells. Given these additional considerations, we processed the aircraft region data with the following process, based on Platis et al. (2018) and

Larsén and Fischereit (2021). First, we calculated the horizontal wind speed as:

$$U = \sqrt{u^2 + v^2} \tag{17}$$

where u and v are the zonal and meridional components, respectively, of the wind in m s^{-1} for a given transect with the 100 Hz observations. Then, we resampled the horizontal wind speeds with a moving 2 km window. This 2 km window was first determined by Platis et al. (2018) and later implemented in Larsén and Fischereit (2021) for this case study. This window was selected based on the aircraft speed to yield an average turbulent timescale on the order of a couple of minutes. This integral timescale appropriately separates the small-scale fluctuations from the large-scale turbulent motions (Platis et al., 2018; Larsén and Fischereit, 2021). Additional averaging was then performed, this time across grid cells. The 2 km resolution wind speed calculations were mapped to a corresponding model grid cell based on their latitude and longitude, and all the 2 km resolution wind speed calculations for a given model grid cell were averaged together. The number of 2 km resolution wind speed calculations for a given model grid cell depended on the amount of time that the aircraft spent in that grid cell. Well-sampled grid cells may contain close to 3000 points, whereas less-sampled grid cells may contain only 10 points. These grid-cell-averaged values could then be compared to the relevant model cell value at the closest timestep (Table 2).

We employed a similar process to calculate TKE, also based on Platis et al. (2018) and Larsén and Fischereit (2021). We again isolated the 100 Hz observations for a given transect and resampled the TKE based on a 2 km moving (standard deviation) window (Platis et al., 2018):

$$TKE = \frac{1}{2}(\sigma_u^2 + \sigma_v^2 + \sigma_w^2) \tag{18}$$

where σ_u, σ_v, and σ_w correspond to the standard deviations of the u, v, and w components, respectively. We then averaged these values across each grid cell to compare to the WRF model (Table 2).

Model data and observations from the aircraft vertical profiles were compared to assess the model's ability to capture important local characteristics like atmospheric stability. Because each profile traversed several model grid cells, observations for a given profile were compared to a designated "middle" model grid cell (Table 3). Potential temperature, θ [K], was calculated–both from the 100 Hz observations and from 10 min model data–according to the equation:

$$\theta = (\frac{p_0}{p})^{0.2854} \tag{19}$$

where p_0 is the reference pressure 10000 hPa and p (hPa) is the pressure surrounding the aircraft at a given moment. Horizontal wind speed and TKE for the aircraft profile observations were calculated using the same Eq. 17 and Eq. 18 used with the aircraft transect flights. However, the rolling averaging window for the aircraft profile flights was 10 m (as opposed to 2 km), again as in Larsén and Fischereit (2021).

These wind speed (and corresponding TKE) retrievals were subject to errors in aircraft yaw alignment. Bärfuss et al. (2023) analyzed experimental influences on retrievals from a drone platform and concluded yaw misalignment to be the dominant source of error for wind speed (and TKE) retrievals. Bärfuss et al. (2023) also noted that errors in yaw alignment have opposite signs depending on whether the sensor interacts with the wind from the starboard or backboard side. Because odd transects

315 expose the sensor to one side of the plane and even transects expose the sensor to the other side, this error in yaw alignment could be reflected in systematic differences in wind speed (and TKE) between even transects and odd transects.

2.4.3 Error metrics

The standard (Optis et al., 2020) error metrics of bias, centered root mean square error ($cRMSE$), correlation squared (R^2), and earth movers' distance (EMD) were calculated for each available variable at each site. The bias, $cRMSE$, and EMD 320 all have an optimal value of 0, whereas R^2 has an optimal value of 1. For the FINO1 region, the averages represented time averages, and for the aircraft region, the averages were across grid cells. FINO1 time averages included all 10 min data points for the 12:00:00–00:00:00 UTC period. The bias represents the difference between the modeled and observed means:

$$bias = (\overline{p} - \overline{o}) \tag{20}$$

where $\overline{p}$ represents the modeled mean and $\overline{o}$ represents the observed mean. The $cRMSE$ represents the unbiased component 325 of the model error. The $cRMSE$ in this case is:

$$cRMSE = [\frac{1}{N}\sum_{n=1}^{N}[(p_n - \overline{p}) - (o_n - \overline{o})]^2]^{1/2} \tag{21}$$

where N is the number of data points. The Pearson correlation coefficient (R) represents the correspondence between two variables:

$$R = \frac{\frac{1}{N}\sum_{n=1}^{N}[(p_n - \overline{p}) - (o_n - \overline{o})]}{\sigma_p \sigma_o} \tag{22}$$

330 where σ_p and σ_o represent the standard deviations of the predictions and observations, respectively. Here, we reported the coefficient of determination, R^2, as recommended in Optis et al. (2020). Finally, EMD, also known as the Wasserstein distance, represents the area between two cumulative distribution functions. The EMD can be described by:

$$EMD(u,v) = \inf_{\pi \in \Gamma(u,v)} \int_{\mathbb{R}\times\mathbb{R}} |x - y| \mathrm{d}\pi(x,y) \tag{23}$$

where $\Gamma(u,v)$ is the set of probability distributions on $\mathbb{R} \times \mathbb{R}$ whose marginals are u and v on the first and second factors, such 335 that $u(x)$ is the probability of u at position x and $v(x)$ is the probability of v at position x (Ramdas et al., 2015). Here, the EMD was calculated with the Python function wasserstein_distance() from the SciPy library (Virtanen et al., 2020).

2.4.4 Statistical significance testing

Statistical significance testing was also performed to determine the strength of the differences between simulations. This testing was performed for each error metric separately, with each transect/height representing a value of the appropriate sample. 340 Noting the small number of data points as well as the non-normality of each sample, we performed our statistical testing with a Mann–Whitney U test (Mann and Whitney, 1947). The Mann–Whitney U test is a non-parametric, rank-sum test that relaxes traditional normality and variance requirements with many other statistical tests. Tests were performed with the corresponding Python function from the SciPy stats module (Virtanen et al., 2020), and a result is deemed statistically significant if $|p| < 0.05$.

3 Results

We outline the results of the 3DPBL and MYNN evaluation below. We begin with an overall site characterization, then move to a consideration of systematic influences of model influence, and finish with a statistical evaluation of model performance at both sites.

3.1 Site characterization

3.1.1 Atmospheric stability

The modeled atmospheric stability at both the FINO1 location (Fig. 4a,b) and over the aircraft region (Fig. 4f,g) suggests a weakly stable profile near the surface with stronger stability aloft. MYNN simulations are slightly warmer (according to potential temperature) and more stable near the surface than 3DPBL simulations (Fig. 4b,g). For both sites, these modeled temperature profiles are consistent with available observations (Fig. 4a,f) and the observed potential temperature in the aircraft region is slightly warmer than the modeled potential temperature (Fig. 4g). Because the modeled air temperatures are almost identical between models (Fig. 4a,f), these slight differences in the surface stability could be a consequence of the greater TKE with the 3DPBL scheme that encourages slightly more mixing. The wind direction profiles at both FINO1 (Fig. 4c) and the aircraft transect (Fig. 4h) regions show veering wind (i.e., the wind direction rotates from southwesterly direction near the surface to westerly aloft), suggesting warm-air advection. Simulations with MYNN also tend to have slightly more southerly winds than simulations with the 3DPBL scheme (Fig. 4c,h). Wind speeds are similar near the surface and through the rotor layer with slight variations between the aircraft and FINO1 regions (Fig. 4d,i). Peak wind speeds for both sites are slightly higher than the peak TKE, reflecting a LLJ, as in Larsén and Fischereit (2021). This TKE peak is more accentuated with the 3DPBL scheme for both sites (Fig. 4e,j) and during the odd transect flights (Fig. 4j). The wind direction vertical profiles for both the FINO1 (Fig. 4c) and aircraft (Fig. 4h) regions also suggest inversions at 500 m. Although this separate air mass would not be considered an inversion according to the potential temperature vertical profiles (Fig. 4a,d), the variation in wind speed does support a distinct layer above 500 m (Fig. 4d,i). Further, the modeled PBL height, which in stably stratified conditions is determined as the height at which the TKE profile reaches 5% of its surface value (Olson et al., 2019), corroborates this distinction. Thus, dynamically, the top of the stable boundary layer likely resides around 500 m.

The aircraft profile flights suggest a similar vertical structure to that for the FINO1 and aircraft transect regions. Both MYNN and 3DPBL temperature and potential temperature profiles at the locations of the profiles (Fig. 5a,b) suggest a weakly stable profile near the surface with stronger stability aloft. This stability is reinforced by the aircraft profile observations (Fig. 5f,g). A LLJ again emerges both in wind speed (Fig. 5d,i) and TKE (Fig. 5e,j) maxima. Finally, both modeled (Fig. 5c) and observed (Fig. 5h) wind direction from the aircraft profiles again suggest warm-air advection with an inversion that is supported by the modeled PBL height.

For the aircraft profile flights, the vertical structure suggests slight discrepancies between model and observation both in terms of values and shape (Fig. 5). For example, inversions in temperature (Fig. 5a,f) and potential temperature (Fig. 5b,g) are simulated approximately 100 m higher than observed. The wind direction (Fig. 5c,h) rotation is also modeled at a higher altitude

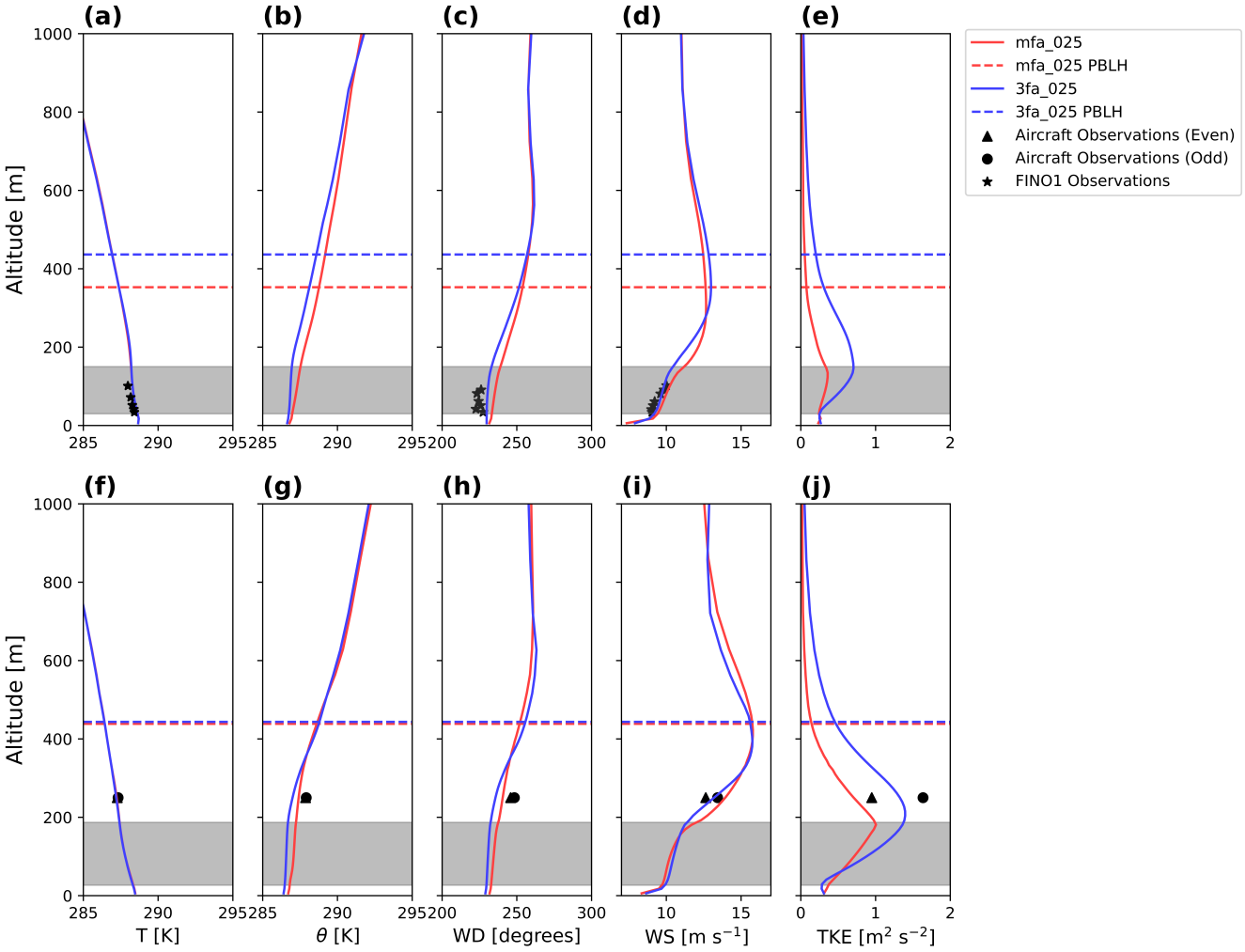

Figure 4. Temperature, potential temperature, wind direction, wind speed, and TKE vertical profiles from observations and WRF simulations for both sites. In all cases, the dashed lines indicate the modeled PBL height and the grey region indicates the turbine rotor region. Observations in the aircraft region are separated between and even and odd transects. (a) FINO1 temperature; (b) FINO1 potential temperature; (c) FINO1 wind direction; (d) FINO1 horizontal wind speed; (e) FINO1 TKE; (f) aircraft temperature; (g) aircraft potential temperature; (h) aircraft wind direction; (i) aircraft horizontal wind speed; (j) aircraft TKE. FINO1 cases are averaged over hours 12:00:00–00:00:00 and the aircraft region cases are averaged over 14:10:00–16:10:00. FINO1 TKE calculations based on observations were not available due to the coarse temporal resolution of the wind speeds. FINO1 potential temperature calculations based on observations were not available due to a lack of pressure observations.

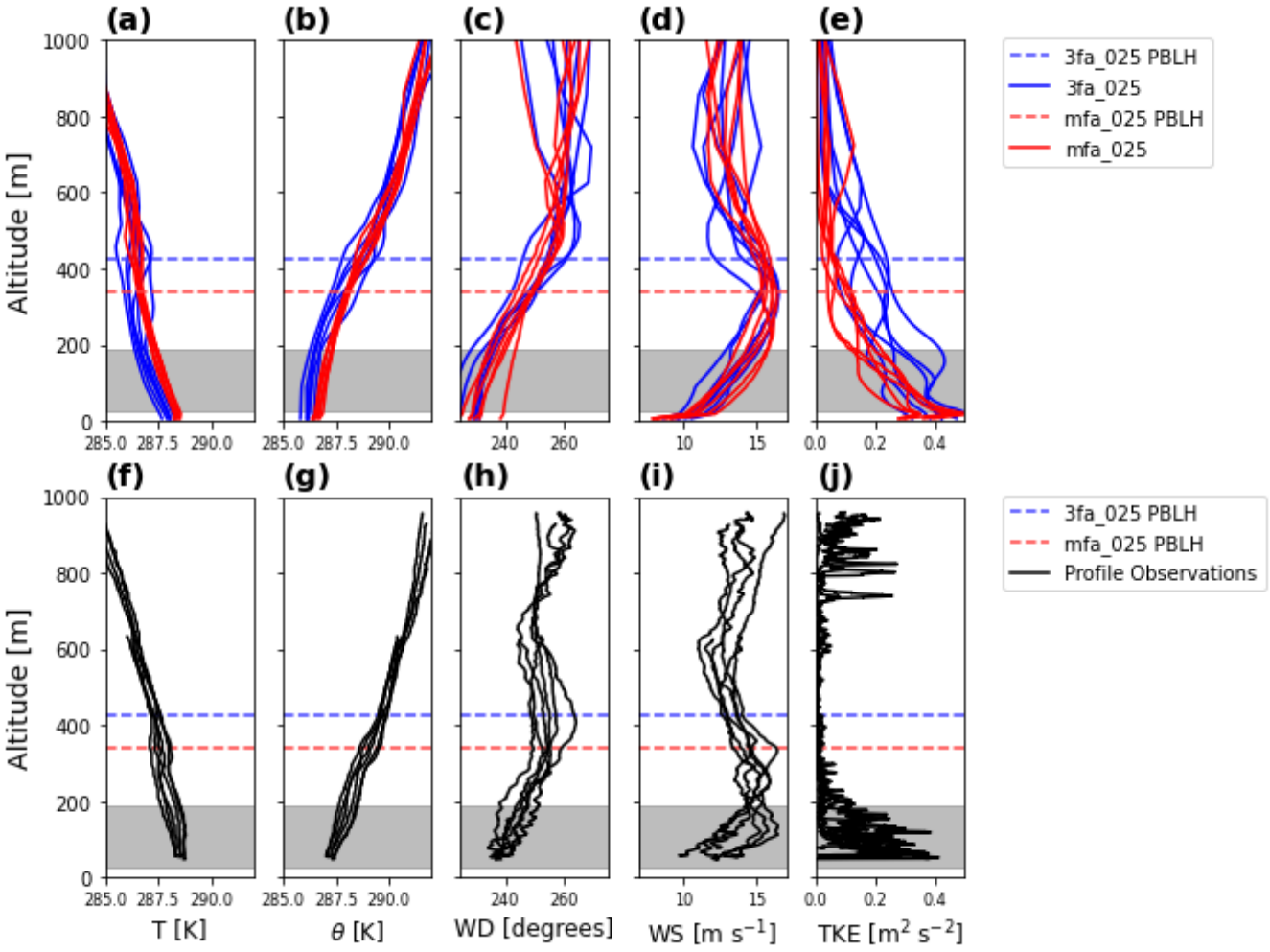

Figure 5. Observed and modeled vertical profiles for the aircraft vertical profile flights (Table 3, Fig. 1b). In all cases, the horizontal line indicates the modeled PBL height, and the color differentiates the PBL scheme. The top row of panels corresponds to modeled output and the bottom row of panels corresponds to the aircraft profile observations. Modeled output are determined to be a given middle cell for each profile as in Larsén and Fischereit (2021) based on the timestep indicated in Table 3. (a) modeled temperature; (b) modeled potential temperature; (c) modeled wind direction; (d) modeled horizontal wind speed; (e) modeled TKE; (f) observed temperature; (g) observed potential temperature; (h) observed wind direction; (i) observed horizontal wind speed; (j) observed TKE.

