Modeling the Convective Boundary Layer in the Terra Incognita: Evaluation of Different Strategies with Real-Case Simulations

PAOLO GIANI,[a] MARC G. GENTON,[b] AND PAOLA CRIPPA[a]

[a] *Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, Indiana*
[b] *Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

(Manuscript received 13 August 2021, in final form 21 January 2022)

ABSTRACT: Modeling atmospheric turbulence in the convective boundary layer is challenging at kilometer and subkilometer resolutions, as the horizontal grid spacing approaches the size of the most energetic turbulent eddies. In this range of resolutions, termed *terra incognita* or *gray zone*, partially resolved convective structures are grid dependent and neither traditional 1D mesoscale parameterizations nor 3D large-eddy simulations closures are theoretically appropriate. Leveraging on a new set of one-way nested, full-physics multiscale numerical experiments, we quantify the magnitude of the errors introduced at gray zone resolutions in a real-case application and we provide new perspectives on recently proposed modeling approaches. The new set of experiments is forced by real-time-varying boundary conditions, spans a wide range of scales, and includes traditional 1D schemes, 3D closures, scale-aware parameterizations, and strategies to suppress resolved convection at gray zone resolutions. The study area is Riyadh (Saudi Arabia), where deep CBLs develop owing to strong convective conditions. Detailed analyses of our experiments, including validation with radiosonde data, calculations of spectral features, and partitioning of turbulent fluxes between resolved and subgrid scales, show that (i) grid-dependent convective structures entail minor impacts on the first-order characteristics of the fully developed boundary layer due to some degree of *implicit scale awareness* of 1D parameterizations and (ii) 3D closures and scale-aware schemes outperform traditional 1D schemes especially in the surface layer, among other findings. The new suite of experiments provides a benchmark of real simulations that can be extended to assess how new turbulence closures perform at gray zone resolutions.

SIGNIFICANCE STATEMENT: As recent advances in high-performance computing are leading to a new era in numerical simulations, regional atmospheric models can now increase their resolution to the widely unexplored kilometer and subkilometer range. While increasing the resolution of atmospheric models is desirable to (i) have more realistic weather and air quality predictions and (ii) better represent boundary conditions for microscale models, kilometer and subkilometer grid spacings pose some theoretical challenges that need to be addressed by the atmospheric modeling community. In this work we run a set of numerical experiments for a real case study that aim to offer new perspectives on recently developed modeling strategies and identify the most promising directions that should be investigated by follow-up studies.

KEYWORDS: Atmosphere; Convection; Turbulence; Eddies; Boundary conditions; Coupled models; Large eddy simulations; Mesoscale models; Model evaluation/performance; Parameterization; Subgrid-scale processes

1. Introduction

Turbulent motions regulate vertical transport of momentum, heat, moisture, and pollutants in the atmospheric boundary layer (Stull 1988). A correct representation of turbulent diffusion in numerical models is key for a wide range of applications, from weather and air quality predictions (Cécé et al. 2016; Chu et al. 2014) to wind power assessments (Mirocha et al. 2018). Depending on the ratio between the model horizontal grid spacing (Δx) and the size of energy-containing turbulent structures (ℓ), two classes of modeling techniques have been explored over the past few decades. If the mesh size is well above the dominant length scale of turbulence (i.e., $\ell/\Delta x \ll 1$), as in operational numerical weather prediction model (NWP) simulations, all turbulent motions occur at the subgrid scale (SGS) and parameterizations are required to model their impact on the mean flow. In the mesoscale modeling community, these parameterizations are usually referred to as 1D planetary boundary layer (PBL) schemes. The second modeling approach, when the model grid spacing is significantly lower than the size of the largest turbulent structures (i.e., $\ell/\Delta x \gg 1$, where Δx is within the inertial range), is known as large-eddy simulations (LES). In LES modeling, the energetic turbulent motions are directly resolved on the numerical grid (Moeng 1984) owing to the relatively small Δx. A subgrid model is required to represent the interaction between the resolved and SGS turbulence (Lilly 1967; Leonard 1975).

