On the Sensitivity of Large-Eddy Simulations of the Atmospheric Boundary Layer Coupled with Realistic Large-Scale Dynamics✏️

PAOLO GIANI[a] AND PAOLA CRIPPA[a]

[a] *Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, Indiana*

ABSTRACT: We present a new ensemble of 36 numerical experiments aimed at comprehensively gauging the sensitivity of nested large-eddy simulations (LES) driven by large-scale dynamics. Specifically, we explore 36 multiscale configurations of the Weather Research and Forecasting (WRF) Model to simulate the boundary layer flow over the complex topography at the Perdigão field site, with five nested domains discretized at horizontal resolutions ranging from 11.25 km to 30 m. Each ensemble member has a unique combination of the following input factors: (i) large-scale initial and boundary conditions, (ii) subgrid turbulence modeling in the *gray zone* of turbulence, (iii) subgrid-scale (SGS) models in LES, and (iv) topography and land-cover datasets. We probe their relative importance for LES calculations of velocity, temperature, and moisture fields. Variance decomposition analysis unravels large sensitivities to topography and land-use datasets and very weak sensitivity to the LES SGS model. Discrepancies within ensemble members can be as large as 2.5 m s^{-1} for the time-averaged near-surface wind speed on the ridge and as large as 10 m s^{-1} without time averaging. At specific time points, a large fraction of this sensitivity can be explained by the different turbulence models in the gray zone domains. We implement a horizontal momentum and moisture budget routine in WRF to further elucidate the mechanisms behind the observed sensitivity, paving the way for an increased understanding of the tangible effects of the gray zone of turbulence problem.

SIGNIFICANCE STATEMENT: Several science and engineering applications, including wind turbine siting and operations, weather prediction, and downscaling of climate projections, call for high-resolution numerical simulations of the lowest part of the atmosphere. Recent studies have highlighted that such high-resolution simulations, coupled with large-scale models, are challenging and require several important assumptions. With a new set of numerical experiments, we evaluate and compare the significance of different assumptions and outstanding challenges in multiscale modeling (i.e., coupling large-scale models and high-resolution atmospheric simulations). The ultimate goal of this analysis is to put each individual assumption into the wider perspective of a realistic problem and quantify its relative importance compared to other important modeling choices.

1. Introduction

The state of the atmospheric boundary layer (ABL) at any space–time point is the result of several atmospheric motions at different scales, from global and mesoscale circulations to turbulent eddies in the near surface. Numerical weather prediction (NWP) models with $\mathcal{O}(10)$ km horizontal grid spacing Δx can resolve synoptic and mesoscale phenomena, and they rely on parameterizations for the unresolved land–atmosphere interactions and boundary layer processes. Recently, there has been a growing interest in large-eddy simulations (LES) that incorporate realistic topography and land use, as well as specified time- and space-varying boundary conditions from lower-resolution NWP simulations (Chow et al. 2019). Realistically forced LES are referred to as nested multiscale simulations in recent literature (Talbot et al. 2012; Haupt et al. 2019) and rely on subgrid-scale (SGS) models to parameterize the impact

of subgrid turbulent scales. Important processes dictated by land surface heterogeneity, topography, and buoyancy can be explicitly resolved by multiscale simulations.

Advances in mesoscale-to-microscale coupling techniques and the expansion in computational resources present the opportunity to overcome existing accuracy limits in operational NWP models and sustain the NWP *quiet revolution* described by Bauer et al. (2015). For instance, the massive simulations presented by Stevens et al. (2020) tested the hypothesis that using cloud-resolving LES over the entirety of Germany ($\Delta x = 156$ m) can leapfrog existing barriers in representing cloud and precipitation fields, achieving promising results. Even with lower computational resources, limited areas of interest can be operationally simulated with fine resolution, resolving the third direction of motion and important local-scale processes dictated by surface heterogeneity, along with large-scale dynamics (Lin et al. 2021).

Refining the numerical grid of NWP models allows studying and understanding local-scale phenomena of interest to researchers and decision-makers. Examples include (i) urban thermal comfort and urban air pollution—Ronda et al. (2017) showed how the rural–urban difference in temperature is not captured with their coarse simulations but is well resolved with urban-scale simulations; (ii) wind farm operation and

---

Corresponding author: Paolo Giani, pgiani@nd.edu

planning, especially in complex terrain where marked differences exist between ridges and valleys at scales of $\mathcal{O}(1)$ km (Giani et al. 2020; Mirocha et al. 2018); (iii) the representation of clouds and precipitation—Shin et al. (2021) and Schemann et al. (2020) illustrated that LES forced by realistic lateral boundary conditions, large-scale forcings, and real topography influenced timing, height, and phase of clouds; and (iv) prediction of the magnitude, track, and timing of tropical cyclones (TCs)—numerically resolving turbulent moisture and energy transfer in the TC boundary layer is crucial to understand the evolution of TCs, as they are challenging to observe (Emanuel 1995; Xu et al. 2021). Other interesting applications include national security problems related to airborne dispersion of contaminants in urban areas (Kochanski et al. 2015) and unmanned aerial vehicle safety due to local wind disturbances and local-scale turbulence (Wang et al. 2019).

Despite the wide range of opportunities, multiscale modeling presents several challenges that hinder the widespread and operational use of such simulations. One of the most demanding problems is the so-called *gray zone* of turbulence (or *terra incognita*), which refers to the challenges in modeling turbulence at intermediate grid spacings that are appropriate neither for ABL parameterizations nor for LES SGS models (Wyngaard 2004). For multiscale nested simulations, the gray zone problem arises in the intermediate domain that forces the innermost LES calculations. Although several researchers investigated the problem in recent years and made significant progress (Chow et al. 2019; Zhou et al. 2018; Honnert 2016; Efstathiou and Plant 2019; Juliano et al. 2022), there is still no widespread consensus on how to deal with turbulence modeling at such resolutions. Another set of challenges include numerical errors and potential instabilities induced by steep slopes in high-resolution simulations that use terrain-following coordinates to simplify the implementation of the bottom boundary conditions. For instance, the Advanced Research version of WRF (ARW) dynamical solver is formulated using a terrain-following hydrostatic-pressure vertical coordinate system (Skamarock et al. 2021), which was shown to introduce truncation errors in the metric terms that can degrade the quality of the numerical solution and lead to numerical instabilities with very steep slopes (Klemp et al. 2003; Lundquist et al. 2010; Daniels et al. 2016). Although innovative strategies to alleviate or eliminate the issue have been proposed in the literature [e.g., the immersed boundary method (IBM) (Lundquist et al. 2010) and truly horizontal diffusion schemes (Arthur et al. 2021)], no operational solution has been achieved yet due to the large complexity of implementation and validation of these methods (Bao et al. 2018). For nonidealized simulations with the innermost domain discretized with $\Delta x \sim \mathcal{O}(1–10)$ m, other challenges involve the choice of realistic high-resolution topography and land-use datasets and the initial (re)analysis to drive the outermost domain as well as appropriate SGS models for LES (Liu et al. 2020). Extensive European literature has identified high-resolution topography and land-use datasets as key factors to improve the performance of high-resolution $\mathcal{O}(1)$ km simulations (Jiménez-Esteve et al. 2018; Schicker et al. 2016; De Meij and Vinuesa 2014), and we expect their effect to be even greater for $\mathcal{O}(1–10)$ m simulations in locally heterogeneous terrain.

Fernández-González et al. (2018) showed that using two different analyses to drive $\mathcal{O}(1)$ km WRF simulations can lead to significant differences in the skill of wind energy forecasts under different meteorological conditions.

While these different aspects have been individually investigated by several studies, the relative importance of the several factors that influence multiscale LES, such as the selection of the initial (re)analysis, land-use datasets, turbulence modeling in the gray zone, and LES subgrid-scale models, remains rather unexplored. In this work, we develop new simulations aimed at comprehensively gauging the global sensitivity of multiscale LES to the different assumptions that real problems in multiscale modeling require. In other words, we aim to put into context the recent challenges highlighted in multiscale modeling (e.g., turbulence modeling in the gray zone) with other important assumptions in real simulations (e.g., synoptic-scale analyses to initialize the real simulations and the selection of different high-resolution land-use and topography input datasets). The ultimate goals are to (i) understand which factors are the most significant for real multiscale problems, (ii) understand the physical/numerical mechanisms of the observed global sensitivity, and (iii) provide evidence-based guidance on where most future research efforts in multiscale modeling should be directed to.

To this end, we design a new ensemble of 36 multiscale experiments with WRF. The dataset consists of nested LES over the double ridge topography in Vale do Cobrão (Perdigão site, Portugal), where an intensive field campaign focused on microscale measurements took place in 2017 (Fernando et al. 2019). The 36 experiments encompass different combinations between different input factors, including initial conditions, land and topography datasets, SGS model of the LES, and turbulence modeling in the gray zone. We analyze the simulation ensemble by (i) decomposing the variance across the ensemble members, to understand the relative importance of each factor in explaining the ensemble variance of simulated winds, moisture, and temperature over the double ridge, and (ii) investigating the physical/numerical mechanisms that explain the observed sensitivity (for instance, how different gray zone schemes can lead to different winds in the LES).

The remainder of the article is organized as follows. Section 2 presents the multiscale numerical experiments with WRF. Section 3 describes the data analysis methods, including the variance decomposition procedure and the inline horizontal momentum budget. Based on the results presented in section 4, section 5 discusses the implications of our findings and summarizes the main conclusions.

2. Numerical experiments

a. Model configuration and simulated episode

We perform a set of 36 multiscale nested numerical experiments using WRF v4.4, which solves the nonhydrostatic compressible Euler equations in hybrid coordinates (Skamarock et al. 2021). The difference between the 36 experiments is related to the different assumptions and modeling choices that real multiscale problems require. Specifically, four different

TABLE 1. Summary of the 36 experiments presented in this work. The first letter in the run label refers to the initial analysis (G = GDAS, E = ERA5, and W = HRES-ECMWF), whereas the second letter refers to the topography datasets (C = coarse and F = fine). Even numbers use the Lilly–Smagorinsky SGS model, whereas odd numbers use the Deardorff scheme.

Run label	IC/LBC (re)analysis	Topography/land-use datasets	Turbulence in gray zone (d03–d04)	LES SGS model (d05)
GC01	GDAS	GMTED/MODIS	1D YSU PBL	Deardorff
GC02	GDAS	GMTED/MODIS	1D YSU PBL	Lilly–Smagorinsky
GC03	GDAS	GMTED/MODIS	1D Shin–Hong PBL	Deardorff
GC04	GDAS	GMTED/MODIS	1D Shin–Hong PBL	Lilly–Smagorinsky
GC05	GDAS	GMTED/MODIS	3D Zhang	Deardorff
GC06	GDAS	GMTED/MODIS	3D Zhang	Lilly–Smagorinsky
GF01	GDAS	SRTM/CORINE	1D YSU PBL	Deardorff
GF02	GDAS	SRTM/CORINE	1D YSU PBL	Lilly–Smagorinsky
GF03	GDAS	SRTM/CORINE	1D Shin–Hong PBL	Deardorff
GF04	GDAS	SRTM/CORINE	1D Shin–Hong PBL	Lilly–Smagorinsky
GF05	GDAS	SRTM/CORINE	3D Zhang	Deardorff
GF06	GDAS	SRTM/CORINE	3D Zhang	Lilly–Smagorinsky
EC01	ERA5	GMTED/MODIS	1D YSU PBL	Deardorff
EC02	ERA5	GMTED/MODIS	1D YSU PBL	Lilly–Smagorinsky
EC03	ERA5	GMTED/MODIS	1D Shin–Hong PBL	Deardorff
EC04	ERA5	GMTED/MODIS	1D Shin–Hong PBL	Lilly–Smagorinsky
EC05	ERA5	GMTED/MODIS	3D Zhang	Deardorff
EC06	ERA5	GMTED/MODIS	3D Zhang	Lilly–Smagorinsky
EF01	ERA5	SRTM/CORINE	1D YSU PBL	Deardorff
EF02	ERA5	SRTM/CORINE	1D YSU PBL	Lilly–Smagorinsky
EF03	ERA5	SRTM/CORINE	1D Shin–Hong PBL	Deardorff
EF04	ERA5	SRTM/CORINE	1D Shin–Hong PBL	Lilly–Smagorinsky
EF05	ERA5	SRTM/CORINE	3D Zhang	Deardorff
EF06	ERA5	SRTM/CORINE	3D Zhang	Lilly–Smagorinsky
WC01	HRES-ECMWF	GMTED/MODIS	1D YSU PBL	Deardorff
WC02	HRES-ECMWF	GMTED/MODIS	1D YSU PBL	Lilly–Smagorinsky
WC03	HRES-ECMWF	GMTED/MODIS	1D Shin–Hong PBL	Deardorff
WC04	HRES-ECMWF	GMTED/MODIS	1D Shin–Hong PBL	Lilly–Smagorinsky
WC05	HRES-ECMWF	GMTED/MODIS	3D Zhang	Deardorff
WC06	HRES-ECMWF	GMTED/MODIS	3D Zhang	Lilly–Smagorinsky
WF01	HRES-ECMWF	SRTM/CORINE	1D YSU PBL	Deardorff
WF02	HRES-ECMWF	SRTM/CORINE	1D YSU PBL	Lilly–Smagorinsky
WF03	HRES-ECMWF	SRTM/CORINE	1D Shin–Hong PBL	Deardorff
WF04	HRES-ECMWF	SRTM/CORINE	1D Shin–Hong PBL	Lilly–Smagorinsky
WF05	HRES-ECMWF	SRTM/CORINE	3D Zhang	Deardorff
WF06	HRES-ECMWF	SRTM/CORINE	3D Zhang	Lilly–Smagorinsky

input features are varied across the 36 experiments, including (i) how we model turbulent fluxes at intermediate gray zone resolutions, (ii) the choice of the SGS model used for LES, (iii) the initial analysis and boundary conditions that drive the simulations in the outermost domain, and (iv) the static datasets for the bottom boundary conditions in the finest domains. In total, we consider two options for static datasets, three options to model turbulence in the gray zone, two SGS LES models, and three initial analyses and boundary conditions. All the possible combinations between the different options are $2 \times 3 \times 2 \times 3 = 36$, which is the number of experiments that we run. Table 1 summarizes all the experiments. The following subsections provide a more detailed description of the four input features and how we varied them, while the current section focuses on the model setup that was consistent across all experiments.

We use five one-way nested domains centered in the Perdigão Valley, where a large field program in 2017 collected detailed measurements to unravel the complexities of microscale winds (Fernando et al. 2019). The local topography resembles a text-book example of a 2D valley located between two parallel ridges, with prevailing winds traversing the ridges perpendicularly on a climatological basis. The quasi-idealized nature of the study area along with the wealth of measurements collected during the intensive operation period (1 May–16 June) makes Perdigão an optimal case study for multiscale real simulations, although detailed observations show that the flow is far from being two-dimensional in reality (Fernando et al. 2019). The outermost domain (d01) is discretized with $N_x = N_y = 196$ grid points and $\Delta x = \Delta y = 11.25$ km in both zonal (x) and meridional (y) directions, with a Lambert Conformal grid that covers western Europe, North Africa, the Mediterranean Sea, and the eastern Atlantic Ocean (Fig. 1). The vertical dimension z is discretized with $N_z = 77$ grid points, reaching a constant isobaric surface of 50 hPa at the model top. The new hybrid coordinates of the ARW solver are used (Park et al. 2013; Beck et al. 2020), where the vertical coordinate represents a pure isobaric surface for $\eta \leq \eta_c = 0.2$. The eta
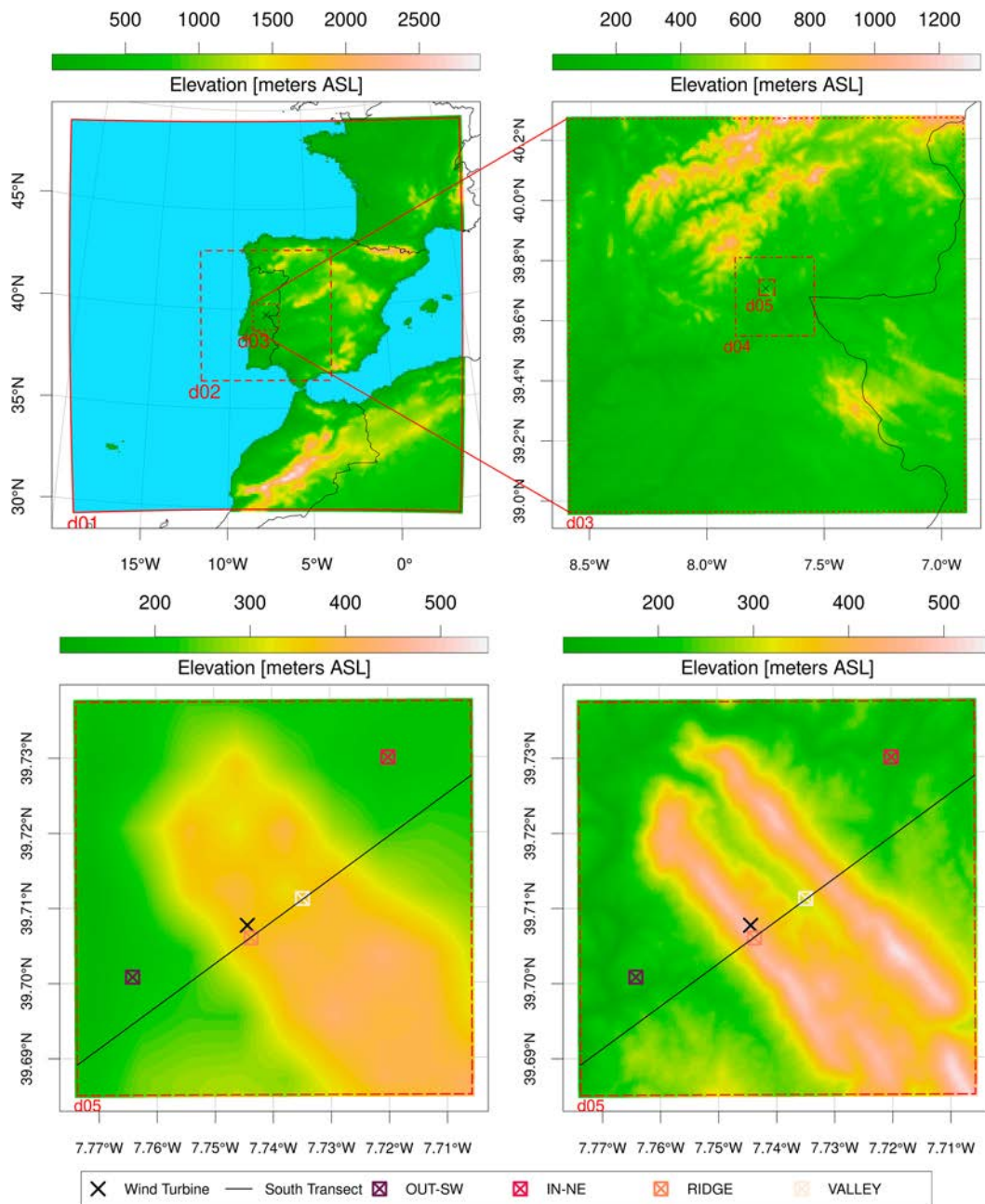
FIG. 1. (top) Elevation contours for (left) d01 and (right) d03, with annotations for the nested domains. (bottom) Close-up look and elevation contours for the d05 domains, with the (left) GMTED topography and (right) SRTM 1-s topography for d05. The transect and the locations annotated on the maps are used in the sensitivity analysis presented in section 4. The same color scale is used only for the bottom two panels. Other geometric characteristics of the domains can be found in Table 2.

levels are chosen to have approximately geometrically increasing Δz in an idealized hydrostatic boundary layer, starting from $\Delta z_S = 36$ m at the surface and reaching an approximately constant $\Delta z_F = 350$ m in the free atmosphere. The first model half level where horizontal velocities, pressure, and scalars are calculated is $z_1 \approx 18$ m. This is a fairly standard choice for the eta levels, and a similar version of this configuration has been used in Giani et al. (2022). Similarly to other multiscale

simulations over urban areas (Ronda et al. 2017) and Perdigão (Wagner et al. 2019), the four nested domains (d02, d03, d04, and d05) have a parent-to-child horizontal resolution ratio of 3:1, 5:1, 5:1, and 5:1, respectively, which correspond to $\Delta x = 3750$, 750, 150, and 30 m. The vertical grid in d02 and d03 is the same as in d01, whereas we refine the vertical levels in d04 and d05 to have a better resolved boundary layer with higher horizontal resolution, using Daniels et al. (2016) vertical

TABLE 2. Main geometric features of the five domains configuration for the numerical experiments. The different topography and land-use datasets are described in section 2b.

Domain	$N_x \times N_y \times N_z$	$\Delta x = \Delta y$	Δz_S	Topography dataset	Land-use dataset	Total domain size	Time step
d01	$196 \times 196 \times 77$	11 250 m	36 m	USGS	MODIS	2205×2205 km^2	75 s
d02	$196 \times 196 \times 77$	3750 m	36 m	USGS	MODIS	735×735 km^2	25 s
d03	$196 \times 196 \times 77$	750 m	36 m	USGS/SRTM	MODIS/CORINE	147×147 km^2	5 s
d04	$196 \times 196 \times 91$	150 m	24 m	USGS/SRTM	MODIS/CORINE	29.4×29.4 km^2	1 s
d05	$196 \times 196 \times 109$	30 m	12 m	USGS/SRTM	MODIS/CORINE	5.88×5.88 km^2	1/5 s

refinement capabilities (and subsequent updates for hybrid coordinates). We choose $\Delta z_S = 24$ m in d04 and $\Delta z_S = 12$ m in d05 and $N_z = 91$ and $N_z = 109$, respectively, which correspond to $\Delta z \approx 80$ m at the typical inversion layer during the afternoon convective boundary layer. Table 2 summarizes all the relevant parameters for the five domain configurations.

All experiments are forced by time- and space-varying surface heat and momentum fluxes dynamically determined with Monin–Obukhov similarity theory from soil characteristics and atmospheric properties at the first model level. Lateral boundary conditions are specified from coarser resolution datasets, described in the next subsection. The Unified Noah Land Surface Model (Ek et al. 2003) evolves soil temperature and moisture prognostically in four ground layers. The shortwave and longwave radiation inputs for the land surface model are calculated with the Dudhia scheme (Dudhia 1989) and the Rapid Radiative Transfer Model (Mlawer et al. 1997), respectively. Other physics modules include the Kain–Fritsch scheme (Kain and Fritsch 1990) to estimate SGS convective precipitation and related redistribution of heat in d01 (turned off in the other domains) and the Thompson scheme (Thompson et al. 2008) to compute moisture, heat, and six hydrometeors tendencies related to ice, snow, and graupel processes from cloud microphysics.

The simulated episode covers four consecutive precipitation-free days, from 0000 UTC 19 May to 0000 UTC 23 May 2017. The integration time step is set to $\Delta t = 75$ s for the outermost domain, and we use a parent-to-child time step ratio that follows exactly the horizontal resolution ratio. All the fields relevant to the analysis are saved with 5-min frequency in the innermost domain. Output is saved at every model time step for locations where flux towers are present. A mixture of clear sky conditions and medium-high clouds was observed during the simulated period (Fig. S1 in the online supplemental material). We chose this specific episode for multiple reasons: (i) There is an extensive coverage of radiosonde launches to characterize the vertical structure of the atmosphere, with as many as 6 launches per day on 21 May (although model verification is not the main focus of this study, but will be investigated in future/ongoing work), (ii) a subset of this period (20–21 May) has been analyzed in detail in previous literature (Connolly et al. 2021), and therefore, a comparison with preexisting work is possible, and (iii) different large-scale dynamics and synoptic conditions are observed during this period, which makes it an interesting case for multiscale experiments. Specifically, high pressure conditions associated with fair weather are observed during the first day of the simulation, while lower pressure, stronger synoptic winds, and high clouds move into the domain from

southwest starting from the evening of 20 May into 21 May. Weaker synoptic winds reach Perdigão on 22 May, when a substantial amount of wind veer is also observed. Figure S2 summarizes the synoptic conditions from three different global datasets, extracted in the grid point nearest to the Perdigão field site from the original (re)analysis datasets.

Every simulation is executed in parallel using the message passing interface (MPI) standard on 240 CPUs, split between five different computing nodes (i.e., 48 CPUs/node). Each node hosts a Dual 24-Core AMD EPYC 7451 workstation. On average, one simulation takes approximately 7200 CPU-hours for completion, which means that the computational cost for the entire ensemble is about 259 200 CPU-hours (10 800 CPU-days). The average wall-clock time per second of numerical integration is about 1.5 s. Note that we have also performed one additional sensitivity simulation with twice as many grid points and half the grid spacing in d05 (15 m), which is more than 10 times more expensive than the base case configuration ($\approx 100\,000$ CPU-hours for an individual simulation). The results from this additional test show a fairly negligible sensitivity to the number of grid points and the grid resolution (Fig. S3), and we accordingly only focus on the configuration reported in Table 2 for the rest of the manuscript.

b. Topography and land-use static datasets

First, we expect that the land-use and topography datasets will have a significant impact for problems in complex terrain. We consider two different datasets: (i) the topography and land-use datasets available in the default static fields of the WRF Preprocessing System (WPS) and (ii) ad hoc high-resolution data for both land use and topography. The default data included in the WPS4.4 release are the Global Multiresolution Terrain Elevation Data 2010 (GMTED) from the U.S. Geological Survey (USGS) for topography and the Noah-modified 21-category land-cover data from the International Geosphere–Biosphere Programme Moderate Resolution Imaging Spectroradiometer (IGBP-MODIS) for land use. The spatial resolution of both datasets is 30 arc s, which corresponds to approximately 900 m at the latitude of Perdigão. Both datasets are rather coarse for the finest resolution domains (d03–d05), but they represent the highest resolution product available by default in the WPS. To test the importance of the static datasets, we implement high-resolution datasets into the WPS that are more suitable to the fine grids in d03–d05. We implement the Shuttle Radar Topography Mission (SRTM) dataset for topography (Werner 2001) and the European Coordination of Information on the Environment (CORINE) land-cover dataset for land use (EEA 2002). The resolutions of the SRTM and CORINE datasets

are approximately 30 and 100 m, respectively. The CORINE land-use categories are recast into USGS categories available in WPS4.4 with the conversion table presented in Pineda et al. (2004). The differences between topography and roughness lengths for the coarse option (MODIS and GMTED) versus the fine option (SRTM and CORINE) are shown in Fig. 1 and Fig. S4, respectively. Note how, with the coarse-resolution dataset, the particular configuration of the double-ridged valley is smoothed into what resembles a 3D hill. Differences between spatially explicit roughness lengths are also important, in terms of both spatial distribution and spatial average values ($\overline{z_0} = 0.326$ m with the coarse dataset and $\overline{z_0} = 0.101$ m with CORINE). Wagner et al. (2019), Wenz et al. (2022), and Wise et al. (2022) argued that the CORINE dataset is likely to underestimate the roughness on the ridge and in the valley substantially because the hills are partially covered with ~20-m-tall eucalyptus trees, which should be represented by much larger values of z_0 [1–2 m, according to Wagner et al. (2019)]. The z_0 underestimation may impact the model performance for surface quantities (e.g., overestimation of near-surface wind speed), which will be addressed in a future study focused on model evaluation and accuracy. In this work, we are instead concerned with the sensitivity that arises because of the differences in z_0 and topography fields.

c. Initial and lateral boundary conditions

Three different global atmospheric (re)analyses are considered to specify initial conditions (ICs) and lateral boundary conditions (LBCs). The three analyses have different spatial and temporal resolutions. We use (i) the Global Data Assimilation System (GDAS) final analysis for the Global Forecast System, which is a gridded product (0.25°) updated every 6 h (NCEP 2015); (ii) the high-resolution operational data from the European Centre for Medium-Range Weather Forecast (HRES-ECMWF), which is on a finer grid compared to GDAS (0.08°) and also updated every 6 h (European Centre for Medium-Range Weather Forecasts 2011); and (iii) ECMWF's ERA5 reanalysis (Hersbach et al. 2020), which has a higher update frequency (1 h) and the same spatial resolution as GDAS (0.25°). All three datasets are valid choices to initialize WRF simulations and have been used by multiple WRF practitioners, and they are all based on observations, satellite measurements, and global simulations. However, they show differences both close to the surface and aloft (e.g., Fig. S2) that may influence the final downscaled LES solution. We are interested in understanding how large their influence is for LES downscaled solutions, compared to other assumptions that real LES of the ABL require. Note, for instance, the difference in surface pressure between datasets over Perdigão in Fig. S2, which is related to the different spatial resolutions (and elevation) of HRES-ECMWF compared to ERA5 and GDAS.

d. Gray zone modeling

Modeling turbulence in the gray zone has emerged as one of the main bottlenecks in multiscale modeling. There is no general consensus on how subgrid turbulent fluxes should be calculated at gray zone resolutions, where the magnitude of resolved turbulent fluctuations is grid-dependent and turbulent structures are not well resolved (Giani et al. 2022; Zhou et al. 2014; Ching et al. 2014). However, significant progress has been achieved in recent years, and several novel schemes have been proposed, based on pragmatic blending approaches (Efstathiou et al. 2018; Beare 2014; Zhou et al. 2018) and scale-aware versions of the 1D column-based schemes (Shin and Hong 2015, 2013; Honnert et al. 2011; Ito et al. 2015), usually referred to as 1D planetary boundary layer (PBL) schemes. We aim to gauge the sensitivity of real downscaled LES to these novel schemes. Specifically, we use a traditional 1D PBL scheme, a scale-aware 1D PBL scheme, and a 3D blending-type closure (Zhang et al. 2018) to model subgrid fluxes in the gray zone intermediate domains (d03 and d04). We select the Yonsei University (YSU) scheme as the 1D PBL scheme in our simulations, which compute subgrid vertical fluxes of a generic quantity C as follows:

$$\overline{w'C'} = -K_c\left(\frac{\partial \overline{C}^\Delta}{\partial z} - \gamma_c\right) + \overline{w'C'_{z_i}}\left(\frac{z}{z_i}\right)^3, \quad (1)$$

where $\overline{C}^\Delta$ is the resolved C at a specific resolution Δ (i.e., $C = \overline{C}^\Delta + C'$), $\overline{w'C'}$ is the subgrid turbulent flux of C, K_c is the eddy diffusivity, γ_c is the countergradient term, and z_i is the boundary layer height. The formulations of K_c, γ_c, and the entrainment flux $\overline{w'C'_{z_i}}$ are documented in Noh et al. (2003) and Hong et al. (2006). In summary, K_c is a height-dependent parabolic function that is zero at the surface and z_i and is based on a mixed layer velocity scale defined with the Deardorff (1970) convective scale w_* and the friction velocity u_*. The term γ_c is the nonlocal term and is calculated from the surface flux of C. The tendency added in the C prognostic equation (e.g., u, v, and potential temperature θ) from the YSU scheme is

$$\left(\frac{\partial C}{\partial t}\right)_{\text{PBL}} = -\frac{\partial}{\partial z}(\overline{w'C'}). \quad (2)$$

As a second option, we use Shin and Hong (2015) extension of the YSU scheme that includes scale awareness for gray zone resolutions:

$$\overline{w'C'} = -K_c\left(\frac{\partial \overline{C}^\Delta}{\partial z}\right)P_L(z, \Delta x_*) + \overline{w'C'_{\text{NL}}}P_{\text{NL}}(u_* w_*^{-1}, \Delta x_*), \quad (3)$$

where P_L and P_{NL} are scale-aware functions that downweigh local and nonlocal PBL fluxes (respectively) depending on the resolution $\Delta x_* = z_i^{-1}\Delta x$ and $\overline{w'C'_{\text{NL}}}$ is the nonlocal flux which has a different formulation compared to the YSU one (countergradient and entrainment). The tendency is computed as in YSU [Eq. (2)].

The third and final option that we test for modeling turbulence in the gray zone is Zhang et al. (2018) 3D blending scheme (i.e., horizontal and vertical fluxes are calculated):

$$\overline{u_i'C'} = -K_c\frac{\partial \overline{C}^\Delta}{\partial x_i} + (\overline{u_i'C'})_{\text{NL}}P_{\text{NL}}(u_* w_*^{-1}, \Delta x_*)\delta_{i3}, \quad (4)$$

where Einstein notation (summation implied over repeated indices) is used and $i = 1$, 2, and 3 for x, y, and z directions.

The term K_c in Eq. (4) is different from the YSU eddy diffusivity and is calculated with Deardorff (1980) turbulent kinetic energy (TKE) closure, which is a fairly standard choice for LES SGS modeling ($K_c = c_k \ell e^{1/2}$, where e is the TKE, ℓ is a grid- and buoyancy-dependent length scale, and $c_k = 0.18$ is a constant). Equation (4) is essentially Deardorff's standard 3D TKE closure for LES SGS turbulent fluxes, with two main differences. First, a nonlocal flux term is added in the vertical direction, to make it appropriate for the mesoscale and gray zone limits where large eddies responsible for nonlocal transport are not explicitly resolved. The nonlocal flux formulation is documented in Zhang et al. (2018) and is conceptually similar to Shin and Hong (2015). In the LES limit, the nonlocal terms vanish because $P_{NL}(u_* w_*^{-1}, \Delta x_*) = 0$, and the standard LES 3D closure is recovered. Second, the length scale ℓ in the local term is obtained by blending Deardorff's length scale for LES and the Mellor and Yamada (1982) level 3 length scale (which is generally used in mesoscale simulations), via local partition functions $P_L(\Delta x_*)$. More details can be found in Zhang et al. (2018). The tendency from the Zhang et al. (2018) scheme is the full divergence of the turbulent fluxes:

$$\left(\frac{\partial C}{\partial t}\right)_{Zhang} = -\frac{\partial}{\partial x_i}(\overline{u_i' C'}). \quad (5)$$

e. LES SGS models

In the LES innermost domain (d05), we calculate subgrid turbulent fluxes with two different widely used closures, the Deardorff (1980) TKE closure and the Smagorinsky (1963) and Lilly (1967) deformation-based closure. Both formulations are based on Eqs. (4) and (5) (with zero nonlocal fluxes), and their difference lies in the eddy viscosity/diffusivity calculations. In the TKE closure, $K_c = c_k \ell e^{1/2}$ and a prognostic equation for TKE is retained, along with additional assumptions to close the dissipation term. The deformation-based closure calculates K_c (for momentum) based on the norm of the deformation tensor (i.e., the amount of strain in the flow):

$$K_c = C_s^2 \ell_S^2 \max[0, (D^2 - P_r^{-1} N^2)^{1/2}], \quad (6)$$

where D is the norm of the deformation tensor, N is the Brunt–Väisälä frequency, P_r is the turbulent Prandtl number, and $\ell_S = (\Delta x \Delta y \Delta z)^{1/3}$ is the length scale in the Smagorinsky–Lilly model (which differs from ℓ in Deardorff's model). More details about the SGS closures (e.g., calculations of D and N^2 in hybrid coordinates) can be found in Skamarock et al. (2021).

3. Data analysis

a. Variance decomposition

The main goal of this study is to analyze the sensitivity of the model experiments. To this end, we use a variance decomposition technique to disentangle the effect of the four different input factors on the LES solution. Let A_{jklm} be an arbitrary variable from the LES in d05, either averaged in time and space or calculated at an instantaneous space–time point. The indices j, k, l,

and m define the ensemble member. Specifically, $j = \{1, 2\}$, $k = \{1, 2, 3\}$, $l = \{1, 2, 3\}$, and $m = \{1, 2\}$ represent the two options for topography and land-use factor (TL), the three options for the initial and boundary conditions (LBC), the three options for gray zone (GZ) modeling, and the two options for LES SGS modeling (SGS), respectively. We can define the sum of squares across the simulation ensemble SSE$_{TOT}$ as follows:

$$SSE_{TOT} = \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{m=1}^{M} (A_{jklm} - A_{....})^2, \quad (7)$$

where $J = 2$, $K = 3$, $L = 3$, and $M = 2$ and the dot subscript ($\cdot$) denotes averaging (i.e., $A_{....}$ represents the average across all experiments). After some algebra, it can be shown that SSE$_{TOT}$ can be decomposed into the sum of five different factors:

$$SSE_{TOT} = SSE_{TL} + SSE_{LBC} + SSE_{GZ} + SSE_{SGS} + SSE_R, \quad (8)$$

where SSE$_{TL}$, SSE$_{LBC}$, SSE$_{GZ}$, and SSE$_{SGS}$ are the portion of the total sum of squares related to the TL, LBC, GZ, and SGS, respectively, and SSE$_R$ is the residual unexplained sum of squares. Their definition is as follows:

$$SSE_{TL} = KLM \sum_{j=1}^{J} (A_{j...} - A_{....})^2, \quad (9)$$

$$SSE_{LBC} = JLM \sum_{k=1}^{K} (A_{\cdot k \cdot \cdot} - A_{....})^2, \quad (10)$$

$$SSE_{GZ} = JKM \sum_{l=1}^{L} (A_{\cdot \cdot l \cdot} - A_{....})^2, \quad (11)$$

$$SSE_{SGS} = JKL \sum_{m=1}^{M} (A_{\cdot \cdot \cdot m} - A_{....})^2. \quad (12)$$

In other words, Eqs. (9)–(12) measure the distance between the factor-specific averages and the overall average. For instance, the sum in Eq. (9) is the sum of two different terms, i.e., (i) the distance between the average of all simulations with coarse datasets (18 experiments) and the overall average and (ii) the distance between the average of all simulations with fine datasets (the other 18 experiments) and the overall average. A large value of SSE$_{TL}$ implies that the average across the simulations with fine-resolution datasets is considerably different from the average across the simulations with coarse datasets; hence, the model is sensitive to TL. Finally, the SSE$_R$ term can be further decomposed into 11 interaction terms, i.e., the sum of squares that derive from changing two or more terms simultaneously (6 two-way terms, 4 three-way terms, and 1 four-way term). As an example, the two-way interaction term related to TL and LBC is written as follows:

$$SSE_{TL,LBC} = LM \sum_{j=1}^{J} \sum_{k=1}^{K} (A_{jk\cdot\cdot} - A_{j...} - A_{\cdot k\cdot\cdot} + A_{....})^2. \quad (13)$$

The remaining interaction terms are similar and not reported here for conciseness but can be calculated similarly. Comparing

the magnitude of individual SSE terms with SSE_{TOT}, normalized by the total number of experiments, $JKLM = 36$ is used as a metric of sensitivity in analyzing the results.

b. Inline budget of horizontal momentum equations

To disentangle the numerical and physical mechanisms of the observed sensitivity, we study the budget of the horizontal momentum equations in the original coordinate system of ARW (x, y, η). In other words, we calculate the magnitude of each tendency term in the horizontal momentum equations to understand where the simulations differ. Let $\mathbf{v} = [u, v, w]$ represent the three velocity components in Cartesian coordinates and $\mathbf{V} = [U, V, W] = \mu_d \mathbf{v}$ be the flux-form momentum variables (conserved variables), where $\mu_d = \partial p_d / \partial \eta$ defines the vertical coordinate metric (p_d is the hydrostatic component of pressure of dry air). The x-direction flux-form horizontal momentum equation in WRF hybrid coordinates reads

$$\frac{\partial U}{\partial t} + (\nabla \cdot \mathbf{V}u) + \mu_d \alpha \frac{\partial p}{\partial x} + \frac{\alpha}{\alpha_d} \frac{\partial p}{\partial \eta} \frac{\partial \phi}{\partial x} = F_U, \qquad (14)$$

where p is the total pressure (including moisture, diagnostic in ARW), $\phi = gz$ is the geopotential (prognostic in ARW), and α_d and α are inverse densities of dry and moist air, respectively (diagnostic in ARW). In Eq. (14), we neglect map-scale factors that account for projections (because of the LES domain size). The advection operator is defined as

$$\nabla \cdot \mathbf{V}u = \frac{\partial Uu}{\partial x} + \frac{\partial Vu}{\partial y} + \frac{\partial \Omega u}{\partial \eta}, \qquad (15)$$

where $\Omega = \mu_d \omega$ and $\omega = d\eta/dt$. The first term in Eq. (14) is the total tendency term (TEND), the second term is advection (ADV), and the sum of the third and fourth terms is the horizontal pressure gradient (PGF). The forcing term F_U includes tendencies from model physics (PHYS), horizontal (HDIFF) and vertical (VDIFF) turbulent mixing, sixth-order numerical diffusion (6DIFF), and the Coriolis (COR) and curvature (CUR) terms. The Coriolis term is fV, where $f = 2\Omega_e \sin\psi$, where ψ is the latitude and Ω_e is the angular rotation rate of Earth. In the y direction, the horizontal momentum conservation equation is similar:

$$\frac{\partial V}{\partial t} + (\nabla \cdot \mathbf{V}v) + \mu_d \alpha \frac{\partial p}{\partial y} + \frac{\alpha}{\alpha_d} \frac{\partial p}{\partial \eta} \frac{\partial \phi}{\partial y} = F_V. \qquad (16)$$

The Coriolis term in the forcing term F_V is $-fU$, neglecting map-scale factor effects.

Closing the budget by postprocessing model output is difficult because of the complex numerics that are implemented in the ARW solver (Moisseeva and Steyn 2014). Instead, we implement an inline budget method directly into the ARW solver (version 4.4), which is conceptually similar to the 2D version presented in Chen et al. (2020). In short, additional variables that track the term-specific tendencies are implemented directly into the WRF code, and they are updated at runtime after their respective call during the Runge–Kutta (RK) time stepping loop (outside the acoustic substeps). As

an extension to Chen et al. (2020), we directly calculate the uncoupled (i.e., divided by μ_d) *accumulated* tendencies over the user-defined output interval to avoid decreasing the output interval to the model time step. To this end, we add a subroutine in the diagnostics module (which is called after all the three steps of the RK loop) that accumulates the term-specific tendencies across the user-defined output interval. The new subroutine integrates the instantaneous tendency over the time step duration and accumulates the tendency value over the output interval (i.e., the accumulated uncoupled momentum tendency units are m s^{-1}). Finally, note that the uncoupling procedure for the flux-form variables is significantly different than Moisseeva and Steyn (2014) and Chen et al. (2020), which use WRF3.4 and WRF3.8 and pure η coordinates (where μ_d represents the mass of dry air per unit area of a certain column); for hybrid coordinates, $\mu_d = \partial p_d / \partial \eta$ varies with height (i.e., it is a 3D variable in space), and its full definition (Skamarock et al. 2021) must be used. With the implemented inline budget procedure, we achieve an almost perfect budget closure, with differences between the total tendencies and the sum of all terms (ADV, PGF, COR, CUR, PHYS, HDIFF, VDIFF, and 6DIFF) on the order of 10^{-6} m s^{-1} over a 5-min accumulation interval (where typical tendencies are on the order of 0.1 m s^{-1}, as shown in the results section).

4. Results

a. Global sensitivity

We evaluate the sensitivity of wind speed, water vapor mixing ratio, and potential temperature in different locations across the LES innermost domain (d05), to highlight different physical processes occurring in d05. Specifically, we show results for four spatial locations in d05, two relatively near the domain boundaries (IN-NE and OUT-SW), one in the valley (VALLEY, collocated with tower tse09), and one location on the ridge (RIDGE, collocated with tower tse04). Their exact locations are shown in Fig. 1, along with the transect location that will be used in the next subsections.

The prevailing near-surface winds blow from the northeast during the simulation period, which makes the northeast location upwind of the double-ridge topography (*inflow*) and the southwest location downwind (*outflow*), although during the last day of the simulation, significant wind rotation is observed. All results in the four locations are calculated at 100 m AGL, which is in the range of a typical wind turbine hub height. In addition, we consider the spatial average on two isobaric surfaces aloft, i.e., 850 hPa (about 1500 m AGL) and 500 hPa (about 5500 m AGL), to assess the effect of the four input factors on quantities at the top of the ABL and well above the ABL.

Figure 2 shows the normalized SSE values [i.e., SSEs defined in Eqs. (7)–(12) divided by $N = JKLM = 36$] for the 4-day time-averaged wind speed ($A = \sqrt{u^2 + v^2}$). The largest total variance (i.e., spread across the ensemble) is observed in near-surface winds on the ridge, which is mostly explained by the topography and land-use factor. Physically, this result implies that the time-averaged winds in the 18 simulations that
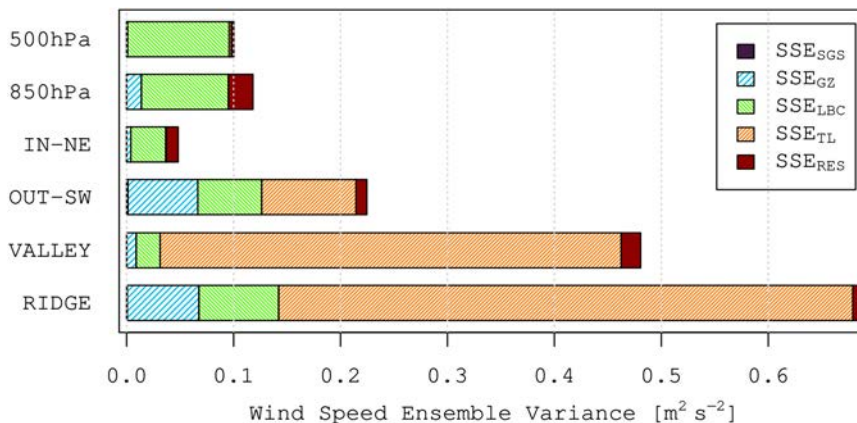
FIG. 2. Partition of the normalized SSE values for wind speeds into the four different input factors considered in this work, at the four different locations highlighted in Fig. 1 and two isobaric surfaces at 500 and 850 hPa.

use the coarse topography (*C** in our four-character naming system, where the wildcard * refers to all other possibilities in the other characters; see Table 1) and land-use datasets are considerably different than the time-averaged winds in the remaining 18 simulations (*F**), which use fine-resolution static datasets. The large differences in surface roughness and topography, described in section 2b, are important drivers for near-surface wind speed, both in the valley and on the ridge. This result underscores the importance of land-use datasets in complex terrain problems, where using the default dataset in WRF leads to significantly different results than using ad hoc high-resolution datasets. To gauge how large is the spread between *F** and *C** simulations, we consider the quantity $\sigma_{TL} = \sqrt{(JKLM)^{-1}SSE_{TL}}$, which on the ridge is equal to 0.73 m s^{-1}. Chebyshev's theorem states that 8/9 of the data lie within 3σ from the mean, which means that with ~88% confidence, we expect *C** and *F** simulations to differ less than $3\sigma_{TL} = 2.19$ m s^{-1}. On a 4-day averaged value, 2.19 m s^{-1} is relatively large and has significant wind power implications. The 3σ values for the other factors on the ridge are $3\sigma_{GZ} = 0.78$ m s^{-1}, $3\sigma_{LBC} = 0.82$ m s^{-1}, and $3\sigma_{SGS} = 0.02$ m s^{-1}, which imply that 100-m winds on the ridge are insensitive to the SGS model but relatively sensitive to the choice of lateral boundary conditions and gray zone model (to a lower extent than topography and land-use datasets). Valley results are similar to the ridge, whereas the inflow and outflow locations show less ensemble variance compared to the two other locations. The variance partitioning is also different, with lateral boundary conditions becoming relatively more important in IN-NE and OUT-SW locations. Notice that the results refer to *time-averaged* winds; as we show in the next subsection, locally other factors can dominate. The gray zone influence on the ridge, which is negligible in the IN-NE location, implies that locally (in time) there can be differences even in the IN-NE location that, once amplified by the topography, can be relevant on the ridge, even on average. We will revisit the local dependence in the next subsection. Aloft, the variance is lower and is mostly explained by the choice of the initial and lateral (re)analysis, as all the other

factors (GZ, SGS, and TL) are expected to have an influence on boundary layer processes.

Interesting results are obtained from a similar analysis on moisture (Fig. S5), where the partition between the different factors is more homogeneous across the four locations compared to wind speed. The largest ensemble variance is found at the isobaric surface of 850 hPa, where $3\sigma_{GZ} = 0.44$ g kg^{-1}. This result is somewhat surprising but can be explained by analyzing the time evolution of moisture during the 4 days of simulation (next subsection). The large spread at 850 hPa is related to water vapor condensation, which occurs in some simulations (and consequently drops the water vapor mixing ratio) and does not in other simulations. This interesting effect is caused by the nonlinearity of the problem, where a small change in moisture transport due to different settings of the simulations can cause condensation during the last day of the simulation and therefore large differences across the ensemble members. At 500 hPa, only the initial and boundary conditions influence the amount of water vapor in the model calculations, as expected.

Overall, a limited spread is found for potential temperature, with the maximum 3σ value equal to 0.75 K on the ridge (Fig. S6). The potential temperature decomposition of variance resembles quite closely the water vapor mixing ratio one presented in Fig. S5. Similarly to the moisture results, the gray zone factor is most relevant at 850 hPa, suggesting a different treatment of turbulent transport at the entrainment zone in the d03 and d04 domains. Topography and land-use datasets influence the average temperature predictions close to the surface, with surprisingly limited influence in the valley. This is likely related to a compensating effect during day and night, which makes the 4-day time-averaged temperatures similar across *C** and *F** simulations. Overall, the large-scale reanalyses explain a significant amount of variance at all locations, becoming dominant at 500 hPa as for moisture and winds. The SGS model of the LES explains a negligible fraction of the total variance at all locations.

Finally, we compare the amount of variability in the model ensemble (quantified by the model spread, i.e., the difference
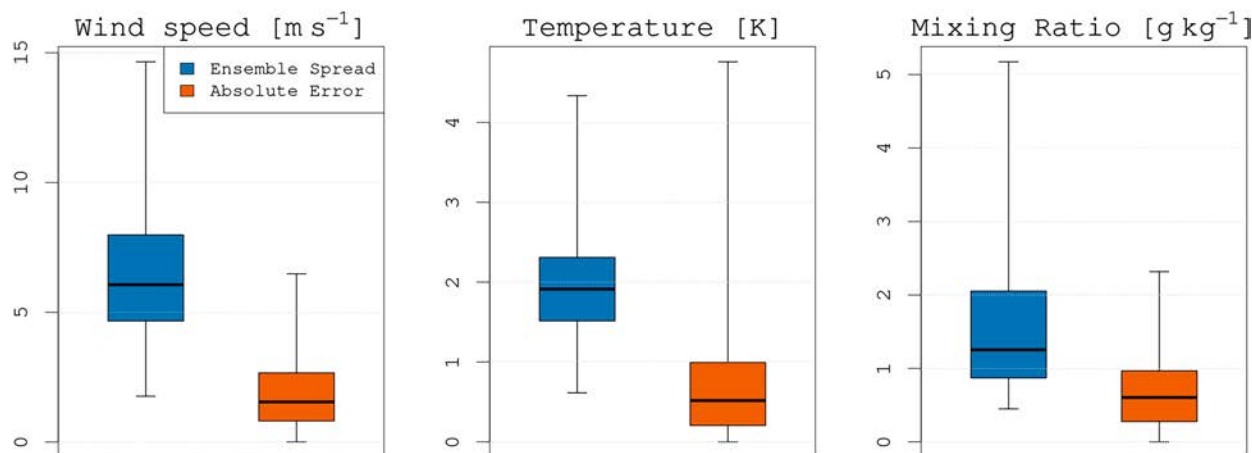
FIG. 3. Time distribution of the ensemble spread, defined as the difference between maximum and minimum values of the ensemble at a given time, and the absolute error, defined as the difference between the ensemble average and collocated tower measurements, (left) for wind speed, (center) temperature, and (right) mixing ratio.

between maximum and minimum values in the ensemble at a given time point) and the absolute difference between the ensemble average and corresponding tower measurements at the RIDGE location. Figure 3 shows the time distribution of these two quantities and indicates that the model spread is consistently larger than the difference between the ensemble average and the corresponding measurements. This comparison indicates that the model uncertainty is on average larger than the ensemble error, thus highlighting that constraining the model uncertainty could potentially result in better model performance. The corresponding time series, where the model ensemble is directly compared against the tower measurements, can be found in Fig. S7.

b. Time-resolved sensitivity

To further elucidate the sensitivity of the ensemble to the different factors, we repeat the same calculations on 15-min averages to reconstruct the entire time series of variance partitioning. Figure 4 shows one example of the time-resolved approach, for 100-m winds on the ridge. Interesting results emerge from the time-resolved analysis. The topography and land-use datasets, which dominate the variance for the ridge winds, do not always explain the largest amount of variance, and at times, simulations are insensitive to TL. The reason why TL becomes dominant on the 4-day averages is the consistent difference between the two sets of simulations, i.e., winds are almost always larger in *F** compared to equivalent *C** simulations, because of the sharper ridge and less rough terrain. Other factors, like LBC and GZ, produce large spread locally that compensate over the 4 days of simulations (i.e., E*** simulations can differ from equivalent G*** and W*** simulations locally, but the sign of the difference flips during the 4 days of simulations, whereas it does not for the difference between *C** and *F**; see Table 1 for naming convention). Other interesting considerations can be deduced from Fig. 4. First, the largest ensemble variances are observed at night, or during the afternoon-to-evening transition or the

morning transition. Mixing during daytime typically reduces the spread among the ensemble, whereas nighttime stratification leads to the opposite effect. The small variances at the end of the simulation are related to weak synoptic conditions (Fig. S2) that lead to lower winds in all simulations and less spread among the ensembles. Note that the local $3\sigma_{TOT}$ values can be as large as 10 m s^{-1} (e.g., during the nights between 19–20 May and 20–21 May), which for wind energy forecasts is a massive difference. In other words, depending on the model configuration, simulations can predict wind speeds
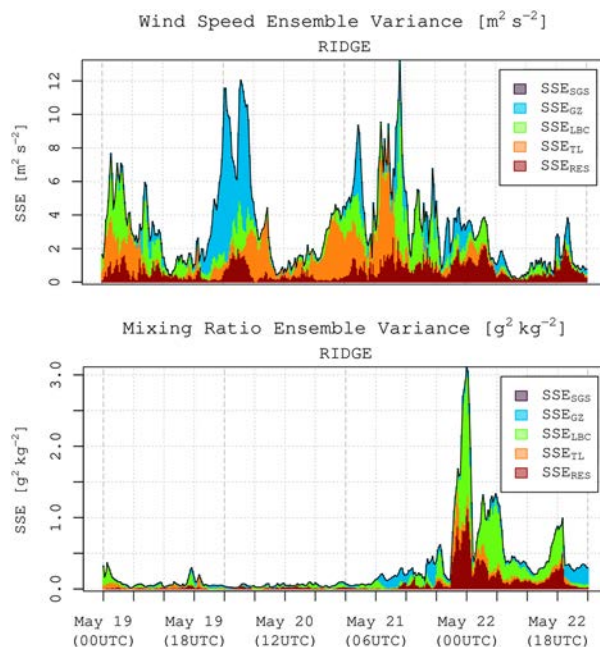

FIG. 4. Time-resolved partition of the normalized SSE values for (top) wind speed and (bottom) mixing ratio at the RIDGE location (100 m AGL).
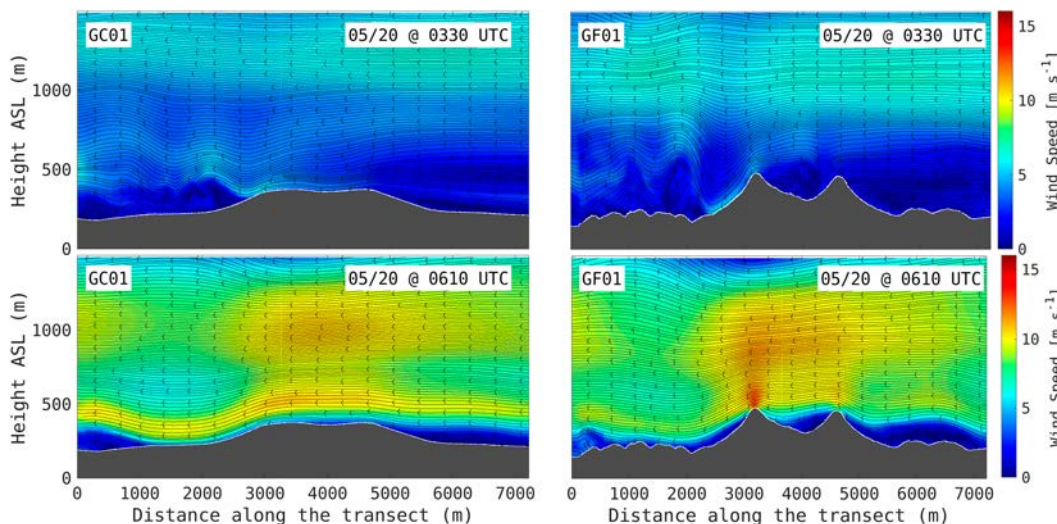
FIG. 5. Streamlines and horizontal velocity magnitude (jet shading) in the 2D transect shown in Fig. 1, at two selected times, for a coarse TL simulation (GC01) and its equivalent fine TL simulation (GF01). Gray fill indicates topography.

on the ridge (where the wind turbine is) that are 10 m s^{-1} different from each other. During the night of 19/20 May, most of this sensitivity is explained by the GZ factor, i.e., the way that turbulence is modeled in the parent domains (d03 and d04) can lead to enormous differences in the LES domain. On the other hand, the SGS closure in the LES domain explains virtually zero variance of the large local spread. We will analyze the strong and weak sensitivity to GZ and SGS during the first night in more detail in the next subsections.

The time-resolved partition of variance for the water vapor mixing ratio is also shown in Fig. 4. An almost mirrored pattern emerges compared to the wind sensitivity, as the largest variance is found toward the end of the simulation time. The large spread observed toward the end of the simulation (peak $3\sigma_{TOT} = 5.19$ g kg^{-1}) is the effect of the synoptic wind shift during the last day of the simulation that brings additional moisture to the Perdigão Valley. As a result, some ensemble members reach supersaturation humidity and condense water vapor with the microphysics scheme, whereas other simulations reach values of relative humidity that are close to supersaturation but not enough to start condensing (not shown). This difference is largely related to LBC and interaction terms, which is reasonable given the nonlinearity of the problem (i.e., a small change in humidity can trigger condensation and large discrepancies between simulations). During the rest of the simulation period (roughly 19–22 May), relative humidity is considerably lower than 100% in all simulations, and differences between mixing ratios remain limited compared to the large peak at the end of the simulation. The time series of the water vapor mixing ratio for the entire ensemble is shown in Fig. S8. For ease of interpretation, a mixing ratio of 6 g kg^{-1} corresponds to approximately 48% relative humidity assuming a temperature of 290 K (typical at night) and a pressure of 960 hPa (typical on the ridge); a mixing ratio of 12 g kg^{-1} corresponds to approximately 95% relative humidity with the

same assumptions. In the simulations where water vapor drops substantially during the night of 21/22 May, low-level clouds and fog (i.e., clouds at the lowest model level) are produced by the microphysics parameterization (not shown).

Finally, the time-resolved partition of variance for potential temperature (not shown) indicates a rather small sensitivity of temperature, with peaks that are below 1.0 K^2, mostly dominated by TL and LBC.

c. Sensitivity to topography and land-use datasets

To understand the topography and land-use influence in more detail, we look at sensible and latent heat fluxes for temperature and moisture sensitivities (Figs. S9 and S10) and analyze the flow in two specific ensemble members for the wind sensitivity (GF01 and GC01). The full ensemble can be found in Fig. S11. The only input differences in GC01 and GF01 are the TL datasets in d03, d04, and d05. Figure 5 shows the difference between GC01 and GF01 flow fields in a 2D transect at two selected times (0330 and 0610 UTC). The streamlines in Fig. 5 are obtained by rotating the wind vector in an orthogonal coordinate system where one of the directions is parallel to the transect. As the wind direction is almost parallel to the transect (or perpendicular to the ridges) during the selected times, the component of the wind into (or out of) the page is small but nonzero, i.e., there is a normal flow component not depicted in the 2D view. During the first selected time, horizontal winds at the ridge location are similar (which is about 3000 m along the transect), although striking differences still exist between the two calculated flows. First, recirculation and cavity zones are observed upstream and within the valley in GF01, whereas they are not produced with the hill-like smoothed topography in GC01. The downslope flow off the southern ridge in GF01, as well as the meandering pattern created by the topography disturbance, is also remarkably different in GC01, which instead shows a distinct wake
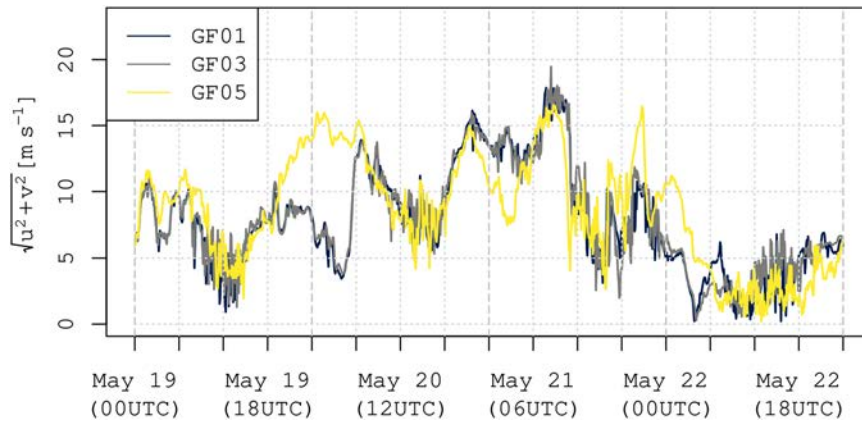
FIG. 6. Wind speed time evolution (d05) at the RIDGE location (100 m AGL) for GF01, GF03, and GF05.

zone and less wave activity. During the second selected time with a low-level jet interacting with the double ridge, the horizontal wind speed magnitude on the southern ridge is considerably different, which is likely related to the different flow accelerations driven by the different pressure fields that develop around the topography (Safaei Pirooz and Flay 2018; Hunt et al. 1988).

d. Sensitivity to gray zone modeling

We analyze two nights (19/20 May and 21/22 May) as interesting episodes of large sensitivity of the LES solution to gray zone modeling. As shown in Fig. 4, differences in wind speed at the RIDGE location can be as large as 10 m s^{-1}, and we consider how such large differences are generated. Figure 6 shows the time series of winds at the RIDGE location for three ensemble members (GF01, GF03, and GF05) in d05. The only difference between the three ensemble members is gray zone modeling in d03 and d04 (i.e., the three d05 simulations are perfectly equivalent except for the boundary conditions). GF01 uses the YSU scheme, GF03 its scale-aware version (Shin–Hong), and GF05 the 3D Zhang scheme. Qualitatively similar results are obtained by analyzing EF01, EF03, and EF05 as well as WF01, WF03, and WF05. First, the sensitivity to gray zone modeling is almost entirely related to the differences between the 3D Zhang and the 1D schemes in d03 and d04, while YSU and its scale-aware version produce very similar time series of winds at RIDGE. The magnitude of the differences during the night between 19 and 20 May is striking, with GF05 predicting maximum winds of 16 m s^{-1} and the 1D schemes predicting minimum winds of 3.5 m s^{-1}, almost at the same time.

The 2D flow transects in Fig. 7 offer an interesting explanation of these differences. The inflow is strikingly different. A low-level jet enters the domain d05 with peak winds of 10 m s^{-1} and interacts with the topography in GF05, amplifying the maximum winds to about 15 m s^{-1} on the southern ridge. In GF01, the inflow carries significantly lower momentum and the interaction with fine-scale topography does not amplify the flow on the ridge to the same level.

We analyze horizontally averaged vertical eddy viscosities K_V and the horizontal momentum budget in d04 to further understand the reasons that lead to such different inflow and the generation of the low-level jet. As expected, average K_V calculated by the YSU and Zhang's schemes are largely different (Fig. S12), given their substantially different formulation based on a diagnostic profile and prognostic TKE, respectively (see section 2d). During the day, average K_V in the YSU's scheme are much larger than those in Zhang's scheme, although their overall impact on the mean flow differences is not as important as during day-to-night transitions and nighttime periods. This occurs for a handful of reasons: (i) during the day, the mixed-layer velocity gradient is close to
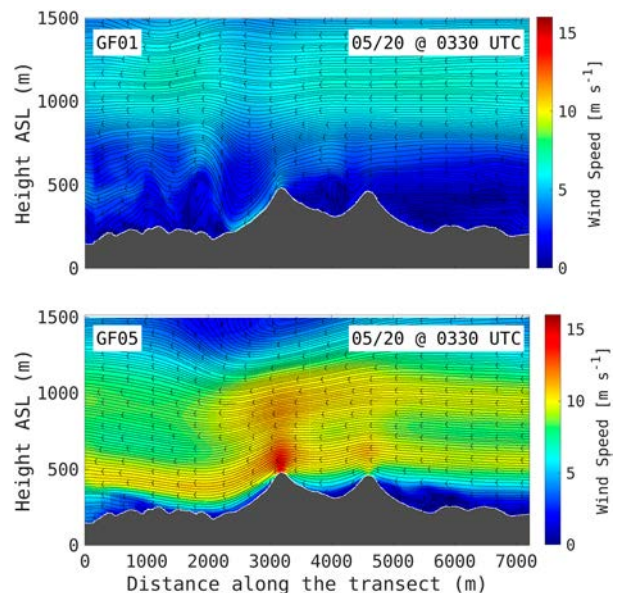

FIG. 7. Streamlines and horizontal velocity magnitude (jet shading) in the 2D transect shown in Fig. 1, at two selected times, for a multiscale simulation using YSU in the gray zone (GF01) and a second one using Zhang's blending closure (GF05). Gray fill indicates topography.
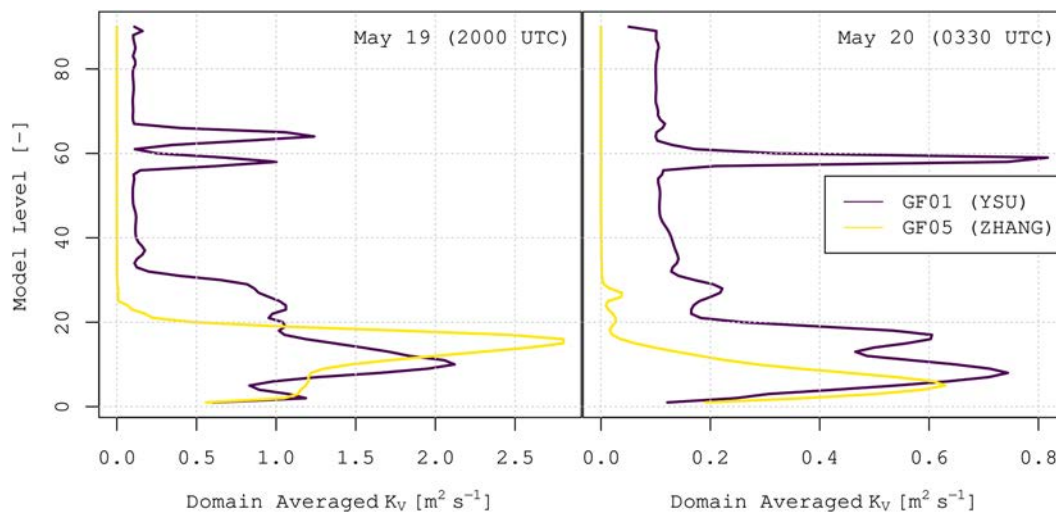
FIG. 8. Vertical profiles of domain-averaged vertical eddy viscosities in the gray zone domain (d04) for GF01 (YSU) and GF05 (Zhang), at (right) 2000 UTC 19 May and (left) 0330 UTC 20 May.

zero, and therefore, different eddy viscosities matter to a lower extent, (ii) part of the vertical transport is done by the nonlocal terms in both YSU and Zhang's schemes, and (iii) lower eddy viscosities allow for more explicitly resolved convection that compensates the lower subgrid mixing. However, the transition and nighttime periods (when $\partial u/\partial z$ and $\partial v/\partial z$ are significantly different than zero) are characterized by differences in K_V that create a significant spread in the simulated mean flow (Fig. 6). Such differences can occur for several different reasons: (i) different definitions of the PBL height in Zhang's and YSU schemes, as shown in more detail in Fig. S13; (ii) no eddy viscosity limiter in Zhang's scheme, with values lower than 0.1 m^2 s^{-1} allowed at night, as opposed to YSU that limits K_V to 0.1 m^2 s^{-1}; and (iii) virtually zero eddy viscosity in the free atmosphere in Zhang's scheme. A combination of all three is likely the culprit for the very different behaviors during nighttime. Specifically, the boundary layer height drops significantly at night (especially in Zhang's scheme, Fig. 8), implying that the majority of vertical model levels experience virtually zero eddy viscosity in Zhang's scheme, whereas YSU calculates nonzero eddy viscosities according to its free atmospheric diffusion scheme even above the boundary layer (Fig. S14). The interface (and shear) between the stable boundary layer and the residual layer becomes very different in the two schemes (as well as the growth of the stable boundary layer), with eddy viscosities at the interface spanning order of magnitudes in the two different schemes, likely generating the different inflow features highlighted in Fig. 7.

Accordingly, the horizontal momentum budget analysis (Fig. 9) shows that the differences in southward acceleration of the flow during the 19/20 May night occur mostly because of differences in the horizontal pressure gradient, which suggests that the presence of inhomogeneous terrain, combined with the different vertical transport of heat and differential nighttime cooling in the two schemes, is what is ultimately responsible for those large differences. In other words, the

effects of the different gray zone closures seem to be of two different types. First, vertical diffusion (or physics in the case of PBL schemes) provides a *direct* effect through its tendency term to the flow at a specified grid point. However, there is also an *indirect*, nonlinear effect that is caused by vertical transport of momentum, heat, and moisture (or lack thereof) which can locally alter pressure gradients, advection, and Coriolis terms (e.g., that can lead to the formation of low-level jets), ultimately affecting the mean flow as well. A combination of the direct and indirect effects is evident in all the nighttime periods and is different considering the time period (Fig. 9). A preliminary comparison with observations (Fig. S15) suggests that Zhang's scheme (without the K_V limiter) captures the transitions better than YSU (and its scale-aware version), in our specific case study. However, more detailed analyses will be presented in a future development of this work to specifically focus on the gray zone sensitivity and look at (i) detailed comparisons with observations, to ensure that results are obtained because of the correct physical reasons and not because of a compensation of different effects and (ii) further identifying what are the conditions that are conducive for strong differences to be formed between the two schemes.

e. Sensitivity to lateral boundary conditions

We turn our attention to how differences in initial and lateral boundary conditions in d01 influence the sensitivity in the innermost LES domain (d05). From the time-averaged sensitivity estimates presented in Fig. 2 and Figs. S5 and S6, it emerges that LBCs dominate the sensitivity for quantities in the middle to upper troposphere, as the three other factors that we considered primarily modulate ABL processes. However, LBCs also have a nonnegligible influence on LES-calculated ABL quantities. This becomes more evident when considering time-resolved sensitivities (Fig. 4). At specific time instances, LBC is the primary factor explaining the ensemble variance, especially for the moisture content. Specifically, such large
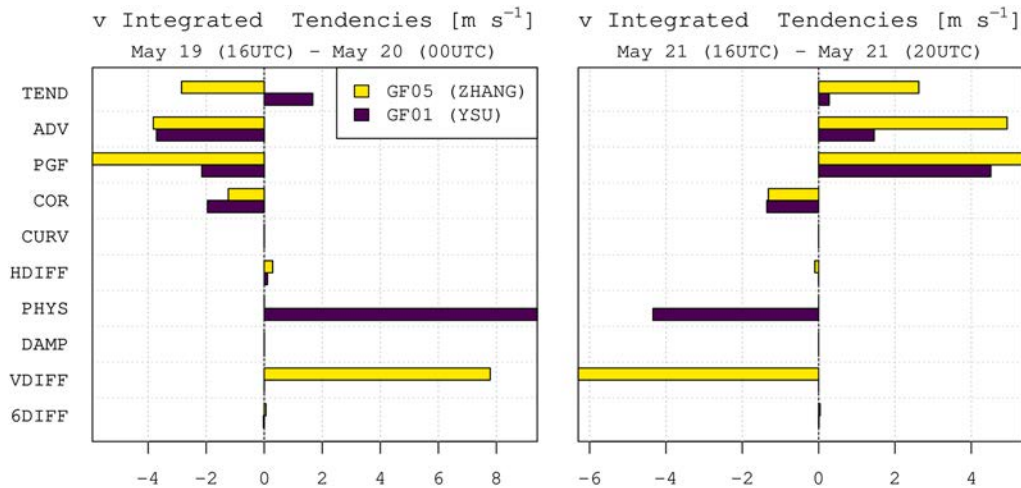
FIG. 9. Inline budget analysis of the horizontally averaged v component of horizontal momentum, for GF01 and GF05 (d04) during two different time periods of high gray zone sensitivity, at the fifth vertical model level (approximately 100 m above ground). The different terms are defined in section 3b and are integrated over the two different time periods directly within the WRF code. TEND therefore represents the net change in v from the start to the end time. (left) Note that for the first time period, v is negative, and so negative tendencies imply accelerations; (right) the opposite is true for the second time period.

sensitivity is observed when synoptic scale features change, e.g., when moisture is advected toward the Perdigão Valley on 21 and 22 May, as well as when the prevailing wind direction significantly shifts. These patterns are described in more detail in De Moliner et al. (2023, manuscript submitted to *Appl. Energy*), whereas here we focus on the specific example of moisture sensitivity, to explain how the model ensemble locally shows $3\sigma_{TOT} = 5.19$ g kg^{-1} at RIDGE (Fig. 4). Using a similar budget technique as for the horizontal momentum budget and the advection operator defined in Eq. (15), we extract the moisture advection term for the Perdigão grid point in d01 and integrate over the lower ABL atmospheric column and the output time interval:

$$QADV(t_* + \Delta\tau) = \int_{t_*}^{t_*+\Delta\tau} \frac{\mu_d}{g} \int_{\eta=1}^{\eta=\eta_{BL}} \left(u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} + \Omega\frac{\partial q}{\partial\eta} \right) d\eta \, dt,$$
(17)

where t_* is the beginning of the output time interval, $\Delta\tau$ is the output time interval, and q is the water vapor mixing ratio. In the d01 simulation, the output time interval $\Delta\tau$ is set to 1 h. The quantity in Eq. (17) represents the mass of water (kg per output time interval) carried into (or away from) the atmospheric column delimited by $\eta = 1$ and $\eta = \eta_{BL} = 0.95$ at the nearest grid point to the Perdigão area. Different analysis products (GDAS, ERA5, and HRES-ECMWF) show interesting differences in QADV in d01 to the Perdigão area that can be linked to what we observe in the fine-scale LES solution (Fig. 10). First, ERA5 data show a rather sharp decrease in cumulated moisture advection during the 21 and 22 May night, which is not present in GDAS and HRES-ECMFW. The water vapor mixing ratios in the LES solutions driven by ERA5 also show a significant drop during the same period,

which suggests that the large-scale dynamics is what drives the moisture drop observed in the LES solution driven by ERA5 (green lines in Fig. S8). Similarly, the more limited and later drop in cumulated advection in the GDAS simulation (i.e., moisture carried away from Perdigão) is also reflected to a lower extent in the LES driven by the GDAS analysis. That period is also the sensitivity peak in Fig. 4, which is accordingly explained by the LBC factor from the variance analysis. We speculate that the reason why ERA5 data show a temporary change of sign in moisture advection is the higher temporal resolution of ERA5 (hourly) compared to GDAS and HRES-ECMWF (6-hourly). In other words, when large-scale conditions change on a fine time scale (e.g., during transition periods), the sensitivity to LBC is enhanced because of the discrepancies in the time resolution of the boundary conditions. This seems to be the case not only for moisture but also for winds, which show large sensitivity during rapid changes in large-scale conditions.
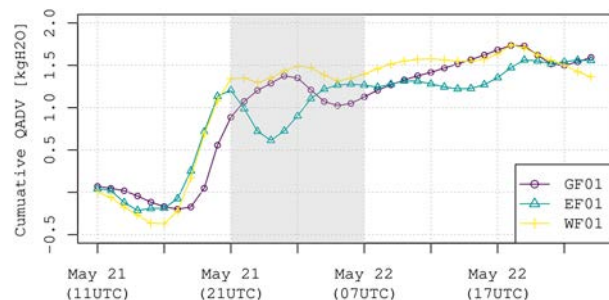

FIG. 10. Cumulated QADV [Eq. (17)] in d01 simulations at the grid point nearest to the Perdigão area. Gray shading represents the period of interest described in the text where moisture is advected to the area.
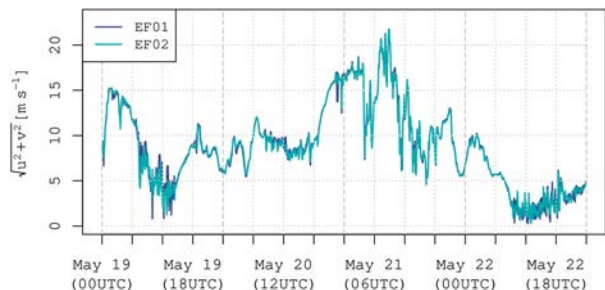
FIG. 11. Wind speed time evolution at the RIDGE location (100 m AGL) for EF01 (TKE scheme) and EF02 (Lilly–Smagorinsky scheme) in d05.

f. Weak sensitivity to the LES subgrid-scale models

An interesting result from our analysis is that the downscaled LES solution is rather insensitive to the two subgrid horizontal and vertical diffusion schemes in the LES domain. Figure 11 presents the RIDGE wind time series of two ensemble members (EF01 and EF02) that only differs for the SGS model in d05, and in the supplemental material, we present the corresponding vertical velocities cross sections and power spectra (Figs. S16 and S17). Note that the comparison would be qualitatively similar for any other pair of simulations that only differs for the SGS model (e.g., WC01 vs WC02, EF03 vs EF04, GF05 vs GF06, and others; see Table 1). The two time series (and daytime spectra) are nearly equivalent, which explains the extremely low SGS variance values that we presented in Figs. 2 and 4. Moisture and temperature also behave similarly (not shown). We analyze the horizontal momentum budget as a starting point to understand why the two different SGS models produce nearly identical results for

winds well within the ABL (100 m above ground). Figure 12 shows the d05 horizontal momentum budget for the same two time periods that we analyzed in the previous subsection, where the total tendency in v (i.e., accumulated change) is not negligible. Interestingly, the total vertical diffusion tendency is different in the two different schemes but is compensated by an equal and opposite difference in the advection term (i.e., resolved vertical motion) tendency. In other words, the amount of subgrid mixing in EF01 is significantly lower than that in the Smagorinsky case (EF02), owing to different eddy viscosity values in the ABL (Fig. S18), but allows for more energetic vertical motions that compensate the lower subgrid mixing. The total vertical mixing effect is therefore similar and produces total tendencies that are comparable. This effect is qualitatively similar to the *implicit scale-awareness* described in Giani et al. (2022), where the same phenomenon occurs at gray zone resolutions. There are three important notes to be added to this analysis. First, note that one could conclude that the LES subgrid-scale model could be neglected altogether from Fig. 11; however, our analysis only suggests that the two subgrid-scale models provide, through different mechanisms, a similar *total* amount of turbulent stress, which is a very important physical driver for boundary layer winds and cannot be neglected (Fig. 12). Second, the EF01 and EF02 eddy viscosity vertical profiles show nearly identical values behavior above the boundary layer in the TKE and the Lilly–Smagorinsky scheme (Fig. S18), which is in stark contrast with our discussion of the gray zone sensitivity, where YSU and Zhang's schemes differ the most above the boundary layer. This is because both LES SGS models limit the eddy viscosity to a minimum of $0.1\Delta z^2/\Delta t$, which becomes active for lower eddy viscosities that would be calculated above the boundary layer. That is an additional reason why low sensitivities are observed,
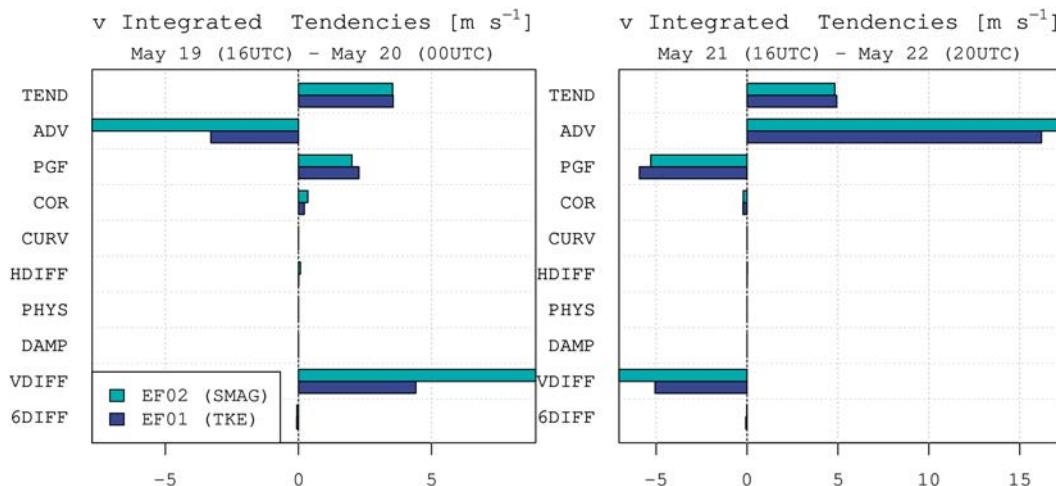

FIG. 12. Inline budget analysis of the horizontally averaged v component of horizontal momentum, for EF01 and EF02 (d05) during two different time periods of interesting changes in v, at the fifth vertical model level (approximately 100 m above ground). The different terms are defined in section 3b and are integrated over the two different time periods directly within the WRF code. TEND therefore represents the net change in v from the start to the end time. (left) Note that for the first time period, v is negative, and so negative tendencies imply accelerations; (right) the opposite is true for the second time period.

compared, for instance, to the gray zone sensitivity. Finally, we have also performed a preliminary test with the less dissipative nonlinear backscatter and anisotropy (NBA) scheme of Kosović (1997) and Mirocha et al. (2010) and found that the NBA scheme produces similar results to the TKE and the Lilly–Smagorinsky one (i.e., the multiscale simulations seem insensitive to the NBA scheme as well, Fig. S19).

5. Discussion and conclusions

In this work, we present a large new dataset of 36 LES experiments that simulate the atmospheric flow at the Perdigão field site, with realistic bottom boundary conditions and time- and space-varying lateral boundary conditions from gray zone and mesoscale simulations. The goal of this work is to present the new ensemble of simulations and to study its global sensitivity to different input factors, i.e., topography and land-use static datasets, modeling turbulence at gray zone resolutions, modeling SGS turbulent fluxes at LES resolutions, and the initial and boundary conditions from global-scale (re)analyses. In other words, we aim to provide a more formal and general sensitivity analysis to complement previous literature that targeted specific individual factors (Liu et al. 2020; Doubrawa and Muñoz-Esparza 2020; Talbot et al. 2012; Kitagawa et al. 2022; Wang et al. 2020) or that focused on flat and homogeneous terrain with doubly periodic boundary conditions, such as the LASSO work (Gustafson et al. 2020). In addition, we isolate a few ensemble members to study the mechanisms that generate the model sensitivity for the four individual factors.

From the analysis presented in this work, we find large sensitivities in wind calculations on the ridge and in the valley of the Perdigão site. Most of the time-averaged wind discrepancy between experiments is explained by the different topography and land-use datasets, although turbulence modeling in the gray zone and the initial analyses are important. Note that the sensitivity to topography and land-use datasets is expected due to the somewhat unrealistic nature of the default WPS datasets when applied to a 30-m resolution grid. However, we aimed to emphasize the significant impact that using the default WPS datasets in high-resolution applications can have and stress the importance of using ad hoc land surface datasets in high-resolution simulations. Discrepancies for the 4-day average 100 m AGL wind speed can be as large as 2.5 m s^{-1} on the ridge. For a flat terrain case, we expect the model sensitivity to TL datasets to be lower, although the different roughness in different datasets can still play a large role for near-surface quantities. The model appears insensitive to the SGS model choice (Smagorinsky–Lilly and Deardorff TKE) in the innermost domain, with our WRF multiscale configuration. Without averaging in time, model discrepancies can be as large as 10 m s^{-1} on the ridge, and locally, the gray zone problem can dominate the model sensitivity. We further study how different gray zone models can produce large discrepancies in the LES nested domain, and we find that in the gray zone domains, nighttime stratification can be significantly different using Zhang's 3D blending closure scheme and the 1D YSU column-based scheme. This is an unexpected result given that most of the gray zone literature has focused on the daytime convective boundary

layer. We show that the lack of an eddy viscosity limiter in Zhang's scheme, combined with nonhomogeneous terrain, produces the largest differences with YSU at night. We underscore the importance of extensively testing new gray zone schemes beyond the idealized convective boundary layer canonical simulations since other aspects (such as diffusion in the free atmosphere at night) have a considerable impact on realistic predictions.

Several other research questions can be investigated from this dataset. Model verification and evaluation with flux towers and radiosonde measurements is part of current ongoing work and will be presented in a future study, as it exceeds the scope of this research which is focused on the sensitivity of model simulations. Evaluation with radiosondes and flux towers offers an interesting opportunity to further constrain the gray zone problem with measurements, as both 1D and 3D schemes have been used in the ensemble and can be evaluated with high-frequency flux tower measurements. Specifically, more in-depth analysis is needed to better characterize the differences between specific ensemble members, e.g., the ones using Zhang's gray zone scheme against 1D PBL parameterizations. Other aspects, including (i) the influence of steep terrain representation with terrain-following hybrid coordinates, (ii) the implications of our findings for wind power production and wind resources forecast, and (iii) a detailed spatial analysis of the observed sensitivity, are all subjects of ongoing work with this dataset. Additional in-depth analyses of specific episodes could also provide more insights on the numerical and physical mechanisms of the observed sensitivity. For instance, investigating the pathway to cloud and fog formation during the last day of simulations can be helpful in furthering our understanding of important local-scale moist dynamics, as well as its interaction with larger-scale processes. Finally, we note that our findings are valid for our model configuration at the Perdigão field site, but they serve as useful guidance to understand the sensitivity of realistic multiscale experiments in complex terrain and the mechanisms that generate such sensitivity. Extension of this work in other regions may help identify the generalizability of our findings.

Data availability statement. The numerical model simulations upon which this study is based are too large to archive or to transfer. Instead, we provide all the information needed to replicate the simulations in the manuscript. The namelists of the simulation ensemble are made publicly available through GitHub (https://github.com/Env-an-Stat-group/23.Giani.MWR), along with the modified FORTRAN files in WRFv4.4 for the horizontal inline momentum budget implementation. Initial and boundary conditions data for WRF simulations are obtained from NCAR's Research Data Archive, https://doi.org/10.5065/D65Q4T4Z, https://doi.org/10.5065/D68050ZV, and https://doi.

org/10.5065/BH6N-5N20 for GDAS, HRES-ECMWF, and ERA5 data, respectively. SRTM data are obtained from USGS (https://doi.org/10.5066/F7F76B1X), and CORINE land-cover data are obtained from Copernicus [Corine Land Cover (CLC) 2018, version 2020_20u1]. Specific output from the numerical simulations is made available upon request.

REFERENCES

Arthur, R. S., K. A. Lundquist, and J. B. Olson, 2021: Improved prediction of cold-air pools in the Weather Research and Forecasting Model using a truly horizontal diffusion scheme for potential temperature. *Mon. Wea. Rev.*, **149**, 155–171, https://doi.org/10.1175/MWR-D-20-0234.1.

Bao, J., F. K. Chow, and K. A. Lundquist, 2018: Large-eddy simulation over complex terrain using an improved immersed boundary method in the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **146**, 2781–2797, https://doi.org/10.1175/MWR-D-18-0067.1.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Beare, R. J., 2014: A length scale defining partially-resolved boundary-layer turbulence simulations. *Bound.-Layer Meteor.*, **151**, 39–55, https://doi.org/10.1007/s10546-013-9881-3.

Beck, J., and Coauthors, 2020: An evaluation of a hybrid, terrain-following vertical coordinate in the WRF-based RAP and HRRR models. *Wea. Forecasting*, **35**, 1081–1096, https://doi.org/10.1175/WAF-D-19-0146.1.

Chen, T.-C., M.-K. Yau, and D. J. Kirshbaum, 2020: Towards the closure of momentum budget analyses in the WRF (v3.8.1) model. *Geosci. Model Dev.*, **13**, 1737–1761, https://doi.org/10.5194/gmd-13-1737-2020.

Ching, J., R. Rotunno, M. LeMone, A. Martilli, B. Kosovic, P. A. Jimenez, and J. Dudhia, 2014: Convectively induced secondary circulations in fine-grid mesoscale numerical weather prediction models. *Mon. Wea. Rev.*, **142**, 3284–3302, https://doi.org/10.1175/MWR-D-13-00318.1.

Chow, F. K., C. Schär, N. Ban, K. A. Lundquist, L. Schlemmer, and X. Shi, 2019: Crossing multiple gray zones in the transition from mesoscale to microscale simulation over complex terrain. *Atmosphere*, **10**, 274, https://doi.org/10.3390/atmos10050274.

Connolly, A., L. van Veen, J. Neher, B. J. Geurts, J. Mirocha, and F. K. Chow, 2021: Efficacy of the cell perturbation method in large-eddy simulations of boundary layer flow over complex terrain. *Atmosphere*, **12**, 55, https://doi.org/10.3390/atmos12010055.

Daniels, M. H., K. A. Lundquist, J. D. Mirocha, D. J. Wiersema, and F. K. Chow, 2016: A new vertical grid nesting capability in the Weather Research and Forecasting (WRF) Model. *Mon. Wea. Rev.*, **144**, 3725–3747, https://doi.org/10.1175/MWR-D-16-0049.1.

Deardorff, J. W., 1970: Convective velocity and temperature scales for the unstable planetary boundary layer and for Rayleigh convection. *J. Atmos. Sci.*, **27**, 1211–1213, https://doi.org/10.1175/1520-0469(1970)027<1211:CVATSF>2.0.CO;2.

——, 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495–527, https://doi.org/10.1007/BF00119502.

De Meij, A., and J. F. Vinuesa, 2014: Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. *Atmos. Res.*, **143**, 351–370, https://doi.org/10.1016/j.atmosres.2014.03.004.

Doubrawa, P., and D. Muñoz-Esparza, 2020: Simulating real atmospheric boundary layers at gray-zone resolutions: How do currently available turbulence parameterizations perform? *Atmosphere*, **11**, 345, https://doi.org/10.3390/atmos11040345.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

EEA, 2002: Corine land cover 2000: Technical guidelines. European Union, Copernicus Land Monitoring Service Tech. Rep., 56 pp.

Efstathiou, G. A., and R. S. Plant, 2019: A dynamic extension of the pragmatic blending scheme for scale-dependent sub-grid mixing. *Quart. J. Roy. Meteor. Soc.*, **145**, 884–892, https://doi.org/10.1002/qj.3445.

——, ——, and M.-J. M. Bopape, 2018: Simulation of an evolving convective boundary layer using a scale-dependent dynamic Smagorinsky model at near-gray-zone resolutions. *J. Appl. Meteor. Climatol.*, **57**, 2197–2214, https://doi.org/10.1175/JAMC-D-17-0318.1.

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

Emanuel, K. A., 1995: Sensitivity of tropical cyclones to surface exchange coefficients and a revised steady-state model incorporating eye dynamics. *J. Atmos. Sci.*, **52**, 3969–3976, https://doi.org/10.1175/1520-0469(1995)052<3969:SOTCTS>2.0.CO;2.

European Centre for Medium-Range Weather Forecasts, 2011: ECMWF's operational model analysis, starting in 2011. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, accessed 6 June 2022, https://doi.org/10.5065/D6ZG6Q9F.

Fernández-González, S., M. L. Martín, E. García-Ortega, A. Merino, J. Lorenzana, J. L. Sánchez, F. Valero, and J. S. Rodrigo, 2018: Sensitivity analysis of the WRF Model: Wind-resource assessment for complex terrain. *J. Appl. Meteor. Climatol.*, **57**, 733–753, https://doi.org/10.1175/JAMC-D-17-0121.1.

Fernando, H. J. S., and Coauthors, 2019: The Perdigão: Peering into microscale details of mountain winds. *Bull. Amer. Meteor. Soc.*, **100**, 799–819, https://doi.org/10.1175/BAMS-D-17-0227.1.

Giani, P., F. Tagle, M. G. Genton, S. Castruccio, and P. Crippa, 2020: Closing the gap between wind energy targets and implementation for emerging countries. *Appl. Energy*, **269**, 115085, https://doi.org/10.1016/j.apenergy.2020.115085.

——, M. G. Genton, and P. Crippa, 2022: Modeling the convective boundary layer in the terra incognita: Evaluation of different strategies with real-case simulations. *Mon. Wea. Rev.*, **150**, 981–1001, https://doi.org/10.1175/MWR-D-21-0216.1.

Gustafson, W. I., and Coauthors, 2020: The Large-Eddy Simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) activity for continental shallow convection. *Bull. Amer. Meteor. Soc.*, **101**, E462–E479, https://doi.org/10.1175/BAMS-D-19-0065.1.

Haupt, S. E., and Coauthors, 2019: On bridging a modeling scale gap: Mesoscale to microscale coupling for wind energy. *Bull. Amer. Meteor. Soc.*, **100**, 2533–2550, https://doi.org/10.1175/BAMS-D-18-0033.1.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Honnert, R., 2016: Representation of the grey zone of turbulence in the atmospheric boundary layer. *Adv. Sci. Res.*, **13**, 63–67, https://doi.org/10.5194/asr-13-63-2016.

——, V. Masson, and F. Couvreux, 2011: A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale. *J. Atmos. Sci.*, **68**, 3112–3131, https://doi.org/10.1175/JAS-D-11-061.1.

Hunt, J. C. R., S. Leibovich, and K. J. Richards, 1988: Turbulent shear flows over low hills. *Quart. J. Roy. Meteor. Soc.*, **114**, 1435–1470, https://doi.org/10.1002/qj.49711448405.

Ito, J., H. Niino, M. Nakanishi, and C.-H. Moeng, 2015: An extension of the Mellor–Yamada model to the terra incognita zone for dry convective mixed layers in the free convection regime. *Bound.-Layer Meteor.*, **157**, 23–43, https://doi.org/10.1007/s10546-015-0045-5.

Jiménez-Esteve, B., M. Udina, M. R. Soler, N. Pepin, and J. R. Miró, 2018: Land use and topography influence in a complex terrain area: A high resolution mesoscale modelling study over the eastern Pyrenees using the WRF model. *Atmos. Res.*, **202**, 49–62, https://doi.org/10.1016/j.atmosres.2017.11.012.

Juliano, T. W., B. Kosović, P. A. Jiménez, M. Eghdami, S. E. Haupt, and A. Martilli, 2022: "Gray zone" simulations using a three-dimensional planetary boundary layer parameterization in the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **150**, 1585–1619, https://doi.org/10.1175/MWR-D-21-0164.1.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802, https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.

Kitagawa, Y. K. L., T. T. de Almeida Albuquerque, P. Kumar, E. G. S. Nascimento, and D. M. Moreira, 2022: Coastal-urban meteorology: A sensitivity study using the WRF-urban model. *Urban Climate*, **44**, 101185, https://doi.org/10.1016/j.uclim.2022.101185.

Klemp, J. B., W. C. Skamarock, and O. Fuhrer, 2003: Numerical consistency of metric terms in terrain-following coordinates. *Mon. Wea. Rev.*, **131**, 1229–1239, https://doi.org/10.1175/1520-0493(2003)131<1229:NCOMTI>2.0.CO;2.

Kochanski, A. K., E. R. Pardyjak, R. Stoll, A. Gowardhan, M. J. Brown, and W. J. Steenburgh, 2015: One-way coupling of the WRF–QUIC urban dispersion modeling system. *J. Appl. Meteor. Climatol.*, **54**, 2119–2139, https://doi.org/10.1175/JAMC-D-15-0020.1.

Kosović, B., 1997: Subgrid-scale modelling for the large-eddy simulation of high-Reynolds-number boundary layers. *J. Fluid Mech.*, **336**, 151–182, https://doi.org/10.1017/S0022112096004697.

Lilly, D. K., 1967: The representation of small-scale turbulence in numerical simulation experiments. *Proc. IBM Scientific Computing Symp.*, Yorktown Heights, NY, IBM, 195–210.

Lin, D., B. Khan, M. Katurji, L. Bird, R. Faria, and L. E. Revell, 2021: WRF4PALM v1.0: A mesoscale dynamical driver for the microscale PALM model system 6.0. *Geosci. Model Dev.*, **14**, 2503–2524, https://doi.org/10.5194/gmd-14-2503-2021.

Liu, Y., Y. Liu, D. Muñoz-Esparza, F. Hu, C. Yan, and S. Miao, 2020: Simulation of flow fields in complex terrain with WRF-LES: Sensitivity assessment of different PBL treatments. *J. Appl. Meteor. Climatol.*, **59**, 1481–1501, https://doi.org/10.1175/JAMC-D-19-0304.1.

Lundquist, K. A., F. K. Chow, and J. K. Lundquist, 2010: An immersed boundary method for the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **138**, 796–817, https://doi.org/10.1175/2009MWR2990.1.

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

Mirocha, J. D., J. K. Lundquist, and B. Kosović, 2010: Implementation of a nonlinear subfilter turbulence stress model for large-eddy simulation in the Advanced Research WRF Model. *Mon. Wea. Rev.*, **138**, 4212–4228, https://doi.org/10.1175/2010MWR3286.1.

——, and Coauthors, 2018: Large-eddy simulation sensitivities to variations of configuration and forcing parameters in canonical boundary-layer flows for wind energy applications. *Wind Energy Sci.*, **3**, 589–613, https://doi.org/10.5194/wes-3-589-2018.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682, https://doi.org/10.1029/97JD00237.

Moisseeva, N., and D. G. Steyn, 2014: Dynamical analysis of sea-breeze hodograph rotation in Sardinia. *Atmos. Chem. Phys.*, **14**, 13471–13481, https://doi.org/10.5194/acp-14-13471-2014.

NCEP, 2015: NCEP GDAS/FNL 0.25 degree global tropospheric analyses and forecast grids. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, accessed 6 June 2022, https://doi.org/10.5065/D65Q4T4Z.

Noh, Y., W. G. Cheon, S. Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.*, **107**, 401–427, https://doi.org/10.1023/A:1022146015946.

Park, S. H., W. C. Skamarock, J. B. Klemp, L. D. Fowler, and M. G. Duda, 2013: Evaluation of global atmospheric solvers using extensions of the Jablonowski and Williamson baroclinic wave test case. *Mon. Wea. Rev.*, **141**, 3116–3129, https://doi.org/10.1175/MWR-D-12-00096.1.

Pineda, N., O. Jorba, J. Jorge, and J. M. Baldasano, 2004: Using NOAA AVHRR and SPOT VGT data to estimate surface parameters: Application to a mesoscale meteorological model. *Int. J. Remote Sens.*, **25**, 129–143, https://doi.org/10.1080/0143116031000115201.

Ronda, R. J., G. J. Steeneveld, B. G. Heusinkveld, J. J. Attema, and A. A. M. Holtslag, 2017: Urban finescale forecasting reveals weather conditions with unprecedented detail. *Bull. Amer. Meteor. Soc.*, **98**, 2675–2688, https://doi.org/10.1175/BAMS-D-16-0297.1.

Safaei Pirooz, A. A., and R. G. J. Flay, 2018: Comparison of speed-up over hills derived from wind-tunnel experiments, wind-loading standards, and numerical modelling. *Bound.-Layer Meteor.*, **168**, 213–246, https://doi.org/10.1007/s10546-018-0350-x.

Schemann, V., K. Ebell, B. Pospichal, R. Neggers, C. Moseley, and B. Stevens, 2020: Linking large-eddy simulations to local cloud observations. *J. Adv. Model. Earth Syst.*, **12**, e2020MS002209, https://doi.org/10.1029/2020MS002209.

Schicker, I., D. Arnold Arias, and P. Seibert, 2016: Influences of updated land-use datasets on WRF simulations for two Austrian regions. *Meteor. Atmos. Phys.*, **128**, 279–301, https://doi.org/10.1007/s00703-015-0416-y.

Shin, H. H., and S.-Y. Hong, 2013: Analysis of resolved and parameterized vertical transports in convective boundary layers at gray-zone resolutions. *J. Atmos. Sci.*, **70**, 3248–3261, https://doi.org/10.1175/JAS-D-12-0290.1.

——, and ——, 2015: Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions. *Mon. Wea. Rev.*, **143**, 250–271, https://doi.org/10.1175/MWR-D-14-00116.1.

——, and Coauthors, 2021: Large-scale forcing impact on the development of shallow convective clouds revealed from LASSO large-eddy simulations. *J. Geophys. Res. Atmos.*, **126**, e2021JD035208, https://doi.org/10.1029/2021JD035208.

Skamarock, W. C., and Coauthors, 2021: A description of the Advanced Research WRF Model version 4.3. NCAR Tech. Rep. NCAR/TN-556+STR, 165 pp.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Wea. Rev.*, **91**, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Stevens, B., and Coauthors, 2020: The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *J. Meteor. Soc. Japan*, **98**, 395–435, https://doi.org/10.2151/jmsj.2020-021.

Talbot, C., E. Bou-Zeid, and J. Smith, 2012: Nested mesoscale large-eddy simulations with WRF: Performance in real test cases. *J. Hydrometeor.*, **13**, 1421–1441, https://doi.org/10.1175/JHM-D-11-048.1.

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

Wagner, J., T. Gerz, N. Wildmann, and K. Gramitzky, 2019: Long-term simulation of the boundary layer flow over the double-ridge site during the Perdigão 2017 field campaign. *Atmos. Chem. Phys.*, **19**, 1129–1146, https://doi.org/10.5194/acp-19-1129-2019.

Wang, B. H., D. B. Wang, Z. A. Ali, B. Ting Ting, and H. Wang, 2019: An overview of various kinds of wind effects on unmanned aerial vehicle. *Meas. Control*, **52**, 731–739, https://doi.org/10.1177/0020294019847688.

Wang, Q., K. Luo, R. Yuan, S. Wang, J. Fan, and K. Cen, 2020: A multiscale numerical framework coupled with control strategies for simulating a wind farm in complex terrain. *Energy*, **203**, 117913, https://doi.org/10.1016/j.energy.2020.117913.

Wenz, F., J. Langner, T. Lutz, and E. Krämer, 2022: Impact of the wind field at the complex-terrain site Perdigão on the surface pressure fluctuations of a wind turbine. *Wind Energy Sci.*, **7**, 1321–1340, https://doi.org/10.5194/wes-7-1321-2022.

Werner, M., 2001: Shuttle Radar Topography Mission (SRTM) mission overview. *Frequenz*, **55**, 75–79, https://doi.org/10.1515/FREQ.2001.55.3-4.75.

Wise, A. S., J. M. T. Neher, R. S. Arthur, J. D. Mirocha, J. K. Lundquist, and F. K. Chow, 2022: Meso- to microscale modeling of atmospheric stability effects on wind turbine wake behavior in complex terrain. *Wind Energy Sci.*, **7**, 367–386, https://doi.org/10.5194/wes-7-367-2022.

Wyngaard, J. C., 2004: Toward numerical modeling in the "terra incognita." *J. Atmos. Sci.*, **61**, 1816–1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2.

Xu, H., H. Wang, and Y. Duan, 2021: An investigation of the impact of different turbulence schemes on the tropical cyclone boundary layer at turbulent gray-zone resolution. *J. Geophys. Res. Atmos.*, **126**, e2021JD035327, https://doi.org/10.1029/2021JD035327.

Zhang, X., J.-W. Bao, B. Chen, and E. D. Grell, 2018: A three-dimensional scale-adaptive turbulent kinetic energy scheme in the WRF-ARW model. *Mon. Wea. Rev.*, **146**, 2023–2045, https://doi.org/10.1175/MWR-D-17-0356.1.

Zhou, B., J. S. Simon, and F. K. Chow, 2014: The convective boundary layer in the terra incognita. *J. Atmos. Sci.*, **71**, 2545–2563, https://doi.org/10.1175/JAS-D-13-0356.1.

——, M. Xue, and K. Zhu, 2018: A grid-refinement-based approach for modeling the convective boundary layer in the gray zone: Algorithm implementation and testing. *J. Atmos. Sci.*, **75**, 1143–1161, https://doi.org/10.1175/JAS-D-17-0346.1.