than they are observed. The observed surface wind direction (Fig. 5h) also reflects an approximately 20 degree discrepancy from the modeled surface wind direction (Fig. 5c). Observed wind speeds (Fig. 5i) are also faster than modeled wind speeds (Fig. 5d) with an additional local peak lower than 200 m that is not captured by the model. Finally, observed TKE (Fig. 5j) suggests two peaks, with one peak within the turbine rotor layer and the other peak well above aircraft transect flights and the top of the boundary layer. In contrast, modeled TKE (Fig. 5c) suggests a single TKE peak with a more gradual decrease with height. These discrepancies in agreement could be explained by several factors, such as the larger number of available data points for the aircraft profile flights, a lack of a smoothing time average for the aircraft profile observations, and the spatial variability sampled with the aircraft profile flights.

Despite the minor disagreements between model and observation for the aircraft profile flights, the stable stratification evident in all locations from both observations and simulations influences the simulations for both the FINO1 and aircraft transect regions. The stable stratification may suppress some of the turbine-generated turbulence from extending aloft to the aircraft transects, at least 100 m above the wind turbines (Fig. 6). Both FINO1 and the aircraft transect measurement regions show simulated TKE peaks at altitudes within the rotor region due to wind-farm-generated turbulence. For the FINO1 region, the two simulated TKE maxima align with the Riffgrund and Alpha Ventus wind farms (Fig. 6a). For the aircraft region, the two TKE peaks align with the Nordsee One (left in Fig. 6b) and Gode (right in Fig. 6b) wind farms. Both the FINO1 and aircraft regions also show larger amounts of TKE with the 3DPBL scheme (Fig. 6a,b) than with the MYNN scheme (Fig. 6c,d). These differences in TKE between the PBL schemes, which are consistent with those shown under the idealized, stable conditions simulated in Rybchuk et al. (2022), reflect the fundamental differences between the models. Notably, while the MYNN scheme uses Smagorinsky mixing to characterize horizontal turbulent mixing, the 3DPBL scheme instead calculates the horizontal turbulent flux divergences explicitly. The two models also rely on different length scales and empirical constants. The stronger TKE maxima in the 3DPBL (Fig. 6a,b) at both sites also lead to greater interfarm TKE overlap with 3DPBL than for MYNN (Fig. 6c,d), such that the TKE interactions between the wind farms are more pronounced for 3DPBL. The difference in both the intensity and degree of overlap between the 3DPBL and MYNN TKE maxima is stronger in the aircraft region (Fig. 6f) than in the FINO1 region (Fig. 6e), likely due to the larger number of wind turbines in the aircraft region (Fig. 6a,c). However, whereas the aircraft region's simulation suggests a higher maximum TKE than the FINO1 region, not all of the aircraft region TKE is captured by the measurements (Fig. 6b,d). The turbine-induced turbulence is generated at the turbine rotor level, which is sampled well by the FINO1 tower (Fig. 6a,c). However, some of this turbine-induced turbulence is suppressed from reaching the aircraft region measurement height (Fig. 6b,d). The two PBL schemes also do not respond to this stable stratification in the same manner. The 3DPBL scheme simulates more of this TKE reaching the aircraft measurement height, resulting in a comparatively larger 3DPBL-MYNN TKE difference at the aircraft measurement height (Fig. 6f).

These diverging patterns in TKE characterization between the two sites also have secondary influences on measured wind speeds (Fig. 7). The 3DPBL scheme enhances TKE peaks present with the MYNN scheme for both the aircraft region (Fig. 6b,d) and the FINO1 region (Fig. 6a,c). However, the differing measurement heights within the stable stratification capture these effects differently. In the aircraft region, the enhanced TKE implies greater momentum extraction and, consequently, reduced wind speeds at the measurement site (Fig. 7b,d,f). The rotor layer in the FINO1 region, however, introduces an additional

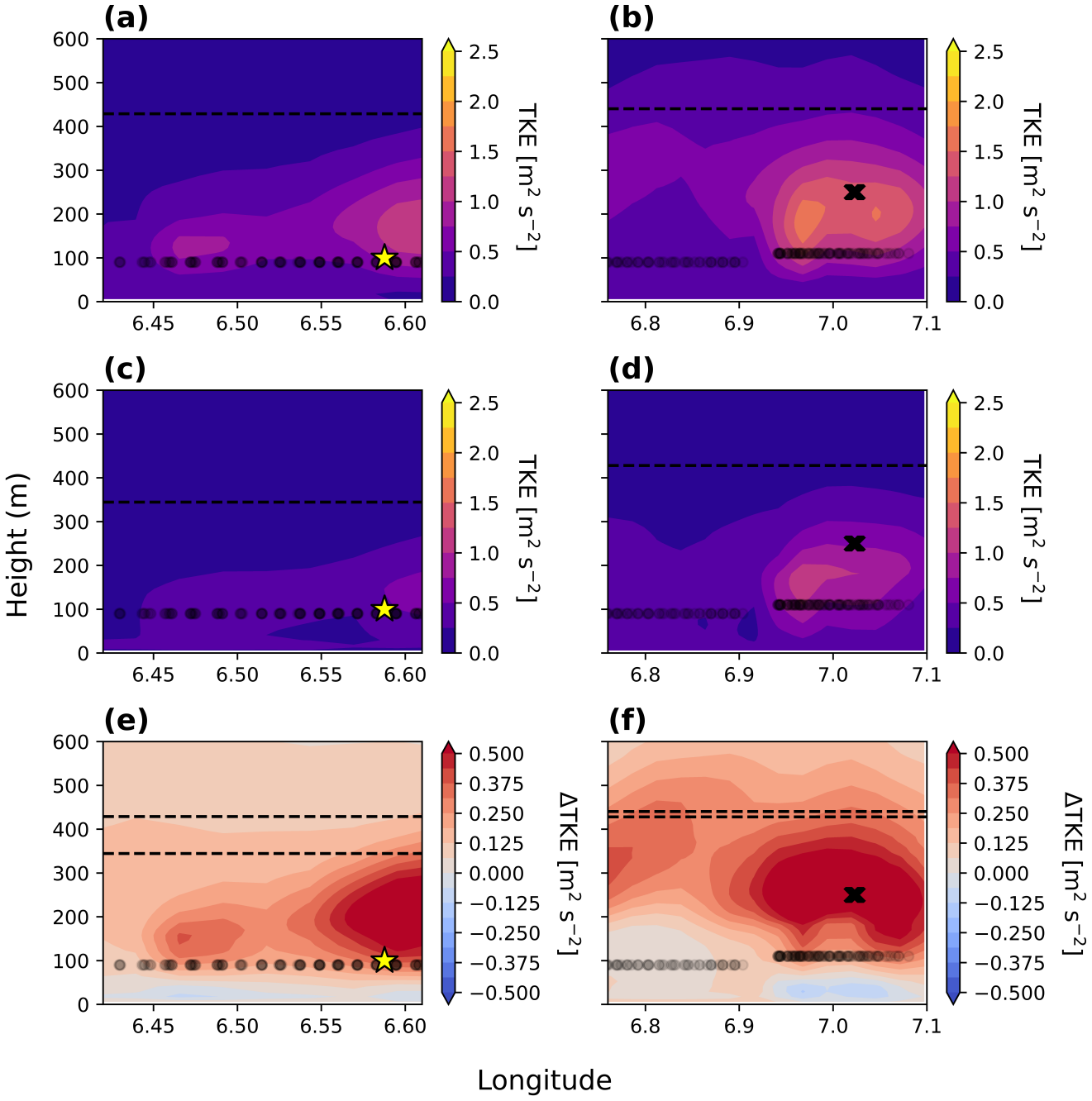

Figure 6. Modeled TKE cross-section at a constant latitude of 54.03. (a) FINO1 3fa_025; (c) FINO1 mfa_025; (e) FINO1 3fa_025 - mfa_025; (b) aircraft 3fa_025; (d) aircraft mfa_025; (f) aircraft 3fa_025 - mfa_025. The horizontal dashed black line denotes the average modeled PBL height, the star indicates the FINO1 tower location, the "X" marks the first transect path, and the black circles indicate the turbine hub height.

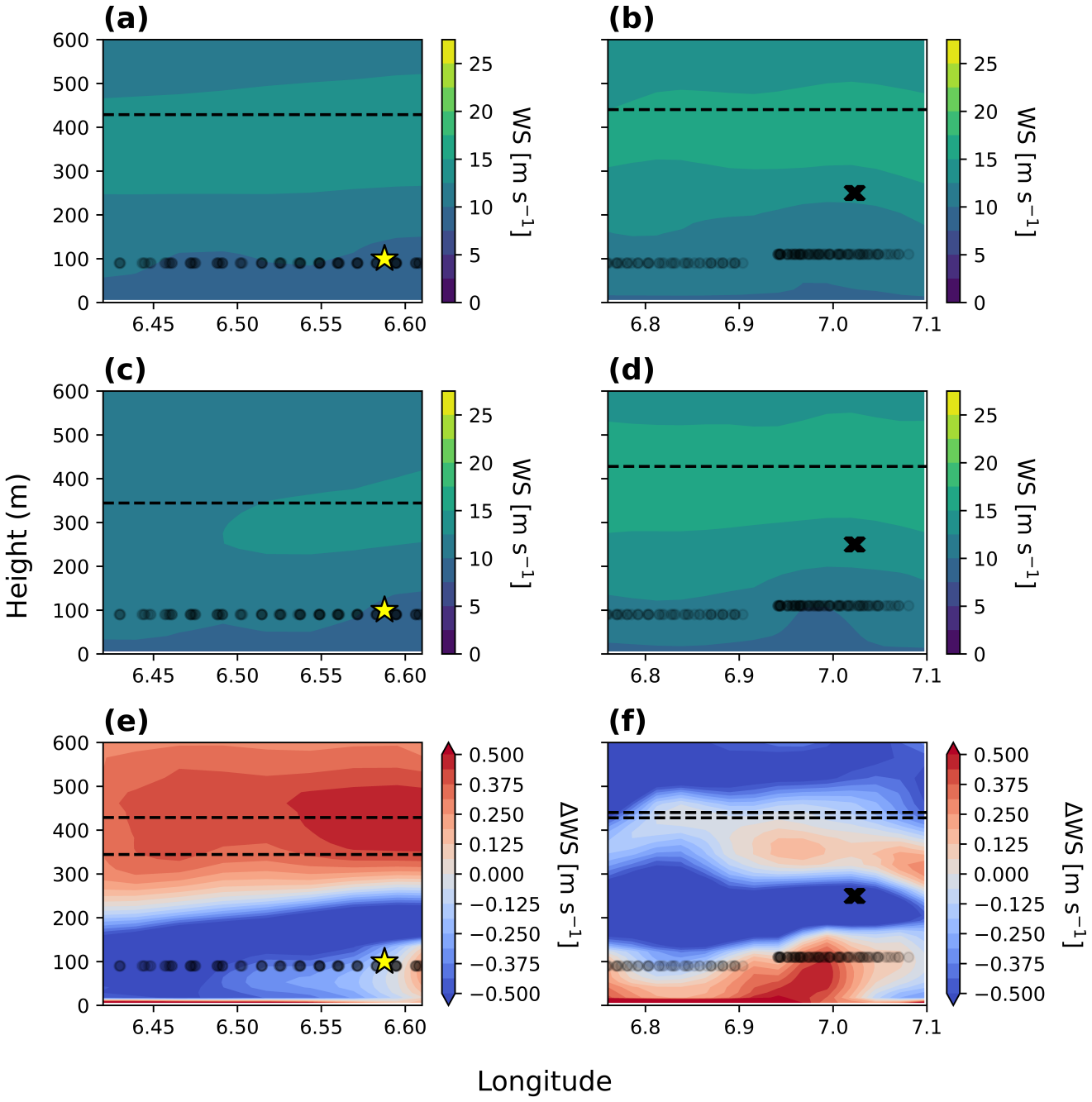

Figure 7. Modeled wind speed cross-section at a constant latitude of 54.03. (a) FINO1 3fa_025; (c) FINO1 mfa_025; (e) FINO1 3fa_025 - mfa_025; (b) aircraft 3fa_025; (d) aircraft mfa_025; (f) aircraft 3fa_025 - mfa_025. The horizontal dashed black line denotes the average modeled PBL height, the star indicates the FINO1 tower location, the "X" marks the first transect path, and the black circles indicate the turbine hub height.

mechanism. Here, TKE from aloft mixes more momentum from aloft into the measurement region, which increases wind speeds (Fig. 7a,c,e). Thus, the differing measurement altitudes between the two sites may impact both the TKE and wind speed assessments presented below.

3.1.2 Spatial variability

Wind field behavior near the turbines differs from that for the rest of the simulation domain, on average. MYNN average wind speeds are higher than 3DPBL average wind speeds outside of the turbine wakes (Fig. 8a). MYNN average wind speeds likely exceed 3DPBL average wind speeds in this area because the 3DPBL scheme has higher TKE (Fig. 8b). This larger TKE with the 3DPBL scheme extracts more momentum from the mean wind, resulting in a greater reduction in wind speeds. This finding that MYNN wind speeds are faster than 3DPBL wind speeds is consistent with other comparisons of these two PBL schemes, completed in both real and idealized conditions (Juliano et al., 2022; Rybchuk et al., 2022; Arthur et al., 2022; Peña et al., 2023; Arthur et al., 2024). This finding is also documented further in Fig. 7 and in Sect. A1.

In contrast, 3DPBL average wind speeds exceed MYNN Fitch average wind speeds in the turbine wakes (Fig. 8a). This distinct behavior in the wakes arises from differences in the drag forces for each PBL scheme: the drag forces are very sensitive to wind speed. Because the MYNN wind speeds are slightly faster when entering the wind farms, the resulting MYNN drag force (Eq. 15) is generally larger than the 3DPBL drag force due to the faster initial wind speeds in MYNN scheme has higher initial wind speeds which are still slower than the rated wind speed. As a consequence, the MYNN scheme shows stronger and longer wakes than the 3DPBL scheme, on average (Fig. 8c). The MYNN average wind speed reduction is sufficiently strong such that 3DPBL average wind speeds exceed MYNN average wind speeds within the turbine wake (Fig. 8a). Further, because 3DPBL average wind speeds exceed MYNN average wind speeds in this region, the 3DPBL scheme also has more turbine-induced TKE than the MYNN scheme (Fig. 8b).

3.2 Measurement variability

3.2.1 FINO1 tower

Simulations capture the slowing winds observed at FINO1. Modeled FINO1 Fitch wind speeds appropriately capture the temporal shifts throughout the observational period (Fig. 9). MYNN wind speeds are also consistently faster than 3DPBL wind speeds, as also shown in the spatial maps (Fig. 8a).

Observational agreement for FINO1 modeled wind speeds differs by measurement altitude. The median FINO1 modeled wind speeds at the highest locations (81 m, 91 m, and 102 m) perform best (compared to those at the lower altitudes of 34 m, 41 m, 51 m, and 61 m) for R^2 (Fig. 10b) and $cRMSE$ (Fig. 10c) and the worst for the bias (Fig. 10a) and EMD (Fig. 10d). The low (34 m and 41 m) and middle (51 m and 61 m) heights show similar performance for the bias (Fig. 10a), R^2 (Fig. 10b), and EMD (Fig. 10d), with some additional shaping to the $cRMSE$ (Fig. 10c). Further, the wind speeds at lower altitudes are also slower, resulting in smaller bias. Whereas 3DPBL EMD shifts from outperforming MYNN EMD at the low and middle heights to underperforming MYNN EMD for the higher heights, the physics-based trends are consistent across the

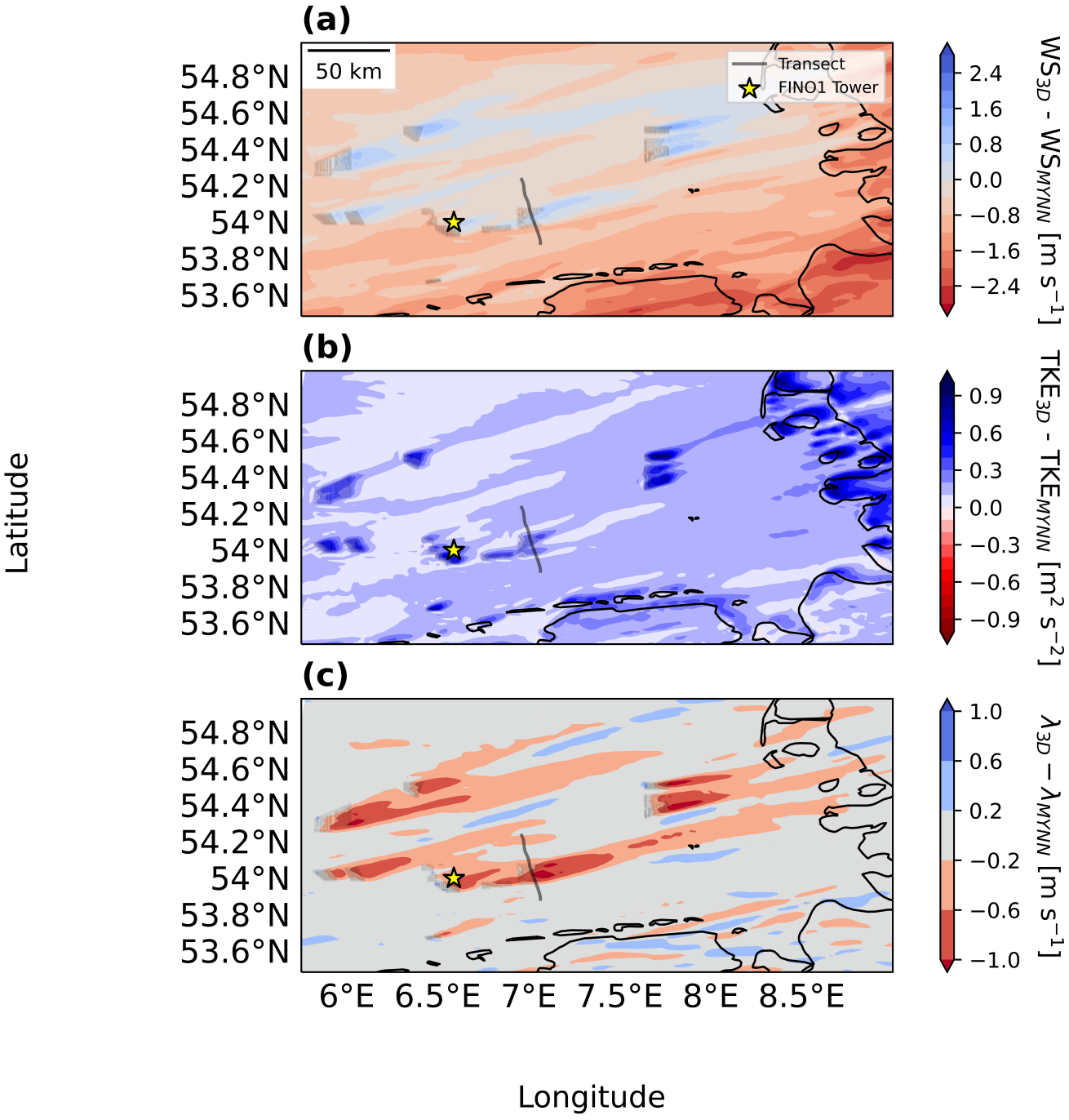

Figure 8. Difference fields for inner region. (a) 3fa_025 - mfa_025 WS; (b) 3fa_025 - mfa_025 TKE; (c) 3fa_025 - mfa_025 wake deficit. Turbines are marked with black circles, the FINO1 tower is marked with a yellow star, and the first transect path is marked with a solid line.

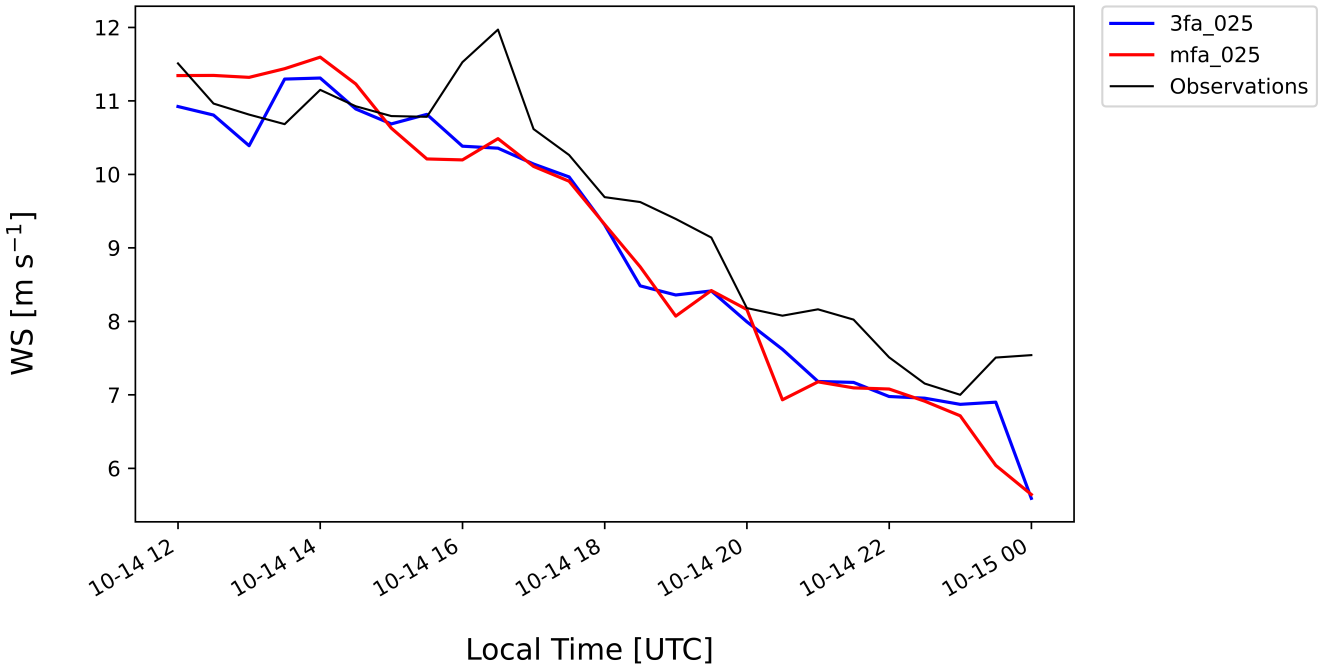

Figure 9. Time series of 76 m modeled horizontal wind speeds (WS) compared to 81 m FINO1 observations for the hours of 12:00:00–00:00:00. Both the modeled wind speeds and observed wind speeds are resampled to 30 minutes.

FINO1 model heights, which are stronger than the variability between heights. 3DPBL cases consistently show larger wind speed biases (Fig. 10a), more variable wind speed R^2 (Fig. 10b), and smaller wind speed $cRMSE$ (Fig. 10c) than the MYNN cases.

3.2.2 Aircraft region

Similar patterns emerge for the wake-affected region of the flight path. The general trends of increased TKE (Fig. 11a) and decreased wind speed (Fig. 11b) above the wind farm clearly emerge. TKE in the wake portion of the flight path is simulated to be stronger in the 3DPBL simulations than in the MYNN simulations (Fig. 11a), and the maximum value in the 3DPBL simulations is similar to the maximum value in the observations, although the location of the TKE maximum is shifted downwind in both simulations. As noted earlier, these differences in TKE between the PBL schemes reflect the fundamental differences between the models in length scales, empirical constants, and horizontal mixing approaches. Because the 3DPBL scheme has higher TKE, the 3DPBL scheme consequently shows slower wind speeds (Fig. 8b, 11a).

Simulated aircraft region wind speeds and TKE are also subject to systematic influences between transects. First, a performance discrepancy between odd and even transects exists. Simulated wind speeds and TKE for southeast-to-northwest transects (1, 3, and 5) have a larger bias and EMD than those for northwest-to-southeast transects (2, 4, and 6) (Fig. 12a,d , Fig. 13a,d).

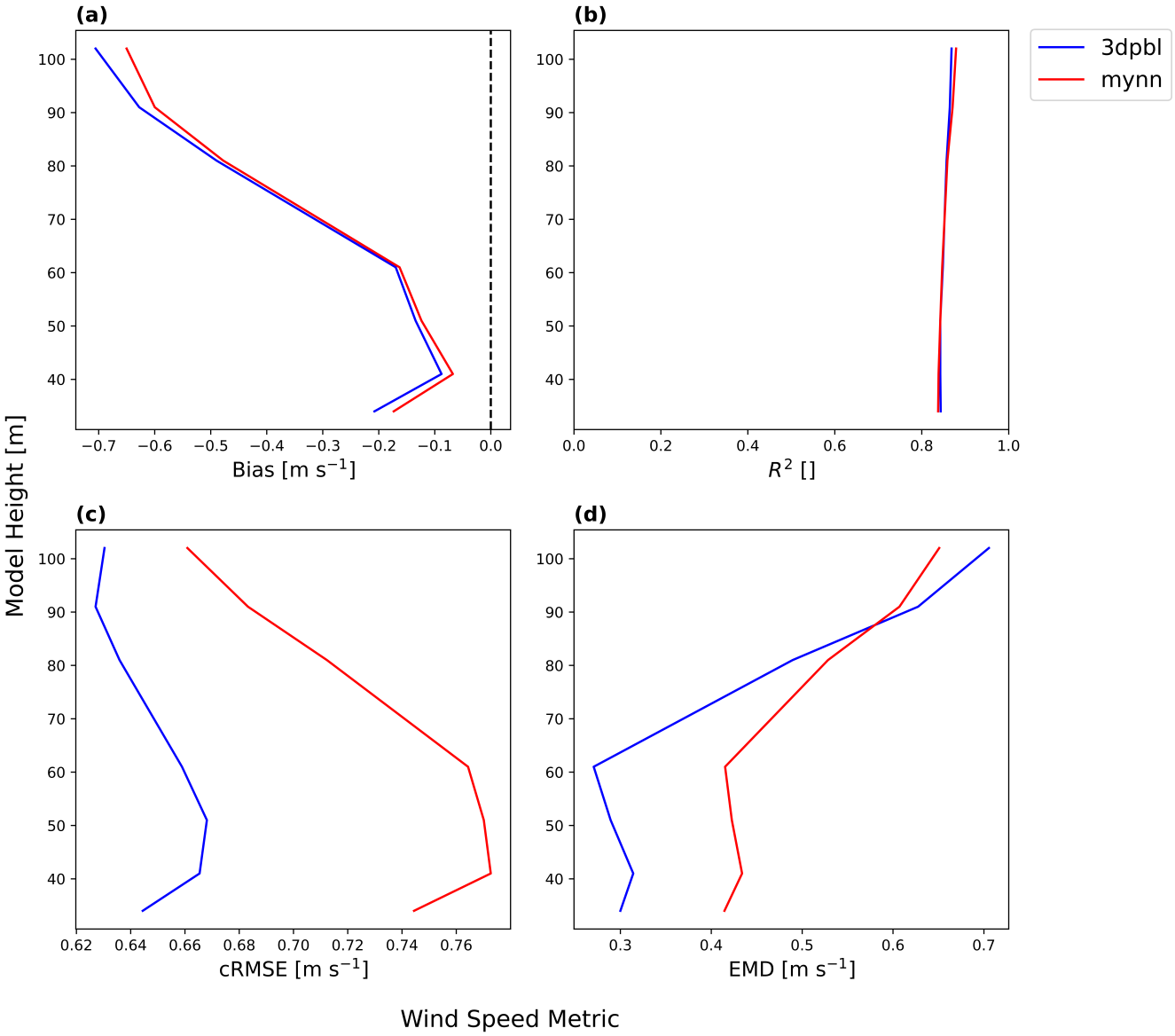

Figure 10. Error metric profile plot for FINO1 region wind speeds across all Fitch simulations with advection on and a wind farm TKE factor of 0.25. (a) Bias; (b) R^2; (c) $cRMSE$; (d) EMD.

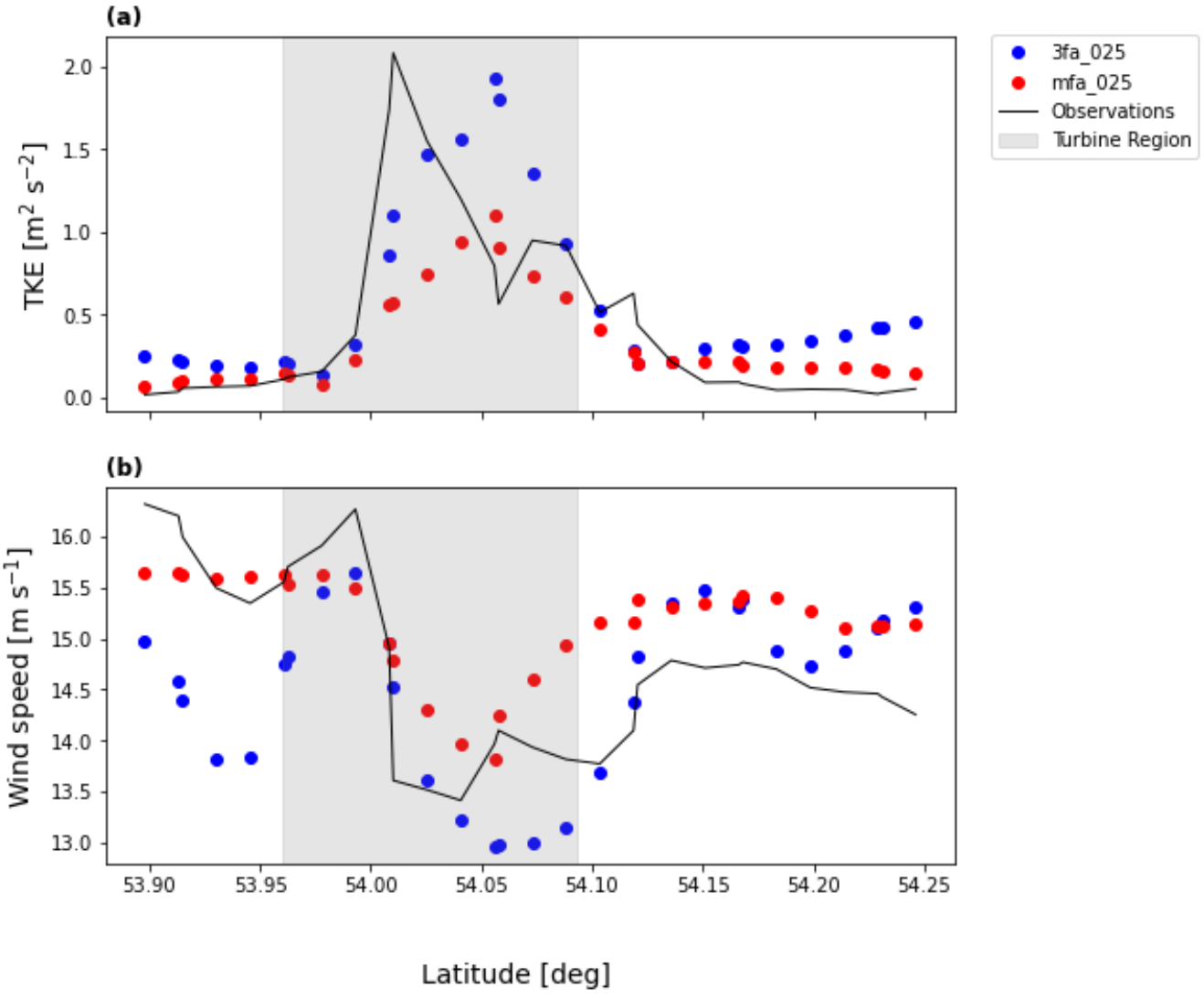

Figure 11. Aircraft region (250 m) simulated and observed (a) TKE and (b) wind speed across the flight path for transect 1. The shaded region incorporates the minimum and maximum latitudes of the turbine locations.

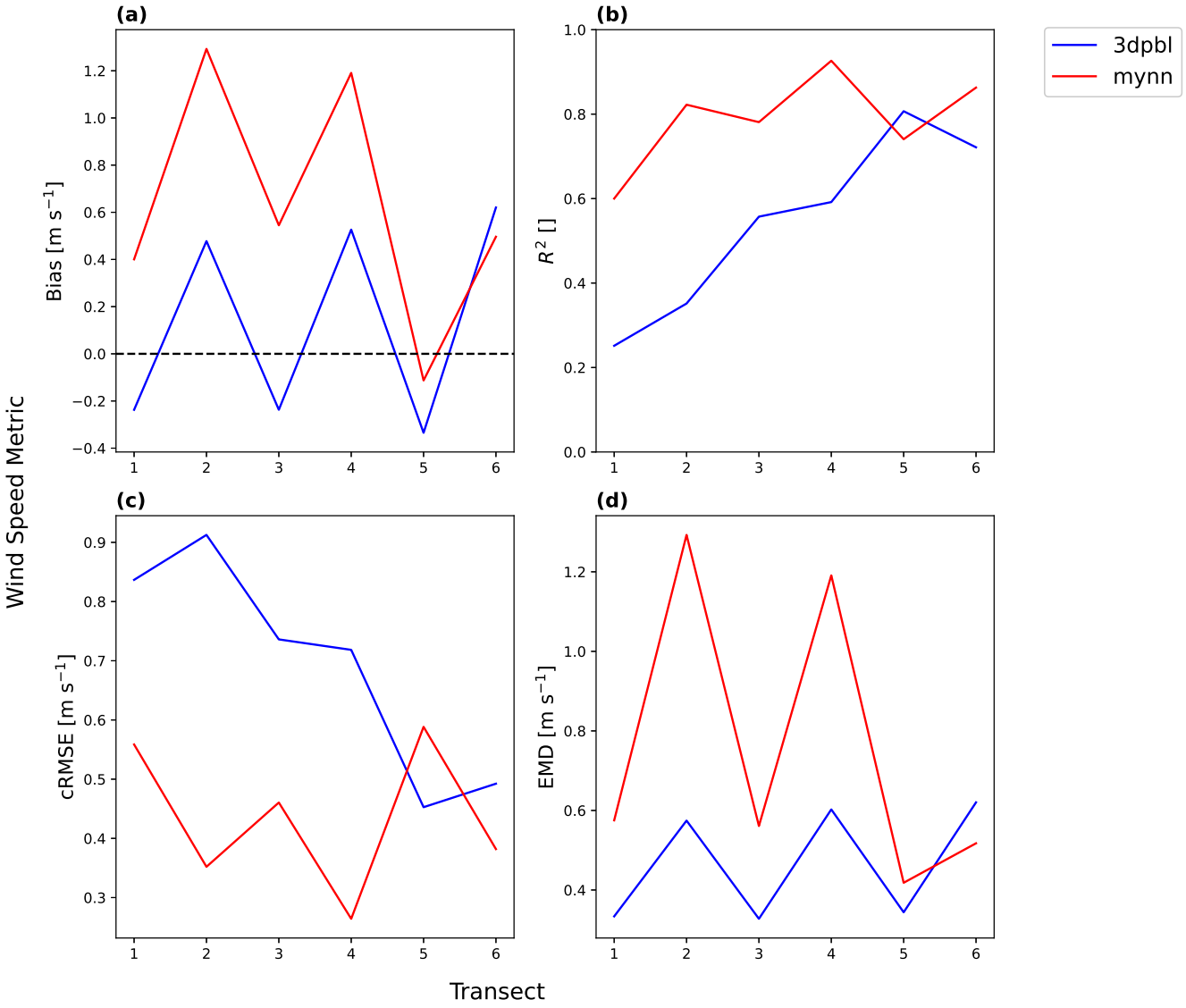

Figure 12. Error metric box plot by transect for aircraft region wind speeds across all Fitch simulations with advection on and a wind farm TKE factor of 0.25. (a) Bias; (b) R^2; (c) $cRMSE$; (d) EMD.

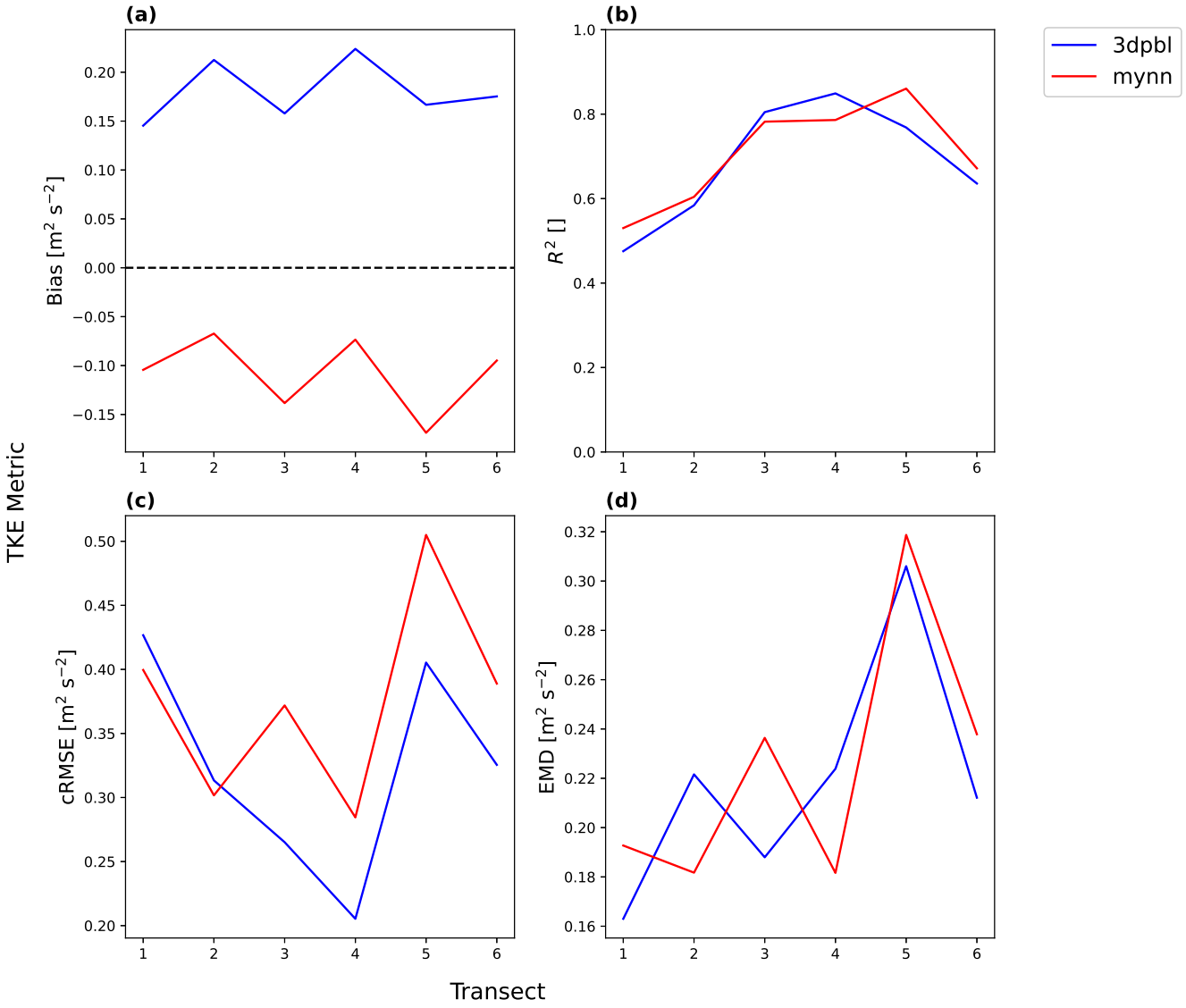

Figure 13. Error metric plot by transect for aircraft region TKE across all Fitch simulations with advection on and a wind farm TKE factor of 0.25. (a) Bias; (b) R^2; (c) $cRMSE$; (d) EMD.

This systematic transect variability corresponds to the reversed directions of the transect paths – whereas transects 1, 3, and 5 are performed in the northwesterly direction, transects 2, 4, and 6 are performed in the southeasterly direction.

This directional variability could potentially be explained by errors in yaw alignment. As noted earlier, the dominant source of error in one set of wind speed retrievals were previously determined to be the yaw measurements (Bärfuss et al., 2023). The error in yaw alignment was also determined to have alternating signs on whether the wind approached the sensor from the starboard or the backboard side (Bärfuss et al., 2023). Because the even and odd transects involved opposite alignments of the aircraft, it would follow that the odd and even transects would then show distinct wind speed (and TKE) errors.

Wind speed and TKE $cRMSE$ and R^2 in the aircraft region also reveal additional systematic influences. While the transect direction does suggest influence on these error metrics (Fig. 12b,c , Fig. 13b,c), these error metrics ($cRMSE$ and R^2) (Fig. 12b,c , Fig. 13b,c) additionally suggest an improvement over time that is not evident with the bias or the EMD (Fig. 12a,d , Fig. 13a,d). This performance improvement occurs after transect 2 and continues until transect 4 or 5, and is most evident from the R^2, although it does extend to the $cRMSE$, especially for the 3DPBL (Fig. 12c, Fig. 13c).

4 Model evaluation

Table 7. p values according to the Mann–Whitney U test. A bold cell indicates statistical significance at $|p| <= 0.05$. Only Fitch simulations with advection on and a wind farm TKE factor of 0.25 are considered.

Parameter	Quantity	Bias	R^2	EMD	$cRMSE$
PBL	FINO wind speed	0.7104	1.000	0.3829	**0.0023**
PBL	Aircraft region wind speed	**0.1797**	**0.0260**	**0.3095**	**0.0411**
PBL	Aircraft region TKE	**0.0022**	0.8182	0.8182	0.5887

A primary goal of this effort is to explore how simulated wake behavior changes based on the PBL scheme. The optimal PBL scheme depends on the site, error metric, and variable considered. The statistical significance of these differences, however, is more constrained. A statistically significant difference between 3DPBL and MYNN exists for wind speeds at the FINO1 site according to the $cRMSE$ (Table 7) but not for the bias, R^2, or EMD. TKE in the aircraft region also shows a statistically significant difference for a single metric – in this case, the bias (Table 7). In contrast, a statistically significant difference between 3DPBL and MYNN is present for all four metrics within the aircraft region (Table 7).

In the FINO1 region, 3DPBL wind speeds outperform MYNN wind speeds in representing wind speeds with respect to $cRMSE$ (Fig. 14d), while the bias (Fig. 14a), R^2 (Fig. 14g), and EMD (Fig. 14j) are comparable between the PBL schemes. In contrast, the MYNN scheme holds a more decisive lead in comparison to observations collected 250 m above the surface and 100 m above a wind farm. MYNN TKE outperforms 3DPBL TKE with respect to bias (Fig. 14c), with no clear winner for $cRMSE$ (Fig. 14f), R^2 (Fig. 14i), or EMD (Fig. 14l). For wind speeds evaluated with the aircraft dataset, MYNN outperforms 3DPBL with respect to $cRMSE$ (Fig. 14e) and R^2 (Fig. 14h), whereas 3DPBL outperforms MYNN with respect to bias (Fig. 14b) and EMD (Fig. 14k).

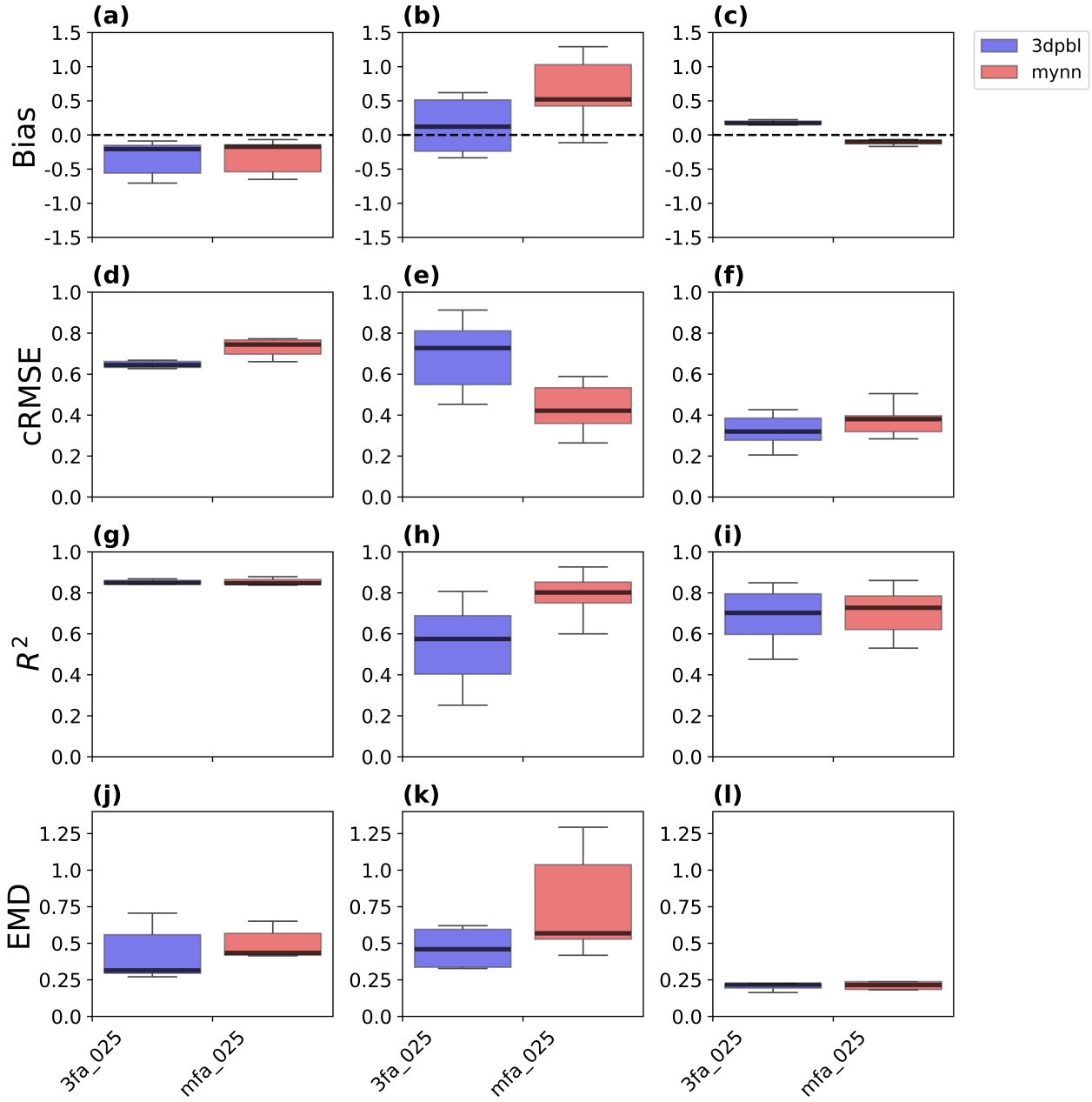

Figure 14. Error metrics across all Fitch simulations with the advection option on and with a wind farm TKE factor of 0.25. The first column corresponds to FINO1 wind speed error metrics, the second column corresponds to aircraft wind speed error metrics, and the third column corresponds to aircraft TKE error metrics. The box encloses the interquartile range (IQR), and the whiskers extend to Q1-1.5*IQR and Q3+1.5*IQR. Units are non-normalized and reflect the units of the evaluated field. (a) FINO1 wind speed bias [m s^{-1}]; (b) aircraft wind speed bias [m s^{-1}]; (c) aircraft TKE bias [m^2 s^{-2}]; (d) FINO1 wind speed $cRMSE$ [m s^{-1}]; (e) aircraft wind speed $cRMSE$ [m s^{-1}]; (f) aircraft TKE $cRMSE$ [m^2 s^{-2}]; (g) FINO1 wind speed R^2 [ ]; (h) aircraft wind speed R^2 [ ]; (i) aircraft TKE R^2 [ ]; (j) FINO1 wind speed EMD [m s^{-1}]; (k) aircraft wind speed EMD [m s^{-1}]; (l) aircraft TKE EMD [m^2 s^{-2}]).

The swapped bias performances between wind speed and TKE in the aircraft region may reflect competing optimizations. The TKE bias differs in sign between the two PBL schemes. The 3DPBL scheme overpredicts TKE (Fig.14c), while the MYNN scheme underpredicts TKE (Fig. 14b). This higher TKE with the 3DPBL scheme implies that the 3DPBL scheme also has slower average wind speeds than the MYNN scheme (Fig. 8a, Fig. 11a,b). These slower average wind speeds translate to an artificially-improved wind speed bias (Fig. 14b). While the 3DPBL TKE overprediction improves wind speed performance on average, this performance enhancement does not necessarily translate to other wind speed error metrics. Notably, the MYNN scheme outperforms the 3DPBL scheme with respect to both wind speed $cRMSE$ (Fig. 14e) and wind speed R^2 (Fig. 14h). The role of TKE overpredictions in wind speed optimizations is further demonstrated in Sect. A2.

Previous studies have also found that the optimal PBL scheme depends on site and metric constraints. Draxl et al. (2014) found that the optimal PBL scheme depended on the atmospheric stability in an analysis of the Høvsøre wind farm. Hahmann et al. (2020) found that the MYNN scheme outperformed the Yonsei University (YSU) scheme in weekly simulations that used spectral nudging in characterizing eight sites in Northern Europe. Hahmann et al. (2020) also performed a series of 25 simulations to compare PBL schemes, land surface models, and surface layer options and found that the differences between the sites were larger than the differences between PBL schemes, regardless of the error metric. Sheridan et al. (2024) found that the MYNN-based dataset both identified more of the extremely LLJ events and had a higher rate of false LLJ identification than the YSU-based dataset in the North Atlantic. Storm et al. (2009) similarly forecast LLJs in the Great Plains and found that, although the YSU scheme outperformed the Mellor–Yamada–Janjic (MYJ) PBL scheme in forecasting the wind direction, the MYJ scheme outperformed the YSU scheme with respect to wind speed for the same West Texas case. Storm et al. (2009) ascribe these differences to be largely site-dependent and argue that the results would likely be different at another site. Olsen et al. (2025) analyzed LLJs from mast and lidar measurements in the North and Baltic seas by varying several modeling parameters including grid cell spacing, surface layer parameterizations, PBL schemes, and vertical resolution. Olsen et al. (2025) found that the PBL scheme was the most significant parameter in determining model results and that the 3DTKE scheme performed better than the other schemes considered.

Our results highlighting the advantages of the 3DPBL scheme within the turbine rotor layer are also supported by previous MYNN and 3DPBL intercomparisons. Arthur et al. (2022) found that the 3DPBL scheme consistently outperformed the MYNN scheme for a Columbia River Gorge site. Similarly, Arthur et al. (2024) compared the MYNN and 3DPBL schemes for an Altamont Pass site and also found the 3DPBL scheme to reduce wind speed error. Our $cRMSE$ results in the turbine rotor layer align with those of Arthur et al. (2022) in that they both support the 3DPBL scheme over the MYNN scheme in stably stratified conditions. Our results also align with those of Arthur et al. (2024) in that they both show the 3DPBL scheme outperforming the MYNN scheme within the turbine rotor layer. Notably, whereas Arthur et al. (2022) and Arthur et al. (2024) were both performed in areas of complex terrain, which has been theorized to provide a theoretical advantage to the 3DPBL scheme (Kosović et al., 2020; Juliano et al., 2022), this work considers the relatively homogeneous offshore area, suggesting that the advantages of the 3DPBL scheme are not confined to complex terrain.

Grid cell spacing may influence the PBL comparison in this work. Notably, the 3DPBL is theorized to improve model performance in the "terra incognita", where the NWP horizontal grid spacing Δx approaches a similar magnitude to the PBL

depth z (Wyngaard, 2004). Indeed, Arthur et al. (2022) found that the 3DPBL scheme outperformed the MYNN scheme in their simulations with a large-eddy simulation (LES) grid that has a finer cell spacing than that used in our mesoscale simulations. Thus, the mesoscale grid cell spacing necessary for this analysis, while vital to contextualize the results of this work with the broader literature, may restrict the potential benefits of the 3DPBL scheme over the MYNN scheme. At the same time, Peña et al. (2023) compared 3DPBL and MYNN to LES output and found that MYNN outperformed 3DPBL when the finest grid cell spacing was used. Thus, grid cell spacing may influence–but not necessarily dictate–the optimal PBL scheme.

Differences in the relative measurement height between the two sites in this case study may also affect the PBL comparison. Whereas the FINO1 tower is within the turbine rotor region, the aircraft measurements are taken more than 100 m above the turbines in a stably stratified boundary layer (Fig. 4) that suppresses some interactions between the atmosphere sampled at the turbine level and the aircraft level (Fig. 6). Thus, the 3DPBL scheme improves $cRMSE$ turbine-induced turbulence characterization in the turbine rotor region as sampled at FINO1 (Fig. 14d) and overpredicts turbine-induced turbulence aloft as sampled by the aircraft measurements, on average (Fig. 14c).

The stable stratification present in this case study also improves the utility of the results of this PBL comparison. By restricting this analysis to time periods considered in previous analyses for this case study, not only are the results contextualized within the broader literature, but the conditions that contribute the strongest and longest wakes are also highlighted. Thus, while other analyses of this region may approach the lack of available in situ observations by introducing statistical downscaling methods to explore scientific questions around diurnal, seasonal, and climatic trends (Fischereit et al., 2022b), this analysis instead addresses scientific questions that are best-suited with in situ observations alone.

Other site-based considerations are less likely to be driving the relative PBL performance trends. The documented even/odd transect variability (Fig. 12, Fig. 13) is not likely to drive the PBL-based differences given that the PBL-based differences are larger than the transect variability. The PBL-based preferences are also reinforced at all heights of the FINO1 tower (Fig. 9a,c,e), even despite performance variability between FINO1 heights (Fig. 10).

5 Conclusions

This work addresses model capabilities for representing wind farm wakes with the Fitch WFP in the WRF model. This question underpins an important and understudied sensitivity of the NWP models that support both wind power forecasting and wind energy assessment. This work explores this question as one of the first comparative evaluations between the 1D MYNN PBL scheme and the newly-Fitch-integrated NCAR 3DPBL scheme against two sets of in situ observations for an offshore case study.

The optimal PBL scheme depends on the site. For wind speeds modeled at the site within the turbine rotor region, 3DPBL outperforms MYNN with respect to the $cRMSE$. In contrast, for TKE modeled within a region 100 m above a wind farm, MYNN outperforms 3DPBL with respect to the bias. Ambiguously, for wind speeds modeled within the region 100 m above a wind farm, MYNN wind speeds outperform 3DPBL wind speeds with respect to R^2 and $cRMSE$ but underperform with respect to bias and EMD. These site-based differences in the optimal PBL scheme likely reflect differences in the relative

measurement height between the two sites. Whereas the FINO1 observations are collected within the turbine rotor region, the aircraft observations are collected 100 m above the turbine rotor layer, such that interactions between the turbine-induced turbulence and the aircraft measurements may be constrained by a stable boundary layer. Thus, 3DPBL simultaneously more appropriately characterizes behavior within the turbine rotor layer and overpredicts the amount of TKE reaching the aircraft region site.

Subsequent investigations could explore other case studies to provide perspective into the generalizability of the results across other sites. Similarly, datasets from the third Wind Forecast Improvement Project (WFIP3) could be useful to explore how offshore wind characterization might differ between the North Sea and the eastern United States (WFIP3). Moreover, datasets from the land-based, horizontally homogeneous American WAKE experimeNt – or AWAKEN – campaign (Moriarty et al., 2024) could be useful to study because previous land-based studies analyzing the 3DPBL scheme have involved complex terrain and far fewer detailed observations.

Further, this study focuses on simulations of the atmosphere alone, without coupling to the ocean surface or the water below. Another example of future work may include performing this analysis with an ocean model coupling (Raghukumar et al., 2022, 2023; Daewel et al., 2022). Although introducing a coupled ocean–atmosphere model does not always improve performance (Gaudet et al., 2024), introducing ocean coupling to this model could help diagnose sources of wind speed error for this model. Introducing a coupled ocean–atmosphere model also provides physical insight by potentially affecting the relative performance between the PBL schemes. Introducing ocean coupling also provides a more direct comparison to the work of Larsén et al. (2024), which included atmosphere–wave coupling, for this case study. Recreating this analysis with an additional ocean model like the Regional Ocean Modeling System (Shchepetkin and McWilliams, 2005) could provide insight into the influence of oceanic forcings on wind wake behavior and the resulting consequences on surface currents.

Data availability. The FINO1 data can be downloaded here: https://insitu.bsh.de/rave/index.jsf?content=insitu, and the WIPAFF aircraft data can be downloaded here: https://doi.pangaea.de/10.1594/PANGAEA.903088. Post-processing code is available here: 10.5281/zenodo.14751600. The WRF code used in this analysis is based on the publicly-available version of 3DPBL code provided here: <update on acceptance>

Appendix A: Additional model runs

A1 Contrast between NWF and WF options

Average wind speeds for the Fitch and NWF cases are similar in regions without wind turbines and differ near the turbines. Both NWF (Fig. A1a,c) and Fitch (Fig. A1b,d) average wind speeds are lower over land than over water. Further, the fastest average wind speeds are immediately to the west of the coast for both the NWF (Fig. A1a,c) and Fitch (Fig. A1b,d) cases. The predominant westerly wind direction and the significant distance of the measurement regions from the coast demonstrate that the measurement regions are not influenced by the coastal decelerations. MYNN average wind speeds (Fig. A1c,d) are

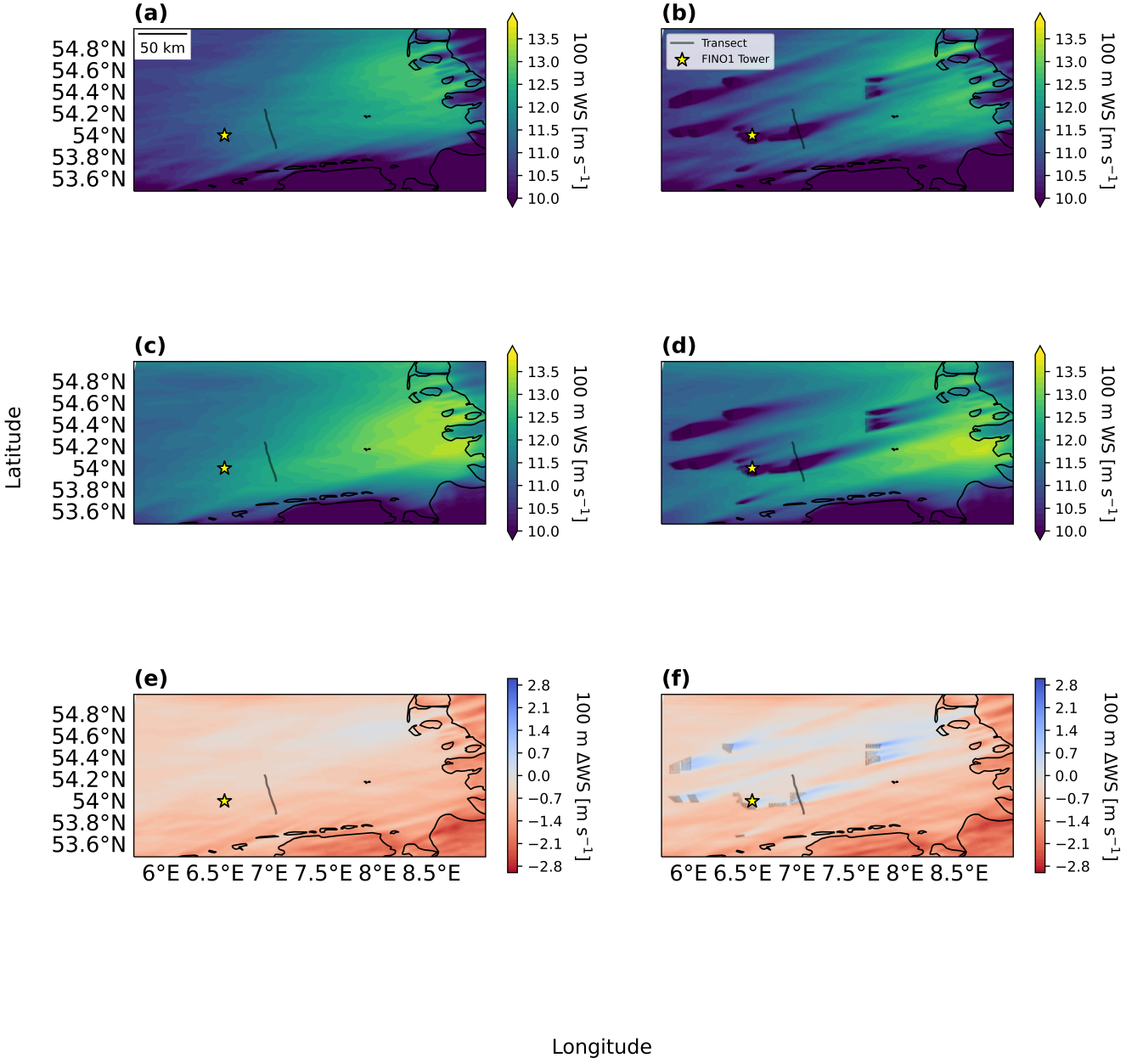

Figure A1. Horizontal wind speeds for the inner 1.67 km domain. (a) 3na_NA. (b) 3fa_025. (c) mna_NA. (d) mfa_025. (e) 3na_NA - mna_NA. (f) 3fa_025 - mfa_025. Turbines are marked with black circles, the FINO1 tower is marked with a yellow star, and the first transect path is marked with a solid line.

also consistently higher than 3DPBL average wind speeds (Fig. A1a,b) for both NWF and Fitch cases outside of the turbine wakes (Fig. A1e,f). This difference in wind speeds between the PBL schemes can be explained by TKE differences between the two PBL schemes. Because the 3DPBL scheme has larger TKE (Fig. 8b), the 3DPBL scheme extracts more momentum and reduces wind speeds further. In contrast, Fitch and NWF wind speeds differ near the turbines. Notably, 3DPBL Fitch average wind speeds exceed MYNN Fitch average wind speeds in the turbine wakes (Fig. A1f). This reversal of which PBL scheme shows the higher average wind speed can be explained by differences in the turbine drag force between the two PBL schemes. The MYNN scheme has a stronger turbine drag force (Eq. 15, Fig. 3b) because of its higher initial wind speeds, which also implies that the MYNN scheme has stronger and deeper wakes (Fig. 8c). This reversal of which PBL scheme shows the higher average wind speed is isolated to the monotonically increasing region of the drag (proxy) curve (Fig. 3b). If the wind speeds were instead within the monotonically decreasing region of the drag (proxy) curve (Fig. 3b), MYNN wind speeds would likely exceed 3DPBL wind speeds even in the wakes. Given that NWF wind speeds mirror Fitch wind speeds outside of the turbine wakes and NWF wind speeds differ from Fitch wind speeds within the wakes, the dominant mechanism for these differences is more likely related to the turbines and not to the underlying meteorology.

A2 Effect of wind farm TKE factor, C_{TKE}

3DPBL mean TKE overpredictions improve wind speed bias performance in the aircraft region. The 3DPBL scheme overpredicts TKE in the aircraft region, regardless of the wind farm TKE factor (Fig. A2a). Further, increasing the wind farm TKE factor exacerbates this TKE overprediction (Fig. A3a). This worsening 3DPBL TKE bias with an increasing wind farm TKE factor is mirrored by an opposing trend in 3DPBL wind speeds. As the wind farm TKE factor increases, the 3DPBL wind speed bias decreases (Fig. A3a). Thus, the additional turbulence implies greater momentum extraction in the aircraft region and this greater momentum extraction slows the winds.

A similar relationship between TKE bias and wind speed bias is observed with the MYNN scheme in the aircraft region. While the MYNN scheme initially underpredicts TKE with a low (i.e. 0 or 0.25) wind farm TKE factor, as the wind farm TKE factor increases, the TKE bias becomes less negative (Fig. A2a) and the wind speed bias improves (Fig. A3a). One key difference between the trends observed with the 3DPBL and MYNN schemes is that the 3DPBL scheme shows a more optimal wind speed bias than the MYNN scheme. By overpredicting TKE, the 3DPBL scheme artificially improves (mean) wind speed performance over the MYNN scheme in the aircraft region.

This artificial performance enhancement due to TKE overprediction does not necessarily translate to other performance metrics. Notably, the wind speed $cRMSE$ (Fig. A3b) and wind speed R^2 (Fig. A3c) with the MYNN scheme outperform the wind speed $cRMSE$ (Fig. A3b) and wind speed R^2 (Fig. A3c) with the 3DPBL scheme in the aircraft region and this performance differential widens as the wind farm TKE factor increases. Thus, while TKE overpredictions with the 3DPBL scheme may artificially improve mean behavior, other metrics may still provide a more physical interpretation of model performance.

The relationship between an increased wind farm TKE factor and mean wind speed performance also differs in the FINO1 region. For the FINO1 region, both PBL schemes initially underpredict the mean wind speeds and improve the wind speed bias with an increasing wind farm TKE factor (Fig. A4a). In the FINO1 region, increased turbulence in the rotor region extracts

more momentum from aloft into the measurement region and increases wind speeds. The discrepancy between wind speed bias and wind speed patterns with other performance metrics is likewise sustained in the FINO1 region. For example, increasing the wind farm TKE factor improves the wind speed bias (Fig. A3a) while worsening the $cRMSE$ (Fig. A4b). Given that metrics may not always agree and some metrics may not fully reflect model performance alone, it is vital to consider a suite of metrics when determining an appropriate model configuration.

A3 Effect of TKE advection

Responses to the advection option may reflect locational differences. The results at the FINO1 location show greater sensitivity to the advection option than those at the aircraft measurement site. First, the FINO1 site is marked by a single grid cell at a wind farm's edge (Fig. 2a) and surrounded by other cells with wind turbines. Slight horizontal adjustment of this FINO1 tower location would capture the local TKE source as opposed to the sink (Fig. A5c,d). In addition, because the FINO1 performance is defined by a single grid cell, local TKE imbalances between grid cells are not compensated in averaging calculations. In contrast, the aircraft measurements experience several competing interfarm advection pools that are compensated across a broader domain (Fig. A5a,b). The two sites are also at different measurement heights. As explored earlier, the FINO1 site more directly interacts with the turbine-induced turbulence at the rotor level than does the aircraft region site (Fig. 6). As such, the FINO1 site is more sensitive to TKE differences. Although introducing advection does support some vertical transport of TKE at the aircraft measurement locations (Fig. A6a,b), maximum TKE is still better physically aligned with the measurement altitude for the FINO1 observations (Fig. A6c,d). As a consequence, the FINO1 region measurements show greater sensitivity to the advection option than the aircraft region measurements.

Differences in the amount of TKE generated by the PBL scheme may also explain responses in the advection option. Local TKE imbalances are stronger for 3DPBL TKE (Fig. A5a,c, Fig. A6a,c) than for MYNN TKE (Fig. A6b,d, Fig. A6b,d). Because the 3DPBL scheme has larger TKE (Fig. 8b), the 3DPBL scheme is more sensitive to TKE movement throughout the region.

Note that all performance differences based on the advection option are statistically insignificant (Table A1). This statistical insignificance is maintained for all evaluated fields, error metrics, and sites. Thus, even when considering TKE characterization for the single-celled FINO1 site, the effects on performance for this analysis are negligible.

Table A1. p values according to the Mann–Whitney U test. A bold cell indicates statistical significance at $|p| <= 0.05$. Only Fitch simulations with a wind farm TKE factor of 0.25 were considered.

Parameter	Quantity	Bias	R^2	EMD	$cRMSE$
Advection	FINO wind speed	0.5972	0.1129	0.3462	0.0769
Advection	Aircraft region wind speed	0.8852	0.5444	0.7508	0.5444
Advection	Aircraft region TKE	0.3408	0.1260	0.4705	0.4357

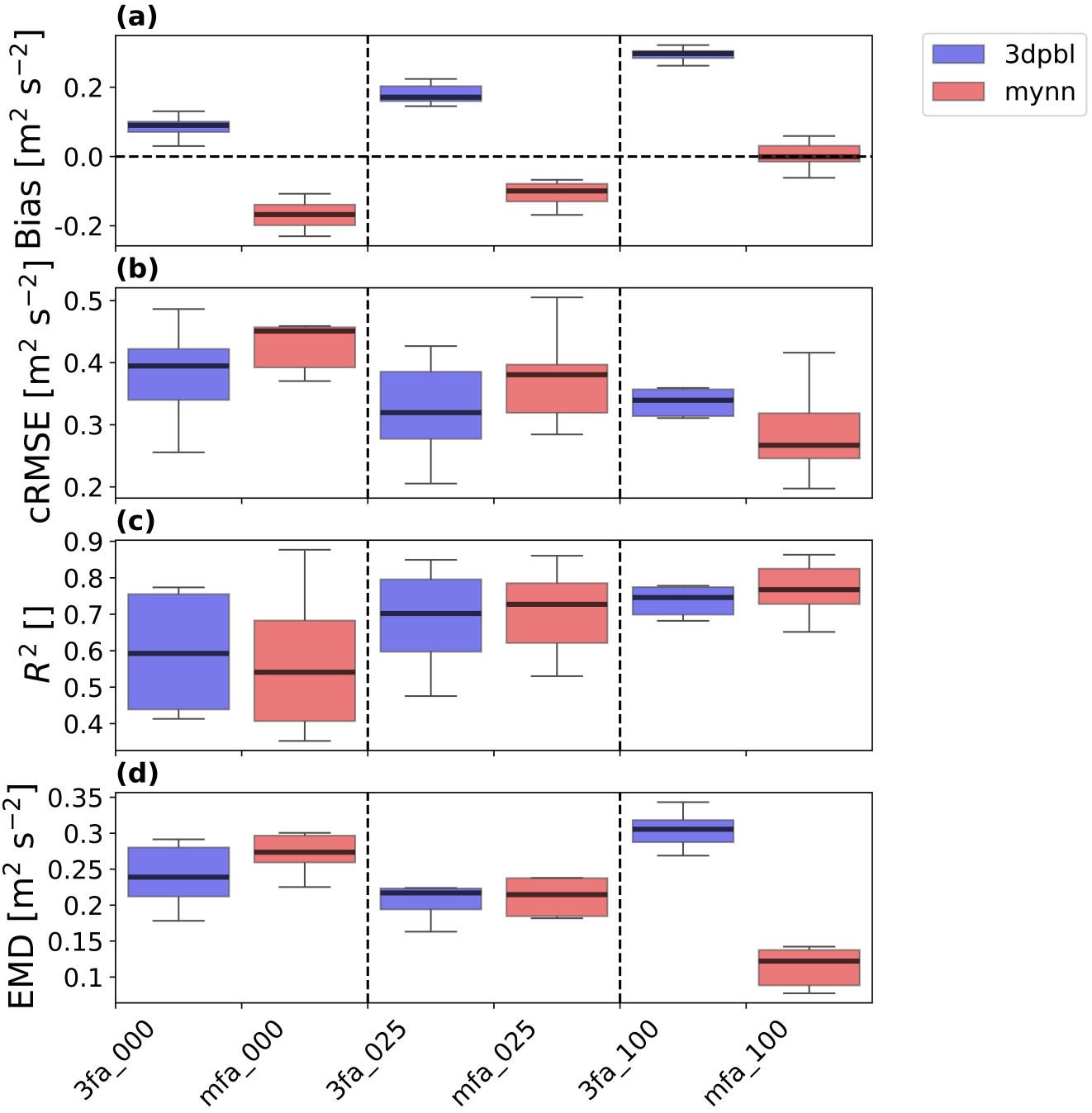

Figure A2. Error metric box plot for aircraft observations collected at 250 m for TKE. The box and whiskers describe aircraft transect variability and are based on Q1 (25th percentile), Q3 (75th percentile), and the interquartile range (IQR) (Q3–Q1). The box encloses the IQR, and the whiskers extend to Q1-1.5*IQR and Q3+1.5*IQR. The simulation names are mapped according to the short names provided in Table 6, and the vertical dotted lines visually separate simulations by wind farm TKE factor. (a) TKE bias; (b) TKE $cRMSE$; (c) TKE R^2; (d) TKE EMD.

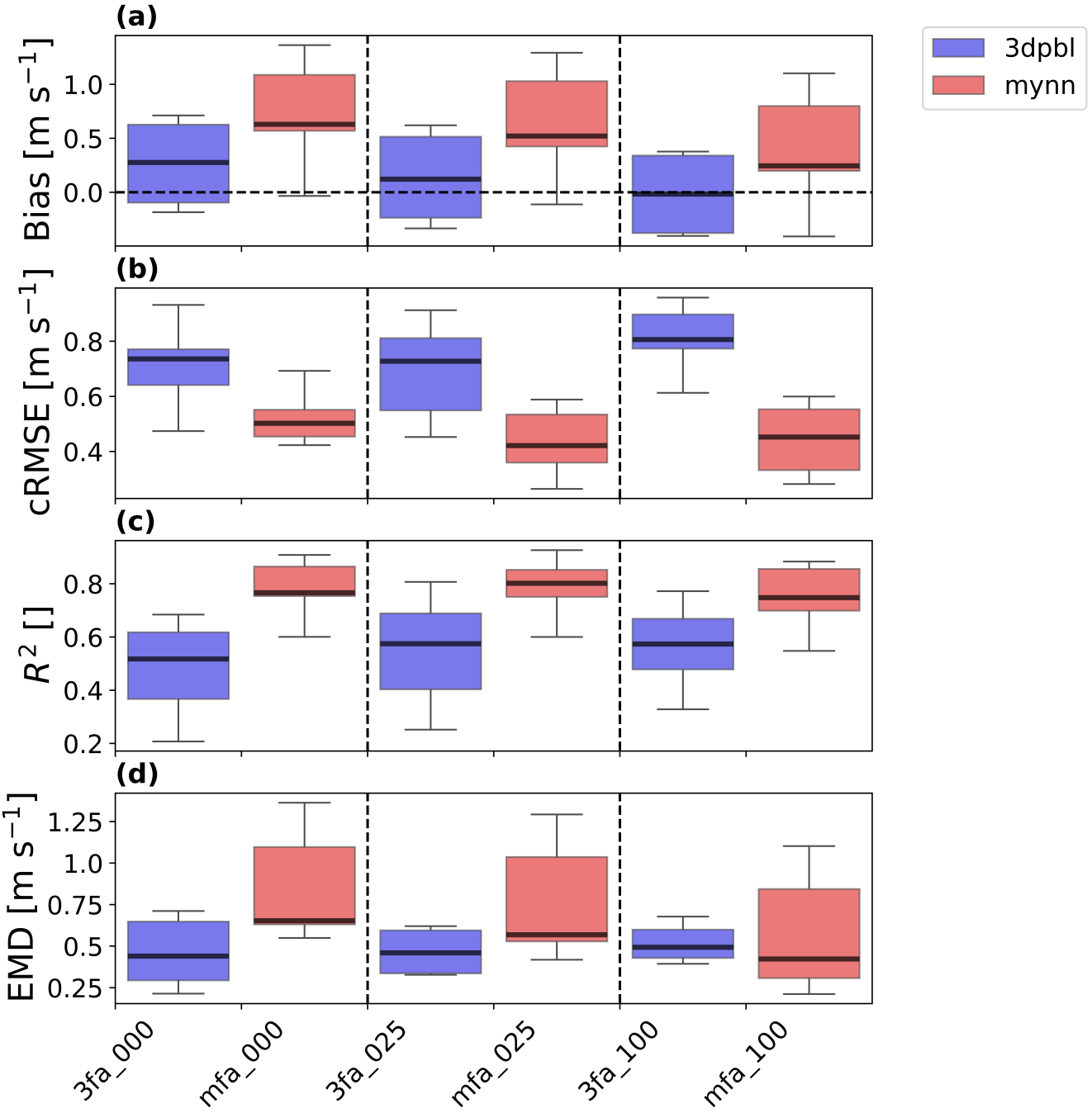

Figure A3. Error metric box plot for aircraft observations collected at 250 m for wind speed. The box and whiskers describe aircraft transect variability and are based on Q1 (25th percentile), Q3 (75th percentile), and the interquartile range (IQR) (Q3–Q1). The box encloses the IQR, and the whiskers extend to Q1-1.5*IQR and Q3+1.5*IQR. The simulation names are mapped according to the short names provided in Table 6, and the vertical dotted lines visually separate simulations by wind farm TKE factor. (a) Wind speed bias; (b) wind speed $cRMSE$; (c) wind speed R^2; (d) wind speed EMD.

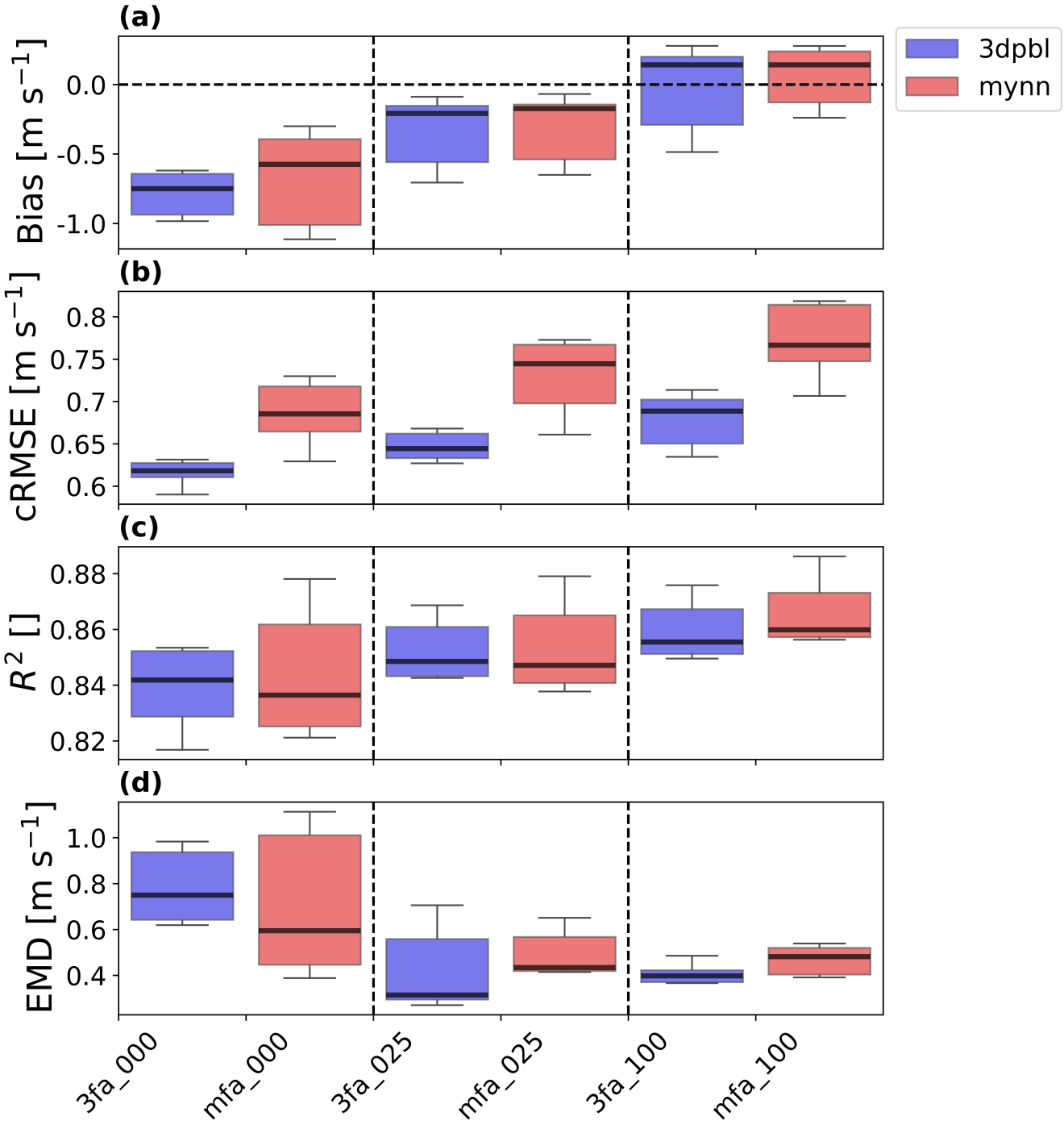

Figure A4. Error metric box plot for the FINO1 tower measurements for wind speed. The box and whiskers describe FINO1 model height variability and are based on Q1 (25th percentile), Q3 (75th percentile), and the interquartile range (IQR) (Q3–Q1). The box encloses the IQR, and the whiskers extend to Q1-1.5*IQR and Q3+1.5*IQR. The simulation names are mapped according to the short names provided in Table 6, and the vertical dotted lines visually separate simulations by wind farm TKE factor. (a) Wind speed bias; (b) wind speed $cRMSE$; (c) wind speed R^2; (d) wind speed EMD.

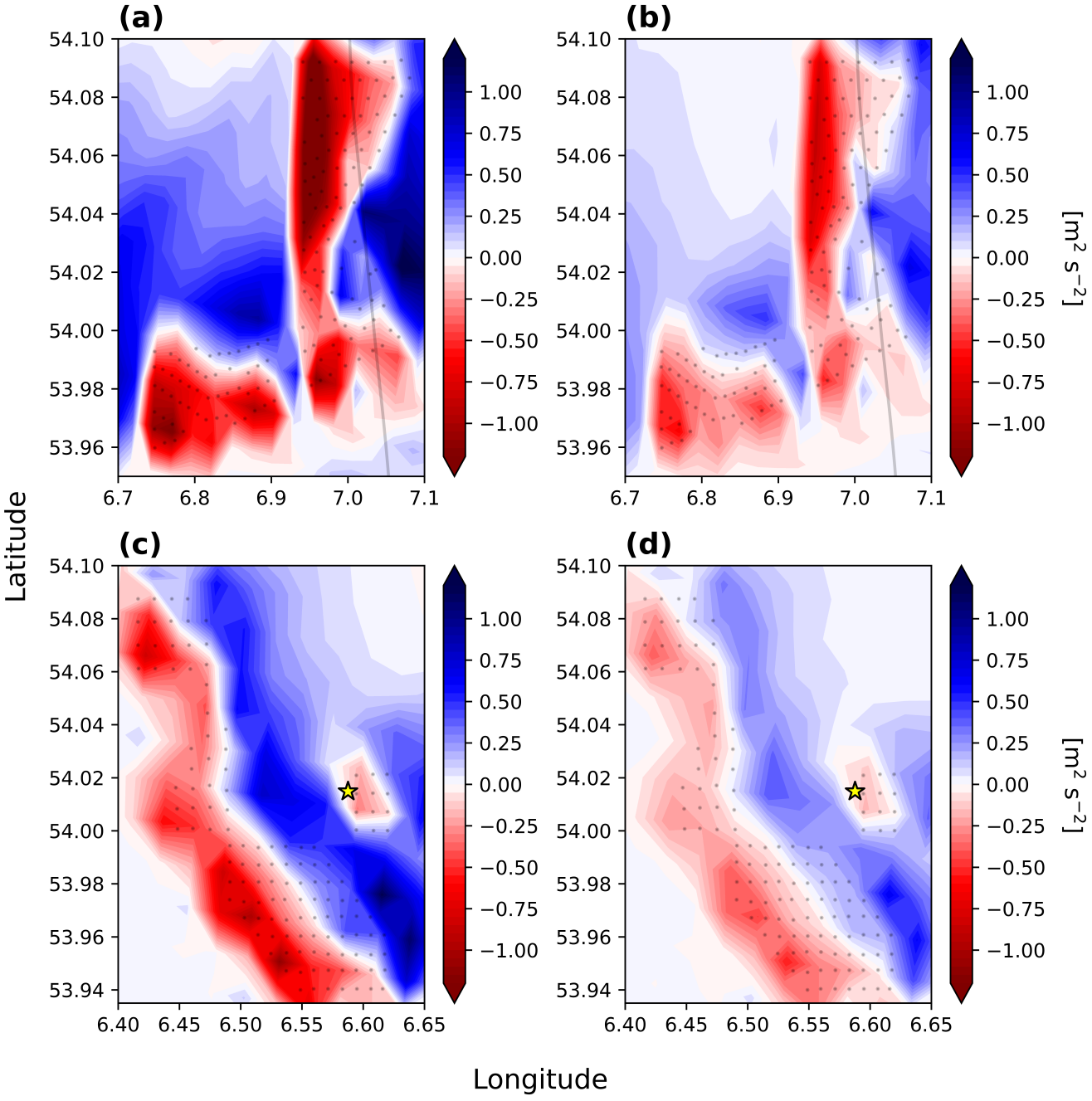

Figure A5. Horizontal slices to mark TKE differences between cases with advection on and cases with advection off. Red indicates that TKE is higher without advection. (a) Aircraft region 3fX_025. (b) Aircraft region mfX_025. (c) FINO1 3fX_025. (d) FINO1 3fX_025. FINO1 cases are averaged over hours 12:00:00–00:00:00, and the aircraft region cases are averaged over 14:10:00–16:10:00. Turbines are marked with black circles, the FINO1 tower is marked with a yellow star, and the first transect path is marked with a solid line.

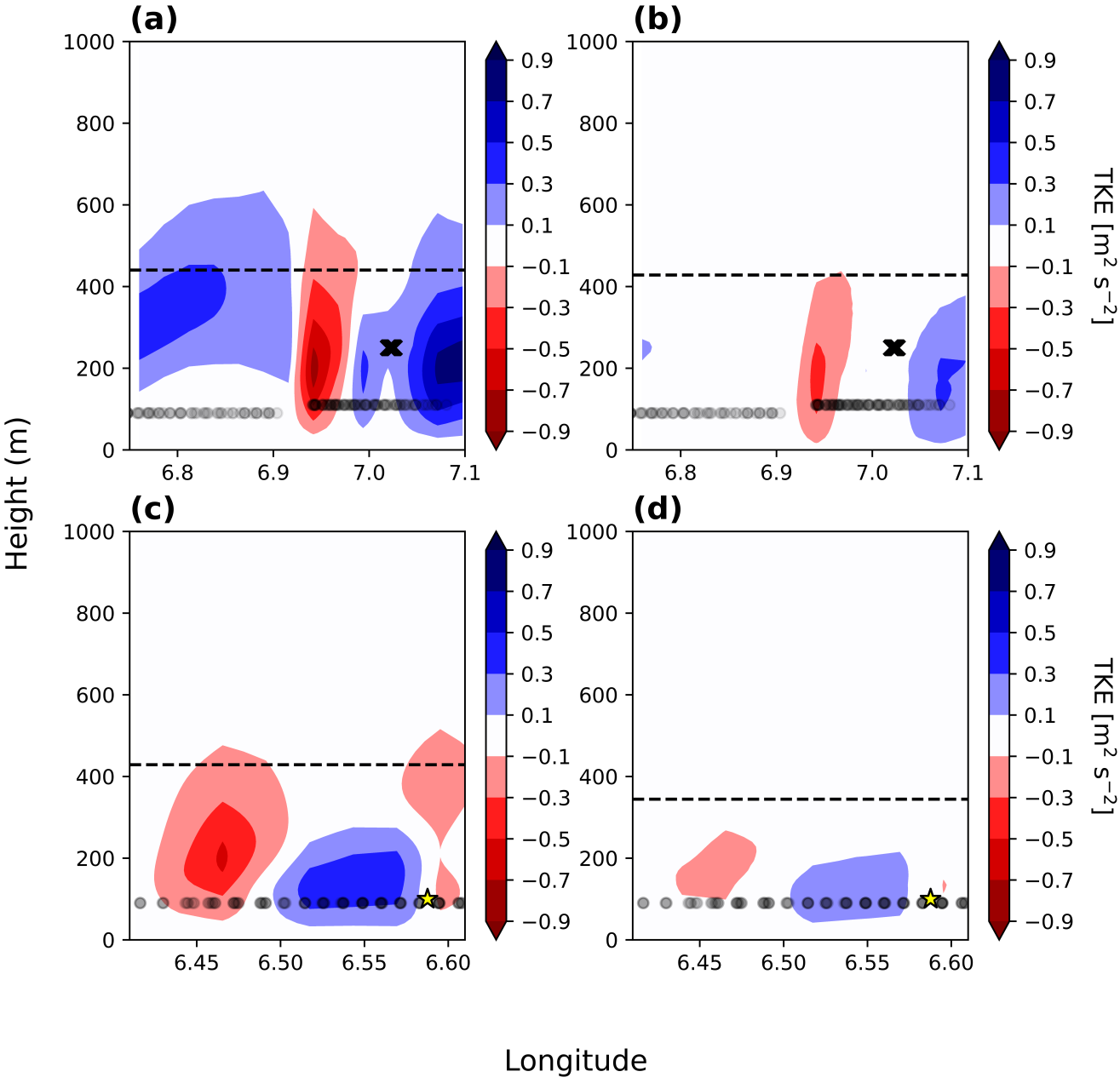

Figure A6. Vertical slices at a constant latitude of 54.03 to mark TKE differences between cases with advection on and cases with advection off. Red indicates that TKE is higher without advection. (a) Aircraft region 3fX_025. (b) Aircraft region mfX_025. (c) FINO1 3fX_025. (d) FINO1 3fX_025. FINO1 cases are averaged over hours 12:00:00–00:00:00, and the aircraft region cases are averaged over 14:10:00–16:10:00. Turbines are marked with black circles, the FINO1 tower is marked with a yellow star, and the first transect path is marked with a solid line.

Author contributions. JKL conceptualized the project and acquired funding and resources for the project. NJA completed the WRF simulations and carried out the formal analysis and investigation, including developing software and carrying out the visualization, with supervision from JKL, TWJ, and AR. NJA and JKL prepared the initial draft. NJA, JKL, TWJ, and AR reviewed and edited the publication.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Wind Energy Science.

Acknowledgements. This work was supported by an agreement with NREL under APUP UGA-0-41026-125. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding was provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office and by the National Offshore Wind Research and Development Consortium under agreement no. CRD-19-16351. The authors acknowledge support from the U.S. Department of Energy (DOE) under DE-EE0009424. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government and the publisher, by accepting the article for publication, acknowledge that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Neither NYSERDA nor OceanTech Services/DNV have reviewed the information contained herein, and the opinions in this report do not necessarily reflect those of any of these parties. Data storage supported by the University of Colorado Boulder 'PetaLibrary'. This work utilized the Alpine high-performance-computing resource at the University of Colorado Boulder. Alpine is jointly funded by the University of Colorado Boulder, the University of Colorado Anschutz, and Colorado State University. A portion of this research was performed using computational resources sponsored by the DOE's Office of Energy Efficiency and Renewable Energy and located at NREL. Author TWJ is grateful for support in part from the U.S. Department of Energy Wind Energy Technologies Office through Contract DE-A05-76RL01830 to Pacific Northwest National Laboratory (PNNL). The U.S. National Science Foundation National Center for Atmospheric Research is a subcontractor to PNNL under Contract 659135. The National Center for Atmospheric Research is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977.

References

Abkar, M. and Porté-Agel, F.: A new wind-farm parameterization for large-scale atmospheric models, Journal of Renewable and Sustainable Energy, 7, 013 121, https://doi.org/10.1063/1.4907600, 2015.

Adams, A. S. and Keith, D. W.: Are global wind power resource estimates overstated?, Environmental Research Letters, 8, 015 021, https://doi.org/10.1088/1748-9326/8/1/015021, 2013.

Ali, K., Schultz, D. M., Revell, A., Stallard, T., and Ouro, P.: Assessment of five wind-farm parameterizations in the Weather Research and Forecasting model: A case study of wind farms in the North Sea, Monthly Weather Review, -1, https://doi.org/10.1175/MWR-D-23-0006.1, 2023.

Archer, C. L., Wu, S., Ma, Y., and Jiménez, P. A.: Two Corrections for Turbulent Kinetic Energy Generated by Wind Farms in the WRF Model, Monthly Weather Review, 148, 4823–4835, https://doi.org/10.1175/MWR-D-20-0097.1, 2020.

Arthur, R. S., Juliano, T. W., Adler, B., Krishnamurthy, R., Lundquist, J. K., Kosović, B., and Jiménez, P. A.: Improved Representation of Horizontal Variability and Turbulence in Mesoscale Simulations of an Extended Cold-Air Pool Event, Journal of Applied Meteorology and Climatology, 61, 685–707, https://doi.org/10.1175/JAMC-D-21-0138.1, 2022.

Arthur, R. S., Rybchuk, A., Juliano, T. W., Rios, G., Wharton, S., Lundquist, J. K., and Fast, J. D.: Evaluating mesoscale model predictions of diurnal speedup events in the Altamont Pass Wind Resource Area of California, Wind Energy Science Discussions, pp. 1–30, https://doi.org/10.5194/wes-2024-137, publisher: Copernicus GmbH, 2024.

Backwell, B., Lee, J., Patel, A., Qiao, L., Nguyen, T., Liang, W., Fang, E., Cheong, J., Tan, W. H., Francisco, A. M., Bui, T. V., Nguyen, T., Fiestas, R., Cox, R., de Lima, J. P., Rabie, H., Ladwa, R., Madan, K., Muchiri, W., Weekes, N., Khinda, N., and Yamamura, J.: Global Offshore Wind Report, 2024.

Baidya Roy, S., Pacala, S. W., and Walko, R. L.: Can large wind farms affect local meteorology?, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2004JD004763, 2004.

Blackadar, A. K.: The vertical distribution of wind and turbulent exchange in a neutral atmosphere, Journal of Geophysical Research (1896-1977), 67, 3095–3102, https://doi.org/10.1029/JZ067i008p03095, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/JZ067i008p03095, 1962.

Bodini, N., Lundquist, J. K., and Optis, M.: Can machine learning improve the model representation of turbulent kinetic energy dissipation rate in the boundary layer for complex terrain?, Geoscientific Model Development, 13, 4271–4285, https://doi.org/10.5194/gmd-13-4271-2020, 2020.

Bärfuss, K., Schulz-Stellenfleth, J., and Lampert, A.: The Impact of Offshore Wind Farms on Sea State Demonstrated by Airborne LiDAR Measurements, Journal of Marine Science and Engineering, 9, 644, https://doi.org/10.3390/jmse9060644, 2021.

Bärfuss, K. B., Schmithüsen, H., and Lampert, A.: Drone-based meteorological observations up to the tropopause – a concept study, Atmospheric Measurement Techniques, 16, 3739–3765, https://doi.org/10.5194/amt-16-3739-2023, publisher: Copernicus GmbH, 2023.

Corsmeier, U., Hankers, R., and Wieser, A.: Airborne turbulence measurements in the lower troposphere onboard the research aircraft Dornier 128-6, D-IBUF, Meteorologische Zeitschrift, pp. 315–329, https://doi.org/10.1127/0941-2948/2001/0010-0315, publisher: Schweizerbart'sche Verlagsbuchhandlung, 2001.

Daewel, U., Akhtar, N., Christiansen, N., and Schrum, C.: Offshore wind farms are projected to impact primary production and bottom water deoxygenation in the North Sea, Communications Earth & Environment, 3, 1–8, https://doi.org/10.1038/s43247-022-00625-0, 2022.

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

Eghdami, M., Barros, A. P., Jiménez, P. A., Juliano, T. W., and Kosovic, B.: Diagnosis of Second-Order Turbulent Properties of the Surface Layer for Three-Dimensional Flow Based on the Mellor–Yamada Model, https://doi.org/10.1175/MWR-D-21-0101.1, 2022.

Fischereit, J., Brown, R., Larsén, X. G., Badger, J., and Hawkes, G.: Review of Mesoscale Wind-Farm Parametrizations and Their Applications, Boundary-Layer Meteorology, 182, 175–224, https://doi.org/10.1007/s10546-021-00652-y, 2022a.

Fischereit, J., Larsén, X. G., and Hahmann, A. N.: Climatic Impacts of Wind-Wave-Wake Interactions in Offshore Wind Farms, Frontiers in Energy Research, 10, https://doi.org/10.3389/fenrg.2022.881459, publisher: Frontiers, 2022b.

Fitch, A. C., Olson, J. B., Lundquist, J. K., Dudhia, J., Gupta, A. K., Michalakes, J., and Barstad, I.: Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model, Monthly Weather Review, 140, 3017–3038, https://doi.org/10.1175/MWR-D-11-00352.1, 2012.

Fitch, A. C., Olson, J. B., and Lundquist, J. K.: Parameterization of Wind Farms in Climate Models, Journal of Climate, 26, 6439–6458, https://doi.org/10.1175/JCLI-D-12-00376.1, 2013.

García-Santiago, O., Hahmann, A. N., Badger, J., and Peña, A.: Evaluation of wind farm parameterizations in the WRF model under different atmospheric stability conditions with high-resolution wake simulations, Wind Energy Science, 9, 963–979, https://doi.org/10.5194/wes-9-963-2024, 2024.

Gaudet, B. J., Medina, G. G., Krishnamurthy, R., Sheridan, L. M., Yang, Z., Newsom, R. K., Pekour, M. S., Gustafson, W. I., and Liu, J.: Assessing Impacts of Waves on Hub-Height Winds off the U.S. West Coast Using Lidar Buoys and Coupled Modeling Approaches, Tech. Rep. PNNL-35856, Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), https://doi.org/10.2172/2337529, 2024.

Hahmann, A. N., Sīle, T., Witha, B., Davis, N. N., Dörenkämper, M., Ezber, Y., García-Bustamante, E., González-Rouco, J. F., Navarro, J., Olsen, B. T., and Söderberg, S.: The making of the New European Wind Atlas – Part 1: Model sensitivity, Geoscientific Model Development, 13, 5053–5078, https://doi.org/10.5194/gmd-13-5053-2020, 2020.

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, 2020.

Hong, S.-Y., Lim, K.-S. S., Lee, Y.-H., Ha, J.-C., Kim, H.-W., Ham, S.-J., and Dudhia, J.: Evaluation of the WRF Double-Moment 6-Class Microphysics Scheme for Precipitating Convection, Advances in Meteorology, 2010, e707 253, https://doi.org/10.1155/2010/707253, 2010.

Jankov, I., Gallus, W. A., Segal, M., Shaw, B., and Koch, S. E.: The Impact of Different WRF Model Physical Parameterizations and Their Interactions on Warm Season MCS Rainfall, Weather and Forecasting, 20, 1048–1060, https://doi.org/10.1175/WAF888.1, 2005.

Juliano, T. W., Kosović, B., Jiménez, P. A., Eghdami, M., Haupt, S. E., and Martilli, A.: "Gray Zone" Simulations Using a Three-Dimensional Planetary Boundary Layer Parameterization in the Weather Research and Forecasting Model, Monthly Weather Review, 150, 1585–1619, https://doi.org/10.1175/MWR-D-21-0164.1, 2022.

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, Journal of Applied Meteorology and Climatology, 43, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kosović, B., Munoz, P. J., Juliano, T. W., Martilli, A., Eghdami, M., Barros, A. P., and Haupt, S. E.: Three-Dimensional Planetary Boundary Layer Parameterization for High-Resolution Mesoscale Simulations, Journal of Physics: Conference Series, 1452, 012 080, https://doi.org/10.1088/1742-6596/1452/1/012080, 2020.

Larsén, X. G. and Fischereit, J.: A case study of wind farm effects using two wake parameterizations in the Weather Research and Forecasting (WRF) model (V3.7.1) in the presence of low-level jets, Geoscientific Model Development, 14, 3141–3158, https://doi.org/10.5194/gmd-14-3141-2021, 2021.

Larsén, X. G., Fischereit, J., Hamzeloo, S., Bärfuss, K., and Lampert, A.: Investigation of wind farm impacts on surface waves using coupled numerical simulations, Renewable Energy, 237, 121 671, https://doi.org/10.1016/j.renene.2024.121671, 2024.

Lee, J. C. Y. and Lundquist, J. K.: Evaluation of the wind farm parameterization in the Weather Research and Forecasting model (version 3.8.1) with meteorological and turbine power data, Geoscientific Model Development, 10, 4229–4244, https://doi.org/10.5194/gmd-10-4229-2017, 2017.

Li, X. and Pu, Z.: Sensitivity of Numerical Simulation of Early Rapid Intensification of Hurricane Emily (2005) to Cloud Microphysical and Planetary Boundary Layer Parameterizations, Monthly Weather Review, 136, 4819–4838, https://doi.org/10.1175/2008MWR2366.1, 2008.

Ma, Y., Archer, C. L., and Vasel-Be-Hagh, A.: The Jensen wind farm parameterization, Wind Energy Science, 7, 2407–2431, https://doi.org/10.5194/wes-7-2407-2022, 2022.

Mangara, R. J., Guo, Z., and Li, S.: Performance of the Wind Farm Parameterization Scheme Coupled with the Weather Research and Forecasting Model under Multiple Resolution Regimes for Simulating an Onshore Wind Farm, Advances in Atmospheric Sciences, 36, 119–132, https://doi.org/10.1007/s00376-018-8028-3, 2019.

Mann, H. B. and Whitney, D. R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, The Annals of Mathematical Statistics, 18, 50–60, https://doi.org/10.1214/aoms/1177730491, 1947.

Mellor, G. L. and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Reviews of Geophysics, 20, 851–875, https://doi.org/10.1029/RG020i004p00851, 1982.

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

Moriarty, P., Bodini, N., Letizia, S., Abraham, A., Ashley, T., Bärfuss, K. B., Barthelmie, R. J., Brewer, A., Brugger, P., Feuerle, T., Frère, A., Goldberger, L., Gottschall, J., Hamilton, N., Herges, T., Hirth, B., Hung, L.-Y. L., Iungo, G. V., Ivanov, H., Kaul, C., Kern, S., Klein, P., Krishnamurthy, R., Lampert, A., Lundquist, J. K., Morris, V. R., Newsom, R., Pekour, M., Pichugina, Y., Porté-Angel, F., Pryor, S. C., Scholbrock, A., Schroeder, J., Shartzer, S., Simley, E., Vöhringer, L., Wharton, S., and Zalkind, D.: Overview of preparation for the American WAKE ExperimeNt (AWAKEN), Journal of Renewable and Sustainable Energy, 16, 053 306, https://doi.org/10.1063/5.0141683, 2024.

Nakanishi, M. and Niino, H.: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer, 87, 895–912, https://doi.org/10.2151/jmsj.87.895, 2009.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD015139, 2011.

Nolan, D. S., Stern, D. P., and Zhang, J. A.: Evaluation of Planetary Boundary Layer Parameterizations in Tropical Cyclones by Comparison of In Situ Observations and High-Resolution Simulations of Hurricane Isabel (2003). Part II: Inner-Core Boundary Layer and Eyewall Structure, Monthly Weather Review, 137, 3675–3698, https://doi.org/10.1175/2009MWR2786.1, 2009.

Olsen, B. T. E., Hahmann, A. N., Alonso-de Linaje, N. G., Žagar, M., and Dörenkämper, M.: Low-level jets in the North and Baltic seas: mesoscale model sensitivity and climatology using WRF V4.2.1, Geoscientific Model Development, 18, 4499–4533, https://doi.org/10.5194/gmd-18-4499-2025, publisher: Copernicus GmbH, 2025.

Olson, J. B., Kenyon, J. S., Angevine, W. A., Brown, J. M., Pagowski, M., and Sušelj, K.: A Description of the MYNN-EDMF Scheme and the Coupling to Other Components in WRF–ARW, https://doi.org/10.25923/N9WM-BE49, 2019.

Optis, M.: Veer_Renewables_WakeMap_Validation_FINO_Towers_2024_05, https://veer.eco/wp-content/uploads/2024/05/Veer_Renewables_WakeMap_Validation_FINO_Towers_2024_05.pdf, 2024.

Optis, M. and Perr-Sauer, J.: The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production, Renewable and Sustainable Energy Reviews, 112, 27–41, https://doi.org/10.1016/j.rser.2019.05.031, 2019.

Optis, M., Bodini, N., Debnath, M., and Doubrawa, P.: Best Practices for the Validation of U.S. Offshore Wind Resource Models, Tech. Rep. NREL/TP–5000-78375, 1755697, MainId:32292, https://doi.org/10.2172/1755697, 2020.

Pan, Y. and Archer, C. L.: A Hybrid Wind-Farm Parametrization for Mesoscale and Climate Models, Boundary-Layer Meteorology, 168, 469–495, https://doi.org/10.1007/s10546-018-0351-9, 2018.

Peña, A., García-Santiago, O., Kosović, B., Mirocha, J. D., and Juliano, T. W.: Can we yet do a fairer and more complete validation of wind farm parametrizations in the WRF model?, Journal of Physics: Conference Series, 2505, 012 024, https://doi.org/10.1088/1742-6596/2505/1/012024, 2023.

Platis, A., Siedersleben, S. K., Bange, J., Lampert, A., Bärfuss, K., Hankers, R., Cañadillas, B., Foreman, R., Schulz-Stellenfleth, J., Djath, B., Neumann, T., and Emeis, S.: First in situ evidence of wakes in the far field behind offshore wind farms, Scientific Reports, 8, 2163, https://doi.org/10.1038/s41598-018-20389-y, 2018.

Pryor, S. C., Shepherd, T. J., Volker, P. J. H., Hahmann, A. N., and Barthelmie, R. J.: "Wind Theft" from Onshore Wind Turbine Arrays: Sensitivity to Wind Farm Parameterization and Resolution, Journal of Applied Meteorology and Climatology, 59, 153–174, https://doi.org/10.1175/JAMC-D-19-0235.1, 2020.

Quint, D., Lundquist, J. K., Bodini, N., and Rosencrans, D.: Meteorological Impacts of Offshore Wind Turbines as Simulated in the Weather Research and Forecasting Model, Wind Energy Science Discussions, pp. 1–34, https://doi.org/10.5194/wes-2024-53, publisher: Copernicus GmbH, 2024.

Raghukumar, K., Chartrand, C., Chang, G., Cheung, L., and Roberts, J.: Effect of Floating Offshore Wind Turbines on Atmospheric Circulation in California, Frontiers in Energy Research, 10, https://doi.org/10.3389/fenrg.2022.863995, 2022.

Raghukumar, K., Nelson, T., Jacox, M., Chartrand, C., Fiechter, J., Chang, G., Cheung, L., and Roberts, J.: Projected cross-shore changes in upwelling induced by offshore wind farm development along the California coast, Communications Earth & Environment, 4, 1–12, https://doi.org/10.1038/s43247-023-00780-y, 2023.

Rajewski, D. A., Takle, E. S., Prueger, J. H., and Doorenbos, R. K.: Toward understanding the physical link between turbines and microclimate impacts from in situ measurements in a large wind farm, Journal of Geophysical Research: Atmospheres, 121, 13,392–13,414, https://doi.org/10.1002/2016JD025297, 2016.

Ramdas, A., Garcia, N., and Cuturi, M.: On Wasserstein Two Sample Testing and Related Families of Nonparametric Tests, https://doi.org/10.48550/arXiv.1509.02237, arXiv:1509.02237 [math], 2015.

Redfern, S., Olson, J. B., Lundquist, J. K., and Clack, C. T. M.: Incorporation of the Rotor-Equivalent Wind Speed into the Weather Research and Forecasting Model's Wind Farm Parameterization, Monthly Weather Review, 147, 1029–1046, https://doi.org/10.1175/MWR-D-18-0194.1, 2019.

Rybchuk, A., Juliano, T. W., Lundquist, J. K., Rosencrans, D., Bodini, N., and Optis, M.: The sensitivity of the Fitch wind farm parameterization to a three-dimensional planetary boundary layer scheme, Wind Energy Science, 7, 2085–2098, https://doi.org/10.5194/wes-7-2085-2022, 2022.

Sanchez Gomez, M., Lundquist, J. K., Mirocha, J. D., and Arthur, R. S.: Investigating the physical mechanisms that modify wind plant blockage in stable boundary layers, Wind Energy Science, 8, 1049–1069, https://doi.org/10.5194/wes-8-1049-2023, 2023.

Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Modelling, 9, 347–404, https://doi.org/10.1016/j.ocemod.2004.08.002, 2005.

Shepherd, T. J., Barthelmie, R. J., and Pryor, S. C.: Sensitivity of Wind Turbine Array Downstream Effects to the Parameterization Used in WRF, Journal of Applied Meteorology and Climatology, 59, 333–361, https://doi.org/10.1175/JAMC-D-19-0135.1, 2020.

Sheridan, L. M., Krishnamurthy, R., Gustafson Jr., W. I., Liu, Y., Gaudet, B. J., Bodini, N., Newsom, R. K., and Pekour, M.: Offshore low-level jet observations and model representation using lidar buoy data off the California coast, Wind Energy Science, 9, 741–758, https://doi.org/10.5194/wes-9-741-2024, 2024.

Shin, H. H. and Hong, S.-Y.: Intercomparison of Planetary Boundary-Layer Parametrizations in the WRF Model for a Single Day from CASES-99, Boundary-Layer Meteorology, 139, 261–281, https://doi.org/10.1007/s10546-010-9583-z, 2011.

Siedersleben, S. K., Lundquist, J. K., Platis, A., Bange, J., Bärfuss, K., Lampert, A., Cañadillas, B., Neumann, T., and Emeis, S.: Micrometeorological impacts of offshore wind farms as seen in observations and simulations, Environmental Research Letters, 13, 124 012, https://doi.org/10.1088/1748-9326/aaea0b, 2018a.

Siedersleben, S. K., Platis, A., Lundquist, J. K., Lampert, A., Bärfuss, K., Cañadillas, B., Djath, B., Schulz-Stellenfleth, J., Bange, J., Neumann, T., and Emeis, S.: Evaluation of a Wind Farm Parametrization for Mesoscale Atmospheric Flow Models with Aircraft Measurements, Meteorologische Zeitschrift, pp. 401–415, https://doi.org/10.1127/metz/2018/0900, 2018b.

Siedersleben, S. K., Platis, A., Lundquist, J. K., Djath, B., Lampert, A., Bärfuss, K., Cañadillas, B., Schulz-Stellenfleth, J., Bange, J., Neumann, T., and Emeis, S.: Turbulent kinetic energy over large offshore wind farms observed and simulated by the mesoscale model WRF (3.8.1), Geoscientific Model Development, 13, 249–268, https://doi.org/10.5194/gmd-13-249-2020, 2020.

Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Liu, Z., Berner, J., Wang, W., Powers, G., Duda, G., Barker, D., and Huang, X.-y.: A Description of the Advanced Research WRF Model Version 4.3, https://doi.org/10.5065/1dfh-6p97, 2021.

Smagorinsky, J.: GENERAL CIRCULATION EXPERIMENTS WITH THE PRIMITIVE EQUATIONS, https://journals.ametsoc.org/view/journals/mwre/91/3/1520-0493_1963_091_0099_gcewtp_2_3_co_2.xml, section: Monthly Weather Review, 1963.

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

Tomaszewski, J. M. and Lundquist, J. K.: Simulated wind farm wake sensitivity to configuration choices in the Weather Research and Forecasting model version 3.8.1, Geoscientific Model Development, 13, 2645–2662, https://doi.org/10.5194/gmd-13-2645-2020, 2020.

Vanderwende, B. J., Kosović, B., Lundquist, J. K., and Mirocha, J. D.: Simulating effects of a wind-turbine array using LES and RANS, Journal of Advances in Modeling Earth Systems, 8, 1376–1390, https://doi.org/10.1002/2016MS000652, 2016.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat,

I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Volker, P. J. H., Badger, J., Hahmann, A. N., and Ott, S.: The Explicit Wake Parametrisation V1.0: a wind farm parametrisation in the mesoscale model WRF, Geoscientific Model Development, 8, 3715–3731, https://doi.org/10.5194/gmd-8-3715-2015, 2015.

Wadler, J. B., Nolan, D. S., Zhang, J. A., Shay, L. K., Olson, J. B., and Cione, J. J.: The Effect of Advection on the Three Dimensional Distribution of Turbulent Kinetic Energy and Its Generation in Idealized Tropical Cyclone Simulations, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 230, https://doi.org/10.1029/2022MS003230, 2023.

WFIP3: Wind Forecast Improvement Project-3, https://www.pnnl.gov/projects/wind-forecast-improvement-project-3.

Wyngaard, J. C.: Toward Numerical Modeling in the "Terra Incognita", Journal of the Atmospheric Sciences, 61, 1816–1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, 2004.

Yang, B., Qian, Y., Berg, L. K., Ma, P.-L., Wharton, S., Bulaevskaya, V., Yan, H., Hou, Z., and Shaw, W. J.: Sensitivity of Turbine-Height Wind Speeds to Parameters in Planetary Boundary-Layer and Surface-Layer Schemes in the Weather Research and Forecasting Model, Boundary-Layer Meteorology, 162, 117–142, https://doi.org/10.1007/s10546-016-0185-2, 2017.

Zhang, D.-L. and Zheng, W.-Z.: Diurnal Cycles of Surface Winds and Temperatures as Simulated by Five Boundary Layer Parameterizations, Journal of Applied Meteorology and Climatology, 43, 157–169, https://doi.org/10.1175/1520-0450(2004)043<0157:DCOSWA>2.0.CO;2, 2004.

Zhang, X., Bao, J.-W., Chen, B., and Grell, E. D.: A Three-Dimensional Scale-Adaptive Turbulent Kinetic Energy Scheme in the WRF-ARW Model, https://doi.org/10.1175/MWR-D-17-0356.1, section: Monthly Weather Review, 2018.