As computational power has substantially increased in the last decade, NWP models can now reach the unexplored range of resolutions termed *terra incognita* (gray zone) by Wyngaard (2004), where the horizontal grid spacing is in the

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/MWR-D-21-0216.s1.

Corresponding author: Paolo Giani, pgiani@nd.edu

DOI: 10.1175/MWR-D-21-0216.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

order of the largest turbulent structures ($\ell/\Delta x \sim 1$). On the one hand, high-resolution NWP simulations are desirable to capture local characteristics of the flow (Sharma et al. 2017; Fernando et al. 2019; Wagner et al. 2019), to predict exposure to air pollutants at neighborhood scales (Fernando et al. 2010; Joe et al. 2014) and to provide boundary conditions to LES codes (Conry et al. 2015; Rai et al. 2019), among other applications. On the other hand, current turbulence parameterizations are not theoretically sound at gray zone resolutions, as large turbulent structures are partially resolved on the grid, breaking the assumptions of both PBL (all turbulent motions occur at SGS scales) and LES SGS schemes (Δx is in the inertial range). In two pivotal studies, Honnert et al. (2011) and Zhou et al. (2014) laid out the foundations of the gray zone problem and showed that NWP models with resolutions in the order of 1 km fall within the gray zone range during daytime convection, as large turbulent eddies driven by convection span the entire convective boundary layer (CBL). In modeling the CBL at gray zone resolutions, Zhou et al. (2014) revealed that the size of resolved convective rolls or cells depends on the grid resolution, rather than the natural state of the flow. At about the same time, Ching et al. (2014) independently discussed the same *grid-dependency* problem from a relatively different standpoint and offered perspectives on how to remove under-resolved convective patterns.

In very recent years, numerous studies have proposed novel techniques to overcome the difficulties laid out by Zhou et al. (2014) and Ching et al. (2014) in characterizing the CBL at gray zone resolutions. Two main avenues can be broadly identified from this growing body of literature. The first approach is based on the idea that poorly resolved convective structures should be suppressed at gray zone resolutions, given the grid-dependency problem (Poll et al. 2017; Muñoz-Esparza et al. 2017). The alternative approach is to retain the grid-dependent convective structures while introducing scale-aware schemes that reduce subgrid fluxes in PBL schemes, to account for mixing done at resolved scales. Honnert et al. (2011) opened the way for scale-aware schemes and modifications in both local and nonlocal PBL schemes were introduced shortly after (Beare 2014; Boutle et al. 2014; Ito et al. 2015; Shin and Hong 2015).

Although the gray zone community has proposed different modeling strategies that can be potentially suitable in the *terra incognita* (e.g., Ito et al. 2015; Kurowski and Teixeira 2018; Honnert et al. 2021), the verification of these novel techniques for real CBLs is still very limited. While a few studies have looked at the performance of non-idealized, full-physics model simulations at gray zone resolutions with conventional PBL schemes (Doubrawa and Muñoz-Esparza 2020; Liu et al. 2020; Rai et al. 2017, 2019; Shin and Dudhia 2016; Xu et al. 2018), the literature lacks a comprehensive comparison and verification of the modeling strategies outlined above, especially in real cases. Understanding strengths and weaknesses of different approaches would benefit both mesoscale modelers seeking guidance on how to perform simulations at gray zone resolutions as well as model developers to broaden the impact of their studies. Building upon previous efforts, in this work we aim to bridge this gap between the recent progress in the gray zone community and the operational needs in the NWP community, which has not comprehensively addressed this issue so far (Doubrawa and Muñoz-Esparza 2020). We leverage on a new set of nine real, full-physics, nested simulations with the Weather Research and Forecasting (WRF) Model to address the following objectives:

1) Identify and quantify potential errors that may arise at gray zone resolutions for first-order statistics of the flow and for turbulent quantities in real CBLs,
2) Evaluate which modeling strategies are more effective in mitigating the problem, among the different avenues indicated by recent idealized simulations,
3) Provide practical recommendations on how to run real high-resolution simulations that potentially fall within gray zone resolutions.

Our nine simulations span the whole spectrum of model resolutions, from traditional mesoscale (Δx = 12 km) to well-resolved LES (Δx = 50 m), passing through the gray zone (Δx = 1350 and 450 m), and they include both traditional and recently developed techniques to model turbulence at gray zone resolutions. All the simulations are centered in Riyadh, Saudi Arabia, which is an ideal area to study the gray zone problem as deep CBLs (up to 4–5 km) form over the desert owing to particularly strong convection (Gamo 1996). The region of Riyadh is also of strategic interest for building wind energy capacity, as shown in previous mesoscale work (Giani et al. 2020). Although drawing specific conclusions for wind energy exceeds the scope of this work, high-resolution simulations over Riyadh are potentially attractive for characterizing the local flow patterns which are key for wind farms planning and managing (Haupt et al. 2019).

The remainder of the article is organized as follows. Section 2 provides a brief description of WRF turbulence closure schemes and describes the recent progress for gray zone resolutions. The numerical simulations and methods are presented in section 3. Based on the results detailed in section 4 and the evidence produced by previous work, section 5 summarizes general recommendations on how to set up full-physics, real simulations at gray zone resolutions.

2. Review of subgrid turbulence parameterizations in WRF

In this section, we review the main schemes commonly used to represent subgrid turbulent mixing in WRF, which is the focus of this work. We do not intend to provide a comprehensive description of all the implemented schemes (which can be found elsewhere, e.g., Skamarock et al. 2008; Cohen et al. 2015 and references therein), but rather to highlight the main features that are needed to explain the rationale behind our experiments (section 3b).

If a PBL scheme is selected in WRF (e.g., for typical mesoscale applications), vertical turbulent diffusion is calculated within PBL routines whereas horizontal diffusion is handled by a separate module based on the Smagorinsky first-order closure (Smagorinsky 1963). Horizontal eddy viscosities K_h

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

are computed as a function of horizontal gradients and the grid size ($\Delta x \Delta y$):

$$K_h = c_s \Delta x \Delta y \sqrt{\frac{1}{4}(D_{11} - D_{22})^2 + D_{12}^2}, \quad (1)$$

where c_s is the so-called Smagorinsky constant (set to 0.25 in WRF) and D_{ij} ($i, j = 1, 2, 3$) is the deformation tensor (Skamarock et al. 2008). In recent literature, increasing the Smagorinsky constant (and therefore the dissipation of thermal instabilities via horizontal diffusion) has been identified as a way of suppressing resolved convection at gray zone resolutions. Muñoz-Esparza et al. (2017) performed some sensitivity tests with artificially amplified Smagorinsky constant and showed that $c_s = 2.0$ removed significant contributions of under-resolved convection from the grid. Although the sensitivity tests were successful in removing resolved convection, c_s was artificially increased to unphysical values.

For vertical turbulent transport, several PBL schemes are implemented in WRF by default and they can be classified in local and nonlocal approaches. In this work, we use the Mellor–Yamada–Janjić (MYJ; Mellor and Yamada 1982; Janjić 2001) scheme and the revised version by Nakanishi and Niino (2004, 2009) (MYNN2.5) as local schemes, and the Yonsei University Scheme (YSU; Hong et al. 2006) and its revised version by Shin and Hong (2015) (SH) as nonlocal ones.

For a generic quantity C, a local PBL scheme calculates subgrid turbulent fluxes ($\overline{w'C'}$) based on local gradients only:

$$\overline{w'C'} = -K_c \frac{\partial \overline{C}^{\Delta}}{\partial z}, \quad (2)$$

where the overbar with Δ represents the resolved part on a grid with horizontal spacing Δ (i.e., $C = \overline{C}^{\Delta} + C'$), following the notation of Shin and Dudhia (2016), z is the vertical coordinate and K_c is the eddy diffusivity. Eddy diffusivities are typically a function of local values of turbulent kinetic energy (TKE), a length scale (sometimes referred to as mixing length, owing to Prandtl's theory) and a stability function. Differences between local schemes are related to the formulation of these quantities. Scale-awareness suitable for gray zone applications can be introduced in local schemes by modifying the mixing length formulation, as proposed by Ito et al. (2015). This modification is hard-coded in the MYNN subroutines, and activating/deactivating this option requires recompiling the WRF code.

Nonlocal approaches (e.g., YSU) explicitly includes a countergradient term γ_c to account for the contribution of large scale eddies and the entrainment flux $\overline{w'C'_{z_i}}(z/z_i)^3$ from the top of the boundary layer z_i (Hong and Pan 1996):

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

Shin and Hong (2015) introduced scale-awareness to this parameterization to address the gray zone issue, by scaling the magnitude of the local flux of Eq. (3) with a grid-size dependency function and revising the nonlocal heat fluxes:

$$\overline{w'C'} = -K_c\left(\frac{\partial \overline{C}^{\Delta}}{\partial z}\right) P_L(z, \Delta x_*) + \overline{w'C'_{\mathrm{NL}}} P_{\mathrm{NL}}\left(u_* w_*^{-1}, \Delta x_*\right), \quad (4)$$

where $\overline{w'C_{\mathrm{NL}}}$ is the nonlocal flux, which differs from the YSU one, Δx_* is the normalized grid spacing ($\Delta x_* = z_i^{-1}\Delta x$), $u_* w_*^{-1}$ is the friction velocity to Deardoff velocity ratio, P_L and P_{NL} are empirical functions bounded within 0 and 1 that reduce the magnitude of the turbulent flux depending on the grid spacing (hence the term scale-awareness).

If a PBL scheme is not selected (i.e., bl_pbl_physics = 0 in WRF), 3D LES turbulence closures can be used to calculate both horizontal and vertical diffusion. The two 3D closures used in this work are the 1.5 TKE anisotropic model of Lilly (1967) and the nonlinear backscatter and anisotropy formulation of Kosović (1997), based on strain rate only, implemented in WRF by Mirocha et al. (2010). The 1.5 TKE model calculates eddy diffusivities based on local TKE values and scale-aware length scales, whereas the approach of Kosović (1997) takes into account energy backscatter to the resolved field through a nonlinear approach.

3. Methods

a. Overall model configuration

We run a set of eight multiscale and nested numerical experiments and one reference simulation with the Weather Research and Forecasting (WRF) Model version V3.9.1.1, which solves the nonhydrostatic compressible Euler equations in terrain-following coordinates (Skamarock et al. 2008). In this section, we describe the common model configuration across all simulations whereas the differences between the nine simulations are explained in section 3b.

All eight numerical experiments share the same set of four one-way nested domains (d01–d04, Fig. 1 and Table 1) centered at Riyadh Airport, Saudi Arabia (24.93°N, 46.71°E), where radiosonde data are available. The four domains have a parent-to-child horizontal resolution ratio of 3:1, with the outermost domain horizontally discretized at a typical mesoscale resolution (Δx = 12 150 m) on a Lambert Conformal grid, covering most of the Arabian Peninsula. The reference simulation includes two additional domains (d05 and d06, with resolutions Δx = 150 and 50 m) with grid spacing that lies within the range of LES modeling (Zhou et al. 2014).

We use high-resolution topography data from the Shuttle Radar Topography Mission (SRTM) at 3 arc-sec resolution (~90 m) to produce elevation data over d04 and d05, 1 arc-sec (~30 m) for d06, and the default U.S. Geological Survey (USGS) dataset at 30 arc-sec resolution (~900 m) for d01, d02, and d03. The SRTM datasets are retrieved from the digital elevation NASA database and are converted into the WRF preprocessing system readable binary format prior to their use (Dimitrova et al. 2016). Mesoscale and gray zone domains (d01–d04) consist of the same 49 vertical levels, with increasing vertical spacing (Δz) in the boundary layer and approximately constant above the boundary layer, up to 50 hPa. Given the vertical nesting capabilities recently implemented in WRF (Daniels et al. 2016), we set a larger number of vertical

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

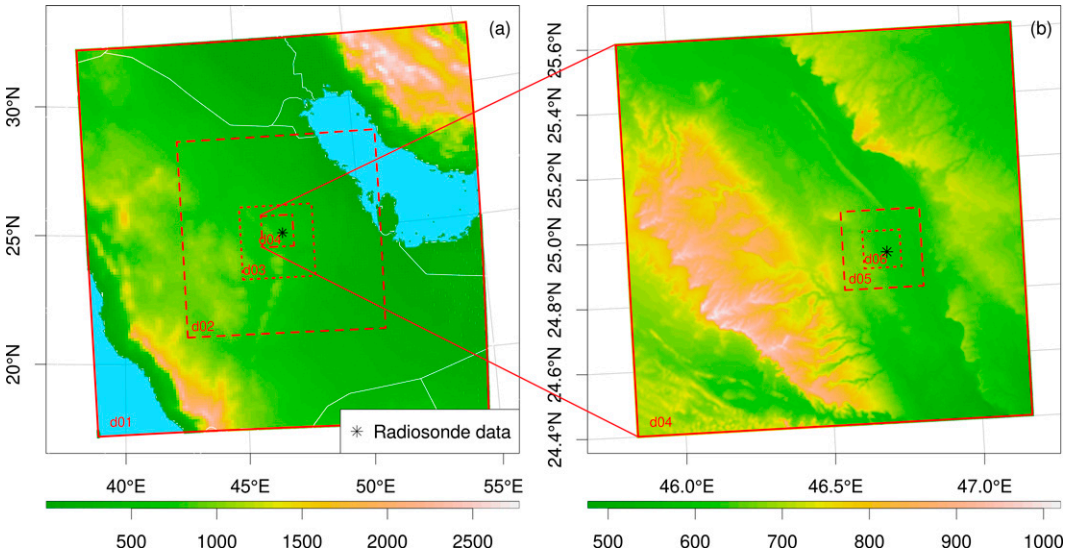

FIG. 1. (a) Mesoscale and gray zone domains (d01–d04) and (b) closeup look to gray zone and microscale domains (d04–d06). Shaded colors represent elevation in meters above ground level, and the star indicates the location of radiosonde data.

levels in d05 and d06 (76 and 106, respectively), as these domains are solved in LES mode where the aspect ratio ($\Delta x/\Delta z$) should be closer to unity (2–4 according to Mirocha et al. 2010). In the boundary layer, we use an approximately constant Δz = 64 and 32 m for d05 and d06, respectively, as done in recent LES works (e.g., Sullivan and Patton 2011). The number of vertical levels, the vertical spacing between them and other aspects of the domains setup are included in the namelist shared through Github (see data availability) and have been object of a thorough sensitivity analysis, which is briefly discussed in the recommendations in section 5.

Simulations are initialized at 0000 UTC 22 July 2016 with the high-resolution operational data from the European Centre for Medium-Range Weather Forecast (HRES-ECMWF). The Euler equations of motion are integrated for 10 days until 0000 UTC 1 August 2016, with a time step Δt = 72 s for the outermost domain and a parent-to-child time step ratio of 3:1. The first day of simulation is discarded from the analysis as model spinup. All the fields relevant to the analysis are saved with hourly frequency. Lateral boundary conditions for the outermost domain are updated every 6 h and are provided by HRES-ECMWF. The simulated episode consists of nine dry cycles of strong convection, with deep boundary layers, high temperatures and large heat fluxes (Fig. S1 in the online supplemental material). All experiments are full-physics, meaning that are driven by heterogeneous heat and momentum fluxes as well as by regional/synoptic conditions through realistic boundary conditions from HRES-ECMWF. The unified Noah land surface model (Ek et al. 2003), in combination with the surface layer schemes shown in Table 2, provide time- and space-varying surface heat, moisture and momentum fluxes. Short and longwave radiation are computed with the Dudhia scheme (Dudhia 1989) and the Rapid Radiative Transfer Model (Mlawer et al. 1997), respectively.

b. *Individual experiments setup*

The set of nested simulations includes four pairs (P1–P4) of numerical experiments (RN01 through RN08), where each pair is designed to assess a different route to address the gray zone issue, and one well-resolved run that is used as a reference (REF). The eight experiments differ in the way turbulence is modeled in the gray zone domains (d03 and d04, Δx = 1350 and 450 m, respectively), as explained below and illustrated in Table 2.

The goal of P1 simulations is to assess the impact of removing resolved convection at gray zone resolutions, as proposed by Ching et al. (2014) and Poll et al. (2017). P1 simulations are traditional mesoscale simulations with MYJ closure for

TABLE 1. Geometric characteristics of the six domains considered in this work and associated model time step.

Domain label	Grid points	Resolution (m)	Topography dataset	Total area covered (km^2)	Time step (s)
d01	140 × 140 × 49	12 150	USGS 30 arc sec	1701 × 1701	72
d02	211 × 211 × 49	4050	USGS 30 arc sec	854.5 × 854.5	24
d03	232 × 232 × 49	1350	USGS 30 arc sec	313.2 × 313.2	8
d04	301 × 301 × 49	450	SRTM 3 arc sec	135.4 × 135.4	8/3
d05	181 × 181 × 76	150	SRTM 3 arc sec	27.15 × 27.15	8/9
d06	256 × 256 × 106	50	SRTM 1 arc sec	12.80 × 12.80	8/27

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

TABLE 2. Main turbulence-related features of the eight WRF experiments and reference simulations: c_s = Smagorinsky coefficient (d01/d02/d03/d04 or only one value for all domains, see section 3b); MYJ = Mellor–Yamada–Janjić scheme; MYNN2.5 = Mellor–Yamada–Nakanishi–Niino scheme level 2.5; SA = scale aware; NSA = not scale-aware; TKE = Turbulent kinetic energy; and NBA = nonlinear backscatter and anisotropy. REF characteristics are related to d05 and d06. The turbulent diffusion scheme refers to the one used in the two innermost domains (d03 and d04 for RN01-RN08, d05 and d06 for REF.

Run label	Pair ID	c_s	Surface layer	Horizontal turbulent diffusion	Vertical turbulent diffusion
RN01	P1	0.25	Janjić Eta	2D Smagorinsky	MYJ
RN02	P1	0.25/0.5/1.0/2.0	Janjić Eta	2D Smagorinsky	MYJ
RN03	P2	0.25	Jimenez	2D Smagorinsky	YSU
RN04	P2	0.25	Jimenez	2D Smagorinsky	Shin–Hong
RN05	P3	0.25	MYNN	2D Smagorinsky	MYNN2.5 (SA)
RN06	P3	0.25	MYNN	2D Smagorinsky	MYNN2.5 (NSA)
RN07	P4	0.25	MYNN	3D based on TKE	3D based on TKE
RN08	P4	0.25	MYNN	3D NBA (strain rate)	3D NBA (strain rate)
REF	—	0.25	MYNN	3D NBA (strain rate)	3D NBA (strain rate)

vertical transport and Smagorinsky for horizontal diffusion. The only difference between RN01 and RN02 is the Smagorinsky coefficient, which is set to the default 0.25 for all domains in RN01 and to 0.25, 0.50, 1.0, and 2.0 for d01, d02, d03, and d04 in RN02, respectively. The specific values of c_s used in this work allow for a significant damping of resolved convection (Fig. S2 and section 4a), similarly to Muñoz-Esparza et al. (2017). The second pair of experiments (P2) addresses the effect of introducing scale-awareness in nonlocal schemes, by comparing a traditional YSU nonlocal closure (RN03) and its scale-aware version (RN04). In a similar fashion, P3 contrasts scale-aware and traditional schemes for local closures (MYNN2.5), as the only difference between RN05 and RN06 is the mixing length revision of Ito et al. (2015) implemented in the MYNN2.5 routines. In this work we use the local-only version of MYNN (Nakanishi and Niino 2004, 2009) to test the impact of scale-awareness on a local scheme, although it should be noted that the MYNN scheme can be run with a nonlocal mass-flux scheme (MYNN-EDMF; Olson et al. 2019) from the release of WRF v3.9. The last pair of simulations (P4) instead uses two different 3D closures in d03 and d04, the traditional 1.5 TKE model in RN07 (Lilly 1967; Skamarock et al. 2008) and the NBA approach of Kosović (1997) in RN08, which was shown to have a better agreement with reference data at coarser resolutions than the 1.5 TKE approach (Mirocha et al. 2010). For P4 simulations, domains d01 to d02 are equivalent to P3 simulations.

The reference simulation (REF) includes two additional domains (d05 and d06) modeled with the 3D NBA closure in LES mode, and the setup for the coarser domains (d01 to d04) is identical to RN05. This specific configuration is based on literature considerations (Mirocha et al. 2010) and the results of P1, P2, P3, and P4 simulations (section 4). A full sensitivity analysis of the reference dataset (i.e., the performance of WRF-LES at these resolutions), including both the effect of lateral boundary conditions and the SGS models (e.g., Liu et al. 2020), would be worth investigating but exceeds the main purpose of this work, which is related to the gray zone domains. However, we compare the results of REF with observed data to ensure that the reference simulation is capturing the mean vertical profiles. For turbulence statistics, previous studies showed that the variability related to different PBL treatments (Liu et al. 2020) and the SGS model (Talbot et al. 2012) in the parent domains is relatively small, as long as there is enough fetch in the innermost domain to break down the under-resolved turbulent structures from the parent domain (Mazzaro et al. 2017). Given the large heat fluxes that drive our case study (Fig. S1), REF achieves its own turbulent motions even close to the boundaries for both d05 and d06 (Fig. S3).

c. *Model evaluation overview*

The model evaluation is based on a direct comparison of the numerical simulations at different resolutions with two reference datasets, i.e., REF and measured radiosonde data at Riyadh Airport. We present results for the mesoscale domain (d01, Δx = 12 km), the gray zone domain (d03, Δx = 1350 m) as well as two microscale domains (d05 and d06, Δx = 150 and 50 m, respectively). For the CBL under investigation, with an average boundary layer height $z_i \approx 3600$ m, d03 can be fully considered in the gray zone during convective instances according to the definitions of Efstathiou et al. (2018); Honnert et al. (2020), as $\Delta x/z_i \approx 0.38$. The analysis of the output from d04 ($\Delta x/z_i \approx 0.13$), which falls into the near gray zone according to the definition of Efstathiou et al. (2018), is mostly deferred to the supplementary material (Fig. S4–S11), although the key differences between d03 and d04 are reported in the results section of the manuscript. Radiosonde data are acquired from the University of Wyoming repository for every day during the period of interest at 1200 UTC [1500 local solar time (LST)], and they are used to assess the model predictions of the mean structure of the CBL. Analyzing daily replicates of multiple convective instances is key to smooth out the natural variability of turbulent processes. For a quantitative assessment of mean quantities, we calculate the root-mean-square error (RMSE) between average radiosonde profiles and the modeled profiles averaged at the same instants in time. Further, spectral analysis is used to characterize horizontal coherent structures at different resolutions, as described in the next subsection (section 3d). Finally, turbulent fluxes are calculated to understand how they partition between the SGS

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

and resolved scales between different resolutions and turbulence closures (details in section 3e).

d. Spectral analysis

The goal of the spectral analysis is to characterize the horizontal convective structures and quantitatively assess their size and type through their spectral characteristics. To allow for a proper comparison between the well-resolved LES and the gray zone solutions, we upscale the REF solution for d05 and d06 onto the d03 grid (example in Fig. S12). The upscaled LES solution, which is directly comparable to the coarser gray zone one, is obtained with a top-hat filter, as done in Honnert (2019). As we are interested in the CBL, we identify regimes of free convection by calculating the following dimensionless parameter ($-z_i/L$) with hourly frequency, similarly to Doubrawa and Muñoz-Esparza (2020):

$$-\frac{z_i}{L} = \frac{k_{\mathrm{VK}} z_i g \left(\overline{w'\theta'_v}\right)_s}{u_*^3 \overline{\theta}_v}, \tag{5}$$

where L is the Obukhov length (m), k_{VK} is the von Kármán constant (0.4), g is the gravitational acceleration (9.81 m s^{-2}), $\left(\overline{w'\theta'_v}\right)_s$ is the virtual potential temperature flux at the ground (m s^{-1} K), u_* is the friction velocity (m s^{-1}), and $\overline{\theta}_v$ is the virtual temperature at the first model level (K). According to previous literature (Deardorff 1972; Honnert et al. 2011), free convective conditions are defined as instances when $-z_i/L > 50$. During such conditions, convection tends to organize in large cells with narrow updrafts on the boundary and wide downdrafts in the middle (Doubrawa and Muñoz-Esparza 2020). Given the heterogeneity in boundary layer heights and heat fluxes in real simulations, we use horizontal averages within the area covered by the d06 domain to characterize the $-z_i/L$ parameter. At each instance in time characterized by free convection and for each model level below the boundary layer height, we measure the size of these convective cells by computing the spectral peak of vertical velocity. The methodology follows de Roode et al. (2004) and is briefly summarized here from a numerical standpoint. We first linearly detrend horizontal slices of vertical velocity $w_D(x, y)$ and get the discrete Fourier transform of the detrended field $\hat{w}_D(k_x,k_y)$ to calculate the 2D matrix of spectral energy $E(k_x, k_y)$:

$$E(k_x,k_y) = \left|\hat{w}_D(k_x,k_y)\right|^2. \tag{6}$$

After transforming the wavenumbers k_x and k_y into cylindrical coordinates, defining $k = (k_x + k_y)^{0.5}$ and $\vartheta = \tan^{-1}(k_y/k_x)(0 \le \vartheta \le \pi)$, the 1D azimuthally averaged density spectrum, which depends on k only, is obtained by integrating out the dependence on ϑ:

$$S(k) = \frac{1}{\Delta k}\int_0^{\pi} E(k,\vartheta)d\vartheta, \tag{7}$$

where Δk is the wavenumber resolution to normalize the energy spectrum into a spectral density. The spectral peak [i.e., the maximum of $S(k)$], that occurs at wavenumber k_c, indicates the most energetic motions and therefore the length scale of the dominant eddies ($\Lambda_c = 1/k_c$). Due to potential noise around the spectral peak, we compute k_c as the wavenumber up to which 2/3 of the total variance (σ_w^2) is contained, as in de Roode et al. (2004) and Zhou et al. (2014):

$$\sigma_w^2 = \int_0^{\infty} S(k)dk, \tag{8}$$

$$\int_{k_c}^{k_N} S(k)dk = \int_{k_c}^{k_N} kS(k)\, d(\ln k) = \frac{2}{3}\,\sigma_w^2, \tag{9}$$

where k_N is the largest resolvable (Nyquist) wavenumber. Figure S13 shows an example of the overall spectral analysis procedure.

Besides their dominant size Λ_c, the type of coherent structures (e.g., rolls versus cells) is also critical in driving the boundary layer dynamics (Williams and Hacker 1992; Poll et al. 2017). In previous literature, the form of horizontal structures generated at gray zone resolutions is often analyzed through visual inspection (e.g., Haupt et al. 2019; Muñoz-Esparza et al. 2017). In this work we propose to characterize their type more rigorously, by calculating the 1D radially averaged spectrum and defining a measure of anisotropy. The 1D radially averaged spectrum $R(\vartheta)$ is obtained from the 2D spectral energy matrix by integrating out the dependence on k in cylindrical coordinates:

$$R(\vartheta) = \int_0^{k_N} E(k,\vartheta)dk. \tag{10}$$

Numerically, we subdivide the full circle into $N = 12$ sectors (30° each), and we calculate the amount of energy in each sector i [$R_i(\vartheta)$]. An appropriate measure of the degree of anisotropy α is the following:

$$\alpha = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[R_i(\vartheta) - \overline{R(\vartheta)}\right]^2}}{\overline{R(\vartheta)}}, \tag{11}$$

which is the coefficient of variation of the 1D radially averaged spectrum. The value of α simply measures the magnitude of $R(\vartheta)$ fluctuations (i.e., its standard deviation) compared to the mean of $R(\vartheta)$. For a perfectly isotropic spectrum, $\alpha = 0$ as $R(\vartheta)$ is a horizontal straight line. The larger the value of α, the more anisotropic the 1D radially averaged spectrum velocity field is.

The degree of anisotropy and the resolved energy values (α and σ_w^2, respectively) are used to compare gray zone vertical velocity fields during strong convection ($-z_i/L > 50$) against the well-resolved filtered LES output (REF). Given the large value of $-z_i/L$, REF produces mostly isotropic convