One-year long turbulence measurements and modeling using large-eddy simulation domains in the Weather Research and Forecasting model

Alfredo Peña (aldi@dtu.dk)
Technical University of Denmark

Jeffrey D. Mirocha
Lawrence Livermore National Laboratory

Research Article

Keywords: Atmospheric turbulence, Large-eddy simulation, Mesoscale modeling, Site conditions, WRF

Posted Date: January 23rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2440654/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.
One-year long turbulence measurements and modeling using large-eddy simulation domains in the Weather Research and Forecasting model

Alfredo Peña · Jeffrey D. Mirocha

Abstract We present an intercomparison of a full year of turbulence measurements and simulations at Østerild, a site in northern Denmark with relatively flat terrain and high surface roughness, where a high-quality tall meteorological mast is deployed. Both sonic and cup anemometers are mounted on booms on the mast from 7 up to 244 m, thus covering the range of heights in which modern wind turbines operate. The simulations were performed using the Weather Research and Forecasting model in a multiscale setup, with large-eddy simulations (LESs) nested one-way within mesoscale simulations. The mesoscale domains thus simulated the evolving weather, while the two innermost domains used a LES closure, permitting the largest scales of turbulence to be explicitly resolved. For a selected day, we show that the simulated turbulence is accurately resolved within the innermost LES domain, and agrees well with the observations at all vertical levels. For the full year, we show that the innermost domain accurately reproduces mean wind speed, direction, and turbulence levels, whereas the mesoscale simulations have difficulties matching the frequency of occurrence of both low and high turbulence ranges when compared to the observations. The largest differences between the simulated turbulence from the innermost domain and the observations are found under low wind speed conditions close to the surface, particularly during nighttime where the simulated mean wind and turbulence levels are higher and lower, respectively, than the observations, potentially due to the inability of the LES mesh to resolve the very small scales of turbulence and associated momentum transport that those features support.

Keywords Atmospheric turbulence · Large-eddy simulation · Mesoscale modeling · Site conditions · WRF
1 Introduction

The use of atmospheric models to generate time series of the wind velocity components at different vertical levels, and over long time periods (e.g., for entire full years) is becoming increasingly common with the advent of affordable computation. These time series outputs are particularly useful for wind resource and energy yield assessments (Fernández-González et al., 2018; Dörenkämper et al., 2020). Long-term time series are typically represented by the instantaneous output of the atmospheric model, e.g., every 30 min or 1 h. This is because such outputs are computationally much less expensive and nearly as accurate, e.g., for mean wind speed and wind power density purposes, as those derived from constructing time series based on time-averages of model output over periods of, e.g., 10 or 30 min, as commonly performed with meteorological observations (Gross et al., 2020). Further, most of these modeling systems are based on mesoscale simulations, an approach that can accurately capture the large-scale weather, but constrains the outputs to spatial and temporal scales of a several hundreds of meters to a few kilometers, and to tens of minutes, respectively, in the best of cases.

However, wind turbines as well as other structures such as buildings and bridges are designed to withstand specific site conditions, which are defined in the standards (e.g., IEC, 2019). An important part of these conditions involves measures of the levels of turbulence, which include processes that occur typically at high spatial resolutions (hundreds of meters at most) and at time scales of minutes to seconds. Turbulence has to be assessed as part of the site conditions because it directly affects the loads and performance of wind turbines (Conti et al., 2021). However, in mesoscale simulations, turbulence cannot be explicitly resolved, and is therefore parameterized as a subgrid-scale (SGS) process. As such, turbulence derived from mesoscale simulations does not reliably provide the information required by the standards. Long-term time series of simulated turbulence measures could therefore offer to, e.g., wind farm project developers, important information useful to select suitable machines that can withstand the turbulence conditions of the site. Further, they could also aid, e.g., turbine manufacturers, in the design and optimization of wind turbines for particular turbulence conditions, sites and other operational considerations.

The Weather Research and Forecasting (WRF) model (Skamarock et al., 2019) is a commonly used atmospheric model, which is typically utilized to simulate mesoscale weather. This is achieved by downscaling the output from global-scale numerical weather prediction models to spatial scales of tens to a few kilometers, at which relevant smaller-scale weather features can be represented. However, mesoscale domains still use meshes that are too coarse to capture atmospheric turbulence directly. Instead, mesoscale simulations rely upon one-dimensional planetary boundary layer (PBL) schemes to parameterize the turbulent fluxes of heat, momentum, and other constituents, in the vertical direction only, within the atmospheric boundary layer. While some mesoscale PBL schemes include a variable to represent SGS turbulence ki-
One-year long turbulence measurements and modeling using WRF-LES

To enable the direct simulation of turbulence motions, the WRF model (and other mesoscale simulation codes) support successive grid nesting, enabling refinement of the mesh to scales capable of resolving the important scales of atmospheric turbulence directly. For this purpose, a number of additional SGS models designed for LES have been implemented into the WRF model (Kosović, 1997; Mirocha et al., 2010). These models allow for the explicit representation of the energetically important scales of atmospheric turbulence, while accounting for the effects of unresolved small-scale fluctuations on those resolved-scale motions. Therefore, using PBL schemes on mesoscale domains, and switching to LES SGS models once the mesh spacing is sufficiently fine to resolve turbulence motions explicitly, provides a methodology for the explicit simulation of turbulence motions occurring within mesoscale simulations, as well as for additional improvements in fidelity from additional fine-scale forcing, such as terrain or surface cover variability. This approach, commonly referred to as WRF-LES when using the WRF model, has been applied previously to wind energy case studies involving a diurnal cycle over simple terrain (Muñoz-Esparza et al., 2017), stable and convective conditions over complex terrain (Wise et al., 2022), and during a mesoscale frontal passage (Arthur et al., 2020), in each case showing very good agreement with observations in both mean and turbulence quantities. WRF-LES was also used to study the development of an atmospheric hydraulic jump and mountain waves, which were observed with long-range scanning lidars (Peña and Santos, 2021). Idealized simulations using the WRF model also employed a LES approach to study atmospheric turbulence under different atmospheric stability conditions (Peña et al., 2021) and to assess the fidelity of mesoscale wind-farm parametrizations by comparison with high resolution LES of wind turbine wakes (Peña et al., 2022).

Here, we show results of analysis of a full year of atmospheric simulations performed with a WRF-LES-based modeling system, with the goal of describing both the mean wind as well as the turbulence characteristics, at a specific site. The site is Østerild, located in northern Jutland, Denmark, which features relatively flat terrain, but with heterogeneous and seasonally varying surface roughness, and where high-quality observations of wind and turbulence are available from instruments deployed at several vertical levels on a 250-m tall meteorological mast, thus covering the range of heights within which modern wind turbines operate. From the author’s knowledge, this is the first time that such an analysis is documented with these details for this relatively long period. In Sect. 2, we introduce the Østerild site and the measurements from the mast. Section 3 provides the details of the numerical setup and the simulations. In Sect. 4, we provide the results both for a selected day (Sect. 4.1) and for the full year (Sect. 4.2). Section 4.3 illustrates results of a sensitivity analysis on different alternatives to instigate turbulence in the boundaries of the LES domains. The two last sections contain a discussion and conclusions of the study.
2 Site and measurements

Measurements from the entire 2017 were collected from instruments installed on a mast, which is deployed at the Østerild test station for large wind turbines in northern Denmark. Østerild is located on a land area between the North Sea and the Limfjorden waters (see Fig. 1). At Østerild, the terrain is flat and heterogeneous from a combination of agricultural and cropland, forest, and urban areas. North of the mast, there were only seven test stands for wind turbines; nowadays there are two more test stands south of the mast.

Fig. 1 The Østerild test station and surrounding area in northern Jutland, Denmark on a digital surface model (UTM WGS84) from 2017. Turbine locations are shown in black circles and the meteorological mast in the red rectangle

We use measurements from Risø P2546A cup anemometers at 10, 40, 70, 106, 140, 178, 210, and 244 m above the ground (agl; to which hereafter all heights are referred) and Metek USA-1 sonic anemometers at 7, 37, 103, 175, and 241 m. We also use measurements from the reference wind vanes at 40 and 244 m. When sonic anemometer measurements are utilized, we remove the instrument’s default flow distortion 2D correction and apply the 3D correction suggested by Metek GmbH (2004). The latter corrects the velocity components so that their spectra show the inertial subrange ratio of 4/3 as expected from Kolmogorov’s hypothesis (Peña et al., 2019). Other details about the area and the measurements are provided in Peña (2019).

3 Numerical simulations

We perform real-time multiscale atmospheric simulations using the WRF model version 4.1.2. We use four telescopic one-way nested domains, which are all centered at the position of the Østerild meteorological mast (see Fig. 2a).
Starting from the outermost (d01) down to the innermost (d04) domain, the horizontal grid spacing of the domains ($\Delta x$) is 6250, 1250, 250, and 50 m. All domains have nearly the same amount of grid points in both horizontal directions ($N_x \times N_y = 201 \times 201$). All domains have 61 vertical levels with the model top boundary at 5000 Pa. The first 20 vertical levels are within the first kilometer from the model bottom; 10 vertical levels are within the first 250 m covering the mast (at ≈11, 32, 48, 65, 81, 100, 125, 154, 186, and 223 m). Table 1 presents a summary of the model domains’ specifications.

![Diagram of nested domains](image)

**Fig. 2** (a) Telescopic nested domains d01–4 (red solid lines) used for the WRF model simulations. All domains are centered on the Østerild meteorological mast in Denmark. (b) Land cover of the innermost domain (d04) from the US Geological Survey classification with the Østerild mast shown in white.

We use the ERA5 reanalysis at 0.3° (C3S, 2018) to initialize and force the model outermost domain. Additional boundary conditions come from the Operational Sea Surface Temperature and Sea Ice Analysis (Donlon et al., 2012). The WRF-default land cover and terrain elevation datasets are used as inputs: Moderate Resolution Imaging Spectroradiometer (MODIS) land
Table 1 Model setup for the numerical simulations

<table>
<thead>
<tr>
<th>domain</th>
<th>$\Delta x$ [m]</th>
<th>nest ratio</th>
<th>$N_x \times N_y$</th>
<th>$\Delta t$ [s]</th>
<th>PBL scheme/turbulence closure</th>
</tr>
</thead>
<tbody>
<tr>
<td>d01</td>
<td>6250</td>
<td>-</td>
<td>202 $\times$ 202</td>
<td>37.5</td>
<td>MYNN2</td>
</tr>
<tr>
<td>d02</td>
<td>1250</td>
<td>5</td>
<td>201 $\times$ 201</td>
<td>7.5</td>
<td>MYNN2</td>
</tr>
<tr>
<td>d03</td>
<td>250</td>
<td>5</td>
<td>201 $\times$ 201</td>
<td>1.5</td>
<td>TKE 1.5</td>
</tr>
<tr>
<td>d04</td>
<td>50</td>
<td>5</td>
<td>201 $\times$ 201</td>
<td>0.3</td>
<td>TKE 1.5</td>
</tr>
</tbody>
</table>

Table 1 Model setup for the numerical simulations

The domain dimensions are chosen to cover classification and global multiresolution terrain elevation data 2010 at 30 arcsec (Danielson and Gesch, 2011). Figure 2b illustrates the land use of the innermost domain together with the position of the Østerild mast. The WRF-default land cover depicts Østerild as mixed forest surrounded by cropland. Although details are missing, the land use description used in the model fits relatively well that of the area and this will be reflected when analyzing the simulations below.

The timestep ($\Delta t$) used for the outermost domain is 37.5 s and with a 1:5 timestep ratio between domains, that of the innermost domain is 0.3 s. The following physical parameterizations are selected for the mesoscale domains: the single moment 5-class scheme for cloud micro-physics (Hong et al., 2004), the Kain–Fritsch scheme for cumulus convection (Kain, 2004) at a 5-min frequency, and only applied to the outermost domain, and the MYNN level-2.5 (MYNN2) planetary boundary layer (PBL) scheme (Nakanishi and Niino, 2009) (PBL domains hereafter). The innermost domains are run in LES mode using the SGS model of Deardorff (1980), which uses a prognostic equation for the SGS turbulent kinetic energy (TKE). All domains employ the MYNN surface-layer scheme (Nakanishi and Niino, 2009) and the Unified Noah Land Surface model (Tewari et al., 2004). For shortwave and longwave radiation, all domains use the RRTMG scheme (Iacono et al., 2008) at a 6-min frequency.

Simulations are performed for the entire year of 2017, and simultaneously for all domains through parallel 10-day long runs, each with a spinup of 12 h. From 37 of these 10-day runs, two experience numerical instabilities and require a reduction of the model timestep. In these cases, $\Delta t = 20$ s is used for the outermost domain. Spectral nudging towards the forcing reanalysis is utilized for the two horizontal wind components and temperature on the outermost domain with a constant of 0.0003 s$^{-1}$. No upper level damping was used. Two-dimensional deformation and second-order diffusion are used for the PBL domains, while three-dimensional turbulence mixing in physical space is used for the LES domains. Sixth-order positive definite numerical diffusion and positive definite advection of moisture and scalars are used for all domains.

Due to the high computational cost of processing these multiple-domain simulations, we extract the output of the simulations from each of the domains at the grid point closest to the Østerild mast. For domains other than the innermost LES domain (d04), the output is instantaneous and produced hourly, except for one particular day that we study in detail, where the instantaneous
output is every 10 min. For the innermost LES domain, instantaneous output from the simulation is produced every 12 s.

When performing multiscale simulations with the WRF model, three-dimensional turbulence within the nested LES domains can be slow to develop, due to the lack of resolved turbulence features contained within the bounding-domain solution, which contains insufficient mesh spacing, and uses a one-dimensional PBL scheme that is not designed to resolve turbulence. The formation of turbulence is therefore accelerated by applying the stochastic cell perturbation method (CPM) (Muñoz-Esparza et al., 2015; Muñoz-Esparza and Kosović, 2018) at the inflow boundaries of the LES domains. The CPM is configured in the standard manner, using perturbation zones of three perturbation cells, each consisting of eight by one grid cells in the horizontal and vertical directions, respectively. The magnitude of the perturbations are chosen to satisfy the optimal turbulent Eckert number (0.2), which spurs the development of hairpin-like vortices that quickly grow to form a realistic three-dimensional turbulence field (Muñoz-Esparza et al., 2015).

Since our simulation spans an entire year, the standard CPM methodology of Muñoz-Esparza and Kosović (2018), which requires performing a mesoscale simulation first to determine the CPM parameters, and then reading those parameters in during a subsequent multiscale simulation with the nested LESs, is impractical. Instead, the CPM used herein runs concurrently within the multiscale simulation, with the required control parameters passed down from the mesoscale domain and its PBL parameterization on d02 during execution. One key parameter is the PBL height, which is used to specify both the depth to apply the perturbations, as well as the perturbation amplitude. The perturbation amplitude, which is computed from the turbulent Eckert number, requires the geostrophic wind speed, the value of which is approximated by the average horizontal wind speed computed at a vertical level that is 10% higher than the maximum PBL height occurring within the footprint of the LES domain. The perturbations are then applied up to 80% of the PBL height.

Another parameter that is adjusted in real time is which lateral boundaries to perturb, determined by the direction of the mean geostrophic wind, with perturbations applied along boundaries with flow oriented into the domain. Finally, the time interval over which to reapply the perturbations is determined using the advective timescale of the perturbation zone, here taken to be 85% of the 24 grid cells perturbed, with the slight reduction accounting for the relaxation zone of 5 grid cells adjacent to each lateral boundary, over which the flow variables within a nested domain are linearly blended with their respective parent-domain values. We assume that this relaxation shortens the effective perturbation zone by one half of the width of the relaxation zone.

The advective time scale is then computed as the time to span the adjusted perturbation zone using the average wind speed at the height of the cells being perturbed. An additional adjustment is made to lengthen the timescale by up to a factor of the square root of two, based on the angle of the flow relative to the computational mesh.
4 Results

For a selected day, which we present in Sect. 4.1, we resample the 20-Hz sonic anemometer data to 12 s for direct comparison with the 12-s output of the innermost LES domain (d04). We also compute 10-min statistics based on the 20-Hz sonic anemometer data. We compare these statistics with the simulations’ 10-min instantaneous output of all domains except for the innermost LES domain (d04). For the latter domain, we use both the 12-s output directly and also compute the corresponding 10-min statistics from the 12-s output. For the entire full year intercomparison (Sect. 4.2), we compute 10-min statistics based on the 1-Hz cup anemometer data, 10-min statistics from the 12-s output of the simulation of the innermost LES domain, and use the instantaneous hourly output of all other domains. All time references are based on Danish standard time (UTC+1).

Cup anemometers measure the horizontal velocity magnitude $U$, which can also be obtained from the two horizontal velocity components measured from the sonic anemometers. On a Cartesian coordinate system:

$$U = (u_x^2 + u_y^2)^{1/2}, \quad (1)$$

where $u_x$ and $u_y$ can be, e.g., the zonal and meridional components of the wind. The cup anemometer velocity variance $\langle U'U' \rangle$, where the prime denotes fluctuation from the mean, which is here denoted within $\langle \rangle$ and represented as a time average within a 10-min period in our analysis, but more precisely its root square, is used to estimate the turbulence intensity (TI) as

$$\text{TI} = \frac{\sqrt{\langle U'U' \rangle}}{U}, \quad (2)$$

which is, e.g., a key parameter for site selection of suitable wind turbines.

The 20-Hz sonic anemometer observations and the 12-s output from the simulations of the innermost LES domain are similarly analyzed. Within each 10-min interval and for each vertical level, the sonic anemometer velocity components are rotated so that the $u$-velocity component (the along-wind component) becomes aligned with the 10-min mean wind, i.e., $\langle u \rangle = \langle U \rangle$. Thus, the $v$-velocity component becomes the cross-wind component with $w$ being the vertical velocity component. In this way, the variance of the horizontal velocity measured by a cup anemometer is very close to that estimated from the along-wind component, which can be derived from a sonic anemometer, i.e., $\langle U'U' \rangle \approx \langle u'u' \rangle$.

The WRF model outputs instantaneous zonal, meridional, and vertical velocity components, the SGS TKE ($e_{\text{sgs}}$), and the deviatoric part of the sub-grid stresses $\tau_{ij}^{\text{dev}}$. From the 12-s output, we first estimate the magnitude of the horizontal velocity, the flow direction, and the SGS stresses. The latter are

$$\tau_{ij}^{\text{sgs}} = (2/3)\delta_{ij}e_{\text{sgs}} + \tau_{ij}^{\text{dev}}, \quad (3)$$
without implicit summation. Then, within each 10-min interval, we compute both the mean horizontal velocity and mean wind direction (from the mean zonal and meridional simulated winds), and the resolved stresses

$$\tau_{ij}^{\text{res}} = -\rho(u_i'u_j'),$$

(4)

where $\rho$ is the air density. The total stress components are computed as

$$\tau_{ij}^{\text{tot}} = \tau_{ij}^{\text{res}} + \langle \tau_{ij}^{\text{sgs}} \rangle,$$

(5)

which are then rotated to the mean wind as $\tau_{ij} = a_{im}a_{jn}\tau_{mn}^{\text{tot}}$, where $a_{ij}$ are the components of the rotation matrix. This procedure is performed at each simulated vertical level.

4.1 Selected day

We select a day, June 12, 2017, to explore the capabilities of WRF-LES in detail. This particular day is selected because it shows a wide range of wind speeds within the operational envelope of wind turbines and winds are from the predominant wind direction (west) most of the day (see Figs. 3 and 4). We show first for three vertical levels covering the extension of the mast, the variability of the 12-s WRF-LES output for the horizontal velocity and that of the resampled observations from the sonic anemometers (Fig. 3).

![Fig. 3](image)

**Fig. 3** Time series of horizontal velocity from 12-s resampled sonic anemometer observations (OBS) and 12-s output from the innermost LES domain (d04) at three different observational heights: 7, 37, and 241 m in panels (a), (b), and (c), respectively, at Østerild. For the simulated winds, the closest vertical levels to the observations are at $\approx 6$, 40, and 244 m.
As illustrated, both the sonic anemometer and WRF-LES high-frequency time series follow the same trends at these three vertical levels. Differences between the simulations and observations are more clearly visible the closer to the ground. At 7 m, the simulated winds tend to be larger than the observations during the morning and the opposite occurs during the afternoon. This behavior becomes less discernible with increasing distance above the ground, and by 241 m is no longer detectable. The variability in the fluctuations of the WRF-LES output similarly shows improving agreement with the observations with increasing distance above the ground. One can notice that the 6-m simulated winds exhibit less fluctuations than those at higher levels: the daily variance of the 6-m simulated wind is $2.74 \text{ m}^2 \text{s}^{-2}$, whereas that of the 40- and 244-m winds are $7.24$ and $5.65 \text{ m}^2 \text{s}^{-2}$, respectively. Also, all simulated winds fluctuate less in the early morning hours than in the rest of the day: at 6-m, the variance of the simulated wind is $0.18 \text{ m}^2 \text{s}^{-2}$ within 00:00–03:00 and 1.65 $\text{m}^2 \text{s}^{-2}$ within 12:00–15:00, whereas the variance of the observed winds is 0.72 and 5.04 $\text{m}^2 \text{s}^{-2}$ within the same two periods, respectively. At 37 m and within 12:00–15:00, the variance of the simulated and observed wind is 4.58 and 6.61 $\text{m}^2 \text{s}^{-2}$, respectively, whereas within the same period but at 241 m, it is 4.60 and 3.31 $\text{m}^2 \text{s}^{-2}$, respectively. Note that this particular day belongs to one of the 10-day simulations used for the year-long analysis; this particular simulation starts on June 10, 2017.

In Fig. 4, we notice the effect of the fluctuating simulated horizontal velocity components on the high-frequency time series of wind direction, the simulations following closely the trend of the sonic anemometer observations for two vertical levels. As expected from the findings in Fig. 3, the simulations capture more closely the variability of the observed winds the farther from the ground. Although there is reduced variability in the simulations compared to the observations, specially in the early morning hours, for this particular day, the simulations capture the large vertical wind veer between the 241- and 37-m sonic anemometers: at 01:30 the simulated instantaneous wind veer and the sonic-derived mean wind veer are $14.52^\circ$ and $13.68^\circ$, respectively.

Figure 5 also shows time series, but at a 10-min frequency, of the winds for the selected day but we include the instantaneous output of the two mesoscale domains (d01 and d02) and the other LES domain (d03). Also for both wind speed and direction, we show the 10-min mean from the output of the innermost LES domain (d04) and the 10-min mean from the cup anemometer observations. Here, we focus on the 110-m height, since this is a common hub height of modern and large wind turbines.

All domains follow the observed winds closely (Fig. 5a), although the two outermost mesoscale domains (d01 and d02) slightly overpredict the wind speed during the early morning. As expected, the mesoscale winds do not capture the observed variability in the 10-min sonic anemometer means. However, the variability of the 10-min mean winds of the two LES domains seems to be larger than that of the observations, particularly from 06:00 onwards. Within the period 00:00–03:00, the variance of the 10-min mean wind measured by the cup anemometer at 106 m is 0.26 $\text{m}^2 \text{s}^{-2}$, whereas that of the
Fig. 4 Similar to Fig. 3 but for the wind direction at two different observational heights: 37 and 241 m in panels (a) and (b), respectively. For the simulated winds, the closest vertical levels to the observations are at ≈40, and 244 m

Fig. 5 Time series of (a) horizontal velocity at ≈110 m and (b) wind direction at ≈40 m from 10-min means of cup anemometer and wind vane observations (OBS in markers) at Østerild. 10-min instantaneous output from the three outermost domains (d01–d03) and 10-min means from the innermost LES domain (d04) are shown in different lines
10-min instantaneous winds from d03 and 10-min mean winds from d04 is 0.15 and 0.27 m² s⁻², respectively, at ≈113 m. Within the period 12:00–15:00, the equivalent variances are 0.69, 5.15, and 1.85 m² s⁻² for the observations, d03, and d04, respectively; the d03 value is clearly dominated by the decreasing trend within that 3-h period but a detrended 3-h time series also shows a high variance (3.39 m² s⁻²).

For the wind direction, we can see all domains’ output closely following the observations from the vane at 40 m (Fig. 5b). The variability in the observed direction (standard deviation of 5.13°) is interestingly close to that of the output of the LES domains, 5.23° and 5.11°, for d04 and d03, respectively. Note that the 10-min direction values of d04 are computed from the 10-min mean zonal and meridional winds.

The selected day also presents some other interesting features (not shown). During the early morning hours (00:00–06:00), the heat flux measured by the sonic anemometer at 37 m is less than −0.03 K m s⁻¹, increasing up to 0.31 K m s⁻¹ at 12:00 and slowly reaching values below zero from 19:00 onwards, which shows the wide range of atmospheric stability conditions. At the closest to observed vertical level (≈40 m), the total simulated sensible heat flux (resolved plus SGS part) from the d04 output agrees well with that of the sonic observations: the mean value within 00:00–03:00 is −0.041 K m s⁻¹. However, the heat flux at this level starts to increase in the simulations much earlier than in the observations, turning positive already at 04:30. It becomes negative and follows well the observed values from 19:00 onwards. Also, during the first five hours, both simulations and observations of the along-wind variance at ≈40 m are quite close and low. At 07:30 and ≈40 m, the simulated sensible heat flux is 0.15 K m s⁻¹, whereas the observed value is nearly zero. We extract output from observations and simulations at 07:30 and Fig. 6 shows vertical profiles of the horizontal wind speed, wind direction and along-wind variance from both the sonic and cup anemometers, and the output of the simulations at Østerild from the different domains.

As illustrated, at this particular time, the wind speeds are strongly overestimated (by more than 5 m s⁻¹ at some vertical levels) by the simulations, in particular those from the LES domains (d03 and d04). The wind direction is well predicted by all domains’ output, with that from the LES d03 domain showing the largest discrepancy from the other domains. Interestingly, when compared to the observations, the along-wind variance is well predicted by the output of the innermost LES domain (d04) both in its value and the behavior of the vertical profile: both observations and simulations show increasing values up to about 70 m, both decreasing slowly upwards. One explanation for the discrepancy of the vertical wind profiles from d03 is the presence of spurious under-resolved convective roll structures that occur during the selected analysis period as a result of its relatively coarse resolution relative to the forcing. These structures can also explain the slightly greater than observed variance within d04, resulting from their advection into and subsequent interactions with the perturbations and evolving turbulence features within d04.
We also select another time (20:00), but at this hour both observed and simulated sensible heat fluxes are closer than in the previous case. At $\approx 40$ m, the simulated sensible heat flux is $-0.016$ K m s$^{-1}$, whereas the observed value is $-0.009$ K m s$^{-1}$. As illustrated in Fig. 7, at this time the vertical profiles of simulated and observed wind speed are in good agreement, both showing a behavior typical of a stable atmosphere. Simulated directions are systematically biased $\approx 10^\circ$ from the observations. However, also in this case, the along-wind variance is well predicted by the LES d04: both observations and simulations show nearly the same value at each vertical level.

4.2 Full year intercomparison

For the analysis of the full year of 2017, we filter out all 10-min periods in which, first, the mean wind speed measured by any of the cup anemometers is below 1 m s$^{-1}$, which helps removing periods, e.g., when the instruments freeze, second, the wind direction measured by any of the wind vanes is within $0 \pm 15^\circ$, which decreases the influence of the wakes of the wind turbines north of the mast, and third, 10-min periods where the 1-Hz records of the cup anemometers are not entirely available. Without such filtering, the dataset consists of 51987 out of 52560 possible 10-min periods within the entire 2017 and after filtering, 45868 10-min periods remain for the analysis.

Figure 8 shows observed and simulated wind roses at 244 m, the height at which the uppermost cup anemometer and wind vane are mounted. As
illustrated, westerlies are predominant at Østerild and the simulations from the innermost LES domain (d04) capture well both the direction and the ranges of wind speed within most of the directions. From the observed wind rose, a ‘notch’ appears at the north direction due to ‘wake’ filtering but these are the least predominant directions at this site.

Fig. 8 Wind roses based on (a) observations from the cup anemometer and vane at 244 m and (b) output from the simulations at \( \approx 238 \) m from the innermost LES domain (d04) at Østerild during 2017.
The excellent agreement between the wind climatology based on the observations and the simulations at Østerild is partly due to a good match between the roughness length used in our WRF model configuration and that of the observations. The roughness length used in the model is read from the vegetation parameter table associated with the MODIS land cover classification with minimum and maximum values of 0.20 and 0.50 m, respectively, and is updated daily at the grid point of Østerild. There are different approaches to estimate the surface roughness length \( z_0 \) from observations; here we assume that the logarithmic wind profile is valid within the first tens of meters of the atmosphere, which results in

\[
z_0 = z \exp \left( \frac{\kappa U}{u_*} \right)^{-1},
\]

where \( z \) is the height of observation, \( u_* \) the friction velocity, and \( \kappa = 0.4 \) the von Kármán constant. By using Eqn. (6) and observations from a single vertical level only, we reduce the uncertainty on the \( z_0 \) estimations. Therefore, and only for deriving \( z_0 \) values from observations for the full year, we use 10-min statistics based on the sonic anemometer at 37 m. Due to the availability of the instrument during 2017, we have only 17852 10-min periods. Since Eqn. (6) is only valid for neutral atmospheric conditions, we select 10-min periods within the range \( |1/L| \geq 0.05 \text{ m}^{-1} \), where \( L \) is the Obukhov length, which is computed from the sonic anemometer means and fluxes as

\[
L = -\frac{(T_s)w_T^3}{\rho g(w'T_s')},
\]

where \( T_s \) is the sonic anemometer temperature, \( g \) the Earth’s gravitational acceleration, and \( (w'T_s') \) the sonic’s kinematic heat flux, which has similar values to the virtual potential kinematic heat flux. We also apply the corrections suggested by Liu et al. (2001) for the computation of the heat flux from this particular sonic anemometer. Further, as demonstrated by Peña (2019), \( z_0 \) largely depends on the wind direction at Østerild. For the predominant westerlies at Østerild, \( z_0 \) derived from Eqn. (6) is the lowest: this is expected as forest surrounds the mast position except west of the mast where there is a clearing (see Fig. 1). Therefore, we also restrict the roughness analysis to winds from the direction 270° ± 15° based on the sonic-derived wind direction. A total of 1428 10-min periods then remain after the directional and atmospheric stability criteria are applied to the observations. Figure 9(a) shows a comparison between the roughness length seen at Østerild by the innermost LES domain (d04) and that derived from the sonic anemometer observations. The 10-min estimates from the observations are averaged monthly and yearly as \( \exp \left( \langle \log z_0 \rangle \right) \). The yearly observed value is 0.2617 m (black dashed line), which is close to both the omnidirectional yearly average from the values at d04 (0.3314 m in the red dashed line) and the value (0.2492 m), which was derived from a similar analysis using the sonic anemometer at 37 m during a 4-year period under the same direction range 270° ± 15° (Peña, 2019). As illustrated, \( z_0 \) at d04 depicts a clear seasonal change peaking during the summer.
months, as expected. The derived monthly averages from the observations are close to those at d04 but they do not show the seasonal trend; the value in July is indeed the lowest. When we estimate $z_0$ from the sonic observations using Eqn. (6), we assume that the roughness includes the displacement height. For the summer months, the displacement height can be large, which decreases the $z_0$ values but we do not include displacement height in the derivation as the model does not include it either.

![Fig. 9](image-url)

**Fig. 9** (a) Roughness length $z_0$ estimated from sonic anemometer observations and that used by the innermost LES domain d04 for 2017 at Østerild (see text for details). The circles represent the value $\exp(\langle \log z_0 \rangle)$ within 30-day periods and the error bar ± one standard error. The dashed lines show the value $\exp(\langle \log z_0 \rangle)$ for 1428 10-min $z_0$ estimates from the observations (in black) and the mean of the values used by the simulations (in red). (b) Time series of the 10-min mean wind speed from the cup anemometer at 106 m and the LES d04 simulations at 110 m.

Figure 9(b) shows the year-long time series of 10-min mean winds from both observations from the cup anemometer at 106 m and output from the LES d04 domain at ≈110 m. As illustrated, the simulations follow well the observations, which portray the wide range of wind speeds at the site. To understand how accurate the simulated time series of wind speed are, we compare the histograms of wind speed based on the observations and simulations for the vertical level at ≈110 m. Figure 10(a) illustrates the year-long histograms of wind speed from the output of the LES domain d04 at ≈ 110 m and the cup anemometer at 106 m.

As illustrated, for this particular height the histograms agree closely, with the simulations generally showing higher frequencies of occurrence for winds $\geq 8$ m s$^{-1}$. For visualization purposes we do not include the histogram based on the output of the mesoscale domain d02, which has a similar behavior as that of domain d04 but it is slightly biased towards higher wind speeds with relative lower frequencies at the histogram’s peak. However, we show the Weibull dis-
One-year long turbulence measurements and modeling using WRF-LES

Fig. 10 (a) Wind speed histograms based on the 10-min means of the 106-m cup anemometer and 10-min means of the ≈110-m output of the LES domain d04. The solid lines are based on maximum likelihood estimates of the parameters of the Weibull distribution using the time series of wind speeds, where we also show the distribution computed from 1-h instantaneous values of the ≈110-m output of the mesoscale domain d02. (b) Vertical profile of the power density computed from the different time series of wind speed (see text for details)

Fig. 10 (a) Wind speed histograms based on the 10-min means of the 106-m cup anemometer and 10-min means of the ≈110-m output of the LES domain d04. The solid lines are based on maximum likelihood estimates of the parameters of the Weibull distribution using the time series of wind speeds, where we also show the distribution computed from 1-h instantaneous values of the ≈110-m output of the mesoscale domain d02. (b) Vertical profile of the power density computed from the different time series of wind speed (see text for details)

Distributions of observed and simulated mean wind speeds. These are computed from the Weibull shape and scale parameters derived via a maximum likelihood estimator that uses the time series of wind speed. Figure 10(a) shows that these all-sectors Weibull distributions fit the wind speed histograms well but underestimate the frequency of the histogram peak. Also, and most importantly for wind energy purposes, both observed and d04-simulated Weibull distributions are close to each other, which shows, first, a very good agreement between the observed and simulated Weibull parameters, and second, that the simulated and observed power density, which is a measure of the energy available in the wind, are very close. The power density is defined as

\[ P_d = \frac{1}{2} \frac{1}{N} \rho \sum_{i=1}^{N} U_i^3, \]  

(8)

where we assume a constant air density \( \rho = 1.225 \text{ kg m}^{-3} \), and \( N \) is the amount of samples. This is further illustrated in Fig. 10(b): power densities computed from time series of the observations and the output of the LES domain d04 are close at each vertical level, and those computed from the output of the mesoscale domain d02, although also close, are the ones generally deviating the most.

In Fig. 11, we also show year-long histograms of a turbulence parameter, the standard deviation of the wind speed. The figure illustrates the results for computations based on the analysis of the 10-min statistics of the 1-Hz cup anemometer observations at 106 m, similarly on the analysis of 10-min statistics of the 12-s output of the LES domain d04, and on estimations based
on the 1-h instantaneous output of the mesoscale domain d02. When using the MYNN2 PBL scheme in the WRF model, one can output a parameter $Q_{KE}$ that corresponds to twice the TKE. Here, to estimate $\sigma_U$ from the output of d02 we assume that

$$\sigma_U \approx \left( \frac{2}{3} e \right)^{1/2},$$

(9)

where $e$ is the TKE; since we use the MYNN2 PBL scheme, a $Q_{KE}$-based prognostic equation is employed.

The histograms in Fig. 11 reflect the wide range of atmospheric turbulence conditions at the site. The histogram based on the output of the LES domain d04 (in red) shows a very close resemblance to that from the observations (in grey). The largest discrepancies between the latter two histograms appear within the values close to zero (in the first bin), where the frequency of simulated values is close to five times that of the observations. This is partly due to the wind speed filter applied on the cup anemometer records, which also results in the filtering of a portion of very low observed turbulent conditions. It is also partly due to deficiencies in resolving turbulence, particularly during nighttime (as we show in more detail in Sect. 4.3), where stable atmospheric conditions dominate at Østerild and turbulence is low. Based on analysis of velocity spectra from the sonic anemometers at Østerild, Peña (2019) showed that under stable conditions, the atmospheric flow has the lowest turbulence length scales within the whole extension of the mast; these cannot be resolved by the innermost LES domain d04, whose effective resolution is $\approx 6\Delta x = 300$ m. Within the high turbulence range ($\sigma_U \geq 2$ m s$^{-1}$), the d04-based turbulence values are also more frequent than those of the observations; the latter

![Fig. 11 Wind speed standard deviation histogram based on 1-year output of the LES domain d04 and mesoscale domain d02 at $\approx 110$ m, and the observations from the cup anemometer at 106 m. The cup- and d04-based results are from the analysis of the 10-min statistics, and the d02-based results from instantaneous 1-h model output.](image-url)
One-year long turbulence measurements and modeling using WRF-LES shows higher frequencies within the range \(0.7 \leq \sigma_U \leq 1.8\) m s\(^{-1}\). The largest differences, when comparing simulation-based values to the observations, clearly appear for the mesoscale-derived turbulence measures from d02. Turbulence levels from d02 concentrate within a narrower portion and closer to the lower side of the histogram \((0 \leq \sigma_U \leq 2)\) m s\(^{-1}\). The mesoscale-based histogram is bimodal: the peak at wind standard deviation values close to zero combines low wind and low turbulence periods mainly occurring during the late night and early morning. For the rest of the vertical levels (not shown), the bimodal behavior of the mesoscale-derived output is less clear closer to the surface and is enhanced the higher the level examined. When comparing the observed histograms with the LES d04-based output (also not shown) and with respect to the results at \(\approx 110\) m in Fig. 11, the simulations are slightly biased towards lower values of wind speed standard deviation closer to the surface, whereas they are slightly biased towards higher values of wind speed standard deviation upwards.

With regards to TI, which combines both the abilities of the model and its configuration to capture both turbulence and mean winds, we analyze its behavior as function of wind speed in Fig. 12 for four of the matching (observations/simulations) vertical levels. There, the full-year 10-min TI values computed for both simulations and observations are binned within wind speed bins to ease the comparison.

As illustrated, for all vertical levels both simulations and observations show a wide range of TI values (0–30%), particularly at low wind speeds, which reflect both the wide variety of turbulence and atmospheric stability conditions at Østerild. Although not that clear at the lowest displayed height (Fig. 12a), the simulated and observed TI behaves similarly as function of wind speed. At high wind speeds, TI tends to stabilize, which is the typical behavior of TI over flat terrain under near-neutral atmospheric stability conditions. When approaching the low wind speed ranges, we observe the effect of atmospheric stability on the TI both on the mean within each wind speed bin and on the spread of the values (shown in the error bars). In general, the observed mean turbulence levels are fairly matched by those computed from the simulations. The simulations show lower TIs for \(U \leq 7\) m s\(^{-1}\), which is partly due to lack of turbulence in the simulations under some late evening and early morning periods mostly, where wind speeds are low, but also because the simulated wind speeds are higher than the observations during these periods. For \(U \geq 7\) m s\(^{-1}\), the simulations tend to show higher TIs, a behavior intensified the higher the vertical level. This is mainly due to an increased bias towards higher turbulence on the simulations compared to the observations as explained when presenting the turbulence histogram in Fig. 11. Figure 12(a), in particular, shows the limitations of the modelling system in its current configuration, as at this level the model has difficulties to resolve turbulence (or cannot at all), particularly under the low wind conditions, where stable atmospheric conditions, characterized by relatively small turbulent eddies, can be dominant.

We explore the behavior of the simulated and observed velocity spectrum for the full year in Fig. 13. Both the time series of 10-min mean wind speeds
Fig. 12 Turbulence intensity (TI) as function of the wind speed at (a) 40 m, (b) 106 m, (c) 178 m, and (d) 244 m based on the 10-min statistics computed using the output of the innermost LES domain d04 and the cup anemometer observations. The TI values are binned within 0.5 m s\(^{-1}\) intervals; the mean for each bin is shown in filled circles and the error bars show ±1 standard deviation within the bin. Each frame title shows the observational height and simulated mean vertical level from the cup anemometer at 106 m and the output of the LES domain d04 at ≈110 m are split into 73 time series that are 5-day long. The spectra computed from these 73 time series are ensemble averaged and the average spectrum is that illustrated in the figure. Similarly, we compute an average spectrum from the 1-h time series of instantaneous winds that are output from the mesoscale domain d02 at ≈110 m.

As illustrated, the observations show the expected behavior within the mesoscale range, i.e., the spectrum follows the \(-5/3\) slope (also shown). The simulated spectrum from the mesoscale domain d02 and the innermost LES domain d04 follows rather well the observed spectrum down to frequencies close to 2 h and 1 h, respectively. There are no signs of decay of any of the simulated spectra at the high resolved frequencies when compared to the observed spectrum, which is a typical feature of mesoscale flow models (Skamarock, 2004). Conversely, the simulated spectrum from the LES domain d04 shows higher energy content than the observed spectrum from frequencies close to 1 h up to 20 min. This is most probably due to artifacts of the simulations: first,
by perturbing the two innermost LES domains, we could artificially increase the variability of the 10-min mean winds, and second, spurious convective roll structures occasionally forming on d03 could increase the variability on d04 as those rolls advect into it and interact with the perturbations and developing turbulence therein.

4.3 Sensitivity to LES inflow perturbations

The high energy content within the high frequencies of the simulated winds from the output of the LES domain d04 when perturbing both LES domains d03/d04, in particular, show us the need to explore the sensitivity on the output of this particular domain to different alternatives for inflow perturbation. Therefore, we also perform two more sets of simulations: one with inflow perturbation at the d04 boundaries only (d04 in Fig. 14) and the other perturbing the d03 boundaries only (d03 in Fig. 14). Since, we cannot afford three sets of simulations for an entire year, we choose a 10-day period, where we observe a variety of wind and turbulence conditions, high and low wind and turbulence events, and clear diurnal cycles. Figure 14 illustrates the time series of three 10-min statistics, namely the mean wind speed, along-wind variance, and turbulence intensity, within the selected 10-day period from both the output of the innermost LES domain d04 from the three sets of simulations at ≈ 110 m and the cup anemometer observations at 106 m.

As illustrated in Fig. 14(a), all three simulations capture well the observed mean wind speed during the 10-day period. During the few events where observations deviate noticeably from the observations (e.g., the late afternoon on June 15), all three of the simulations deviate similarly. Differences become
clever when looking at the along-wind variance time series in Fig. 14(b). All three simulations capture the trends of the observed variance. However, the d03 simulation frequently underpredicts the variance relative to the observations, particularly during the late evening and early morning periods during which the observed levels of turbulence are high. This is also clearly reflected in the TI time series shown in Fig. 14(c). One can also notice in the along-wind variance time series that about midnight during several of the diurnal cycles, all of the simulations underpredict the observed variance during low turbulence periods, e.g., between June 10/June 11 and June 17/June 18. This mismatch is sometimes hidden in the TI time series as the wind speed can be higher from the simulations compared to the observations, e.g., during the early morning of June 11. Further, one can also notice the higher variability of the along-wind variance from the simulation d03/d04 compared to simulation d04, particularly at high turbulence levels. Table 2 provides the root mean square error (RMSE) between the 10-min time series of observations and the output of the three simulations. As shown, the largest RMSEs for the mean wind speed, along-wind variance and TI are computed using the d03 simulation. For the along-wind variance, the d04 simulation has the lowest RMSE, whereas the simulation d03/d04 has the lowest RMSE for the mean wind speed. For this particular 10-day period, if we want to reproduce the best TI estimates, perturbing both LES domains provide the lowest error.

For completeness, we illustrate the velocity spectra within the 10-day period computed from both the observations and the three different simulations at ≈ 110 m in Fig. 15. Since we restrict the analysis to a single 10-day period, we are able to use the time series of 12-s simulated and 10-Hz observed wind speeds. The spectrum of the d04 simulation, where the perturbations are ap-
Table 2 Root mean square error (RMSE) of three flow characteristics between the simulated and observed 10-min time series at ≈110 m for the three simulations within the inflow perturbation sensitivity study.

<table>
<thead>
<tr>
<th>Variable/simulation</th>
<th>d03/d04</th>
<th>d03</th>
<th>d04</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle U \rangle$ [m s$^{-1}$]</td>
<td>1.540</td>
<td>1.573</td>
<td>1.647</td>
</tr>
<tr>
<td>$\langle u'^{2} \rangle$ [m$^2$ s$^{-2}$]</td>
<td>1.443</td>
<td>1.639</td>
<td><strong>1.263</strong></td>
</tr>
<tr>
<td>TI</td>
<td><strong>0.061</strong></td>
<td>0.077</td>
<td>0.068</td>
</tr>
</tbody>
</table>

plied on the boundaries of the innermost LES domain d04 only, follows the observed spectrum the closest within the range of frequencies 6 h to 10 min, with no clear signs of higher wind variability within the range of frequencies between 10 min and 1 hour. The d03 and d03/d04 simulations follow closely each other up to a frequency close to 10 min and both portray higher variability than the observations within the range of frequencies between 10 min and 1 hour, as similarly illustrated for the year-long analysis in Fig. 13. From a frequency close to 10 min to the highest frequencies, the variability of the d03 simulation drops the fastest, whereas both the d03/d04 and d04 simulations exhibit similar behavior, only dropping relative to the observations at frequencies corresponding to a few minutes minutes or less.

Fig. 15 Wind speed averaged power spectrum based on the 12-s time series of the output of the LES domain d04 at 111 m from three configurations of the inflow perturbations, and the 10-Hz time series of the cup anemometer observations at 106 m. The mean spectrum is derived from 6-h long time series for a 10-day period. The spectral slope of $-5/3$ is shown in the grey line.

5 Conclusions and discussion

A full-year multiscale atmospheric simulation combining both mesoscale weather and explicit turbulence simulation is performed at a largely flat and rough site...
in northern Denmark. Results from an LES domain nested within a mesoscale
simulation, all within the WRF model, are assessed by comparing both mean
winds and turbulence measures with observations recorded with a heavily in-
strumented 250-m high-quality meteorological mast.

Post-processed 10-min mean wind speeds and directions from the output
of the simulation’s innermost LES domain show very good agreement with
cup anemometer and wind vane observations for the entire year and for the
extension of the mast. The wind climatology, both in terms of wind roses and
histograms of mean winds, is well represented by the simulation; further the
observed power density at the different vertical levels is better predicted by
the innermost LES domain output than by the mesoscale domain output with
the finest resolution.

Histograms of post-processed 10-min along-wind variances from the output
of the simulation’s innermost LES domain agree very well with those from the
observations. However, the instantaneous hourly output of the high-resolution
mesoscale domain has difficulties capturing the high levels of variance, fre-
quently producing very low variances when compared to the observations.
The behavior of observed turbulence intensity with wind speed is very well
captured by the innermost LES domain output at most vertical levels, with
the best agreements at the greatest heights, and deteriorating with increasing
proximity to the surface. Turbulence intensity also tends to be underestimated
and overestimated by the simulations under low and high wind speed ranges,
respectively.

Spectral analysis shows that the innermost LES domain output presents
larger variability than the observations at high frequencies. A sensitivity study
reveals that the increased variability is partly due to application of the CPM
to both LES domains, rather than just the innermost one. For this particular
10-day period, the observed variability can be better matched by perturbing
the innermost LES domain only.

This study demonstrates a capacity to use a multiscale modeling system
to reproduce both the mean wind and the turbulence conditions (which are
part of the so-called site conditions) at this particular site with high fidelity.
More evaluations of the modeling system under different topographical and
climatological conditions are envisioned for the future to fully understand the
capability of the system. This will be highly beneficial for the exploitation of
wind energy and for the understanding of the atmospheric turbulence condi-
tions also where measurements are not available. It should be noted that
the ability to reproduce flow characteristics of this modeling system depends
strongly on the topographical inputs to the model. For Østerild, the rough-
ness length, which is used in the model, matches well that estimated from the
observations but this might not be the case for many other sites.

Acknowledgements The work is partly funded by Independent Research Fund Denmark
through the ‘Multi-scale Atmospheric Modeling Above the Seas’ (MAMAS) project. This
work is also partly funded by the Ministry of Foreign Affairs of Denmark and administered
by Danida Fellowship Centre through the ‘Multi-scale and model-chain Evaluation of Wind
One-year long turbulence measurements and modeling using WRF-LES

Atlases’ (MEWA) project. JDM’s contribution is supported by LLNL under contract DE-AC52-07NA27344 and by the US Department of Energy’s Wind Energy Technologies Office.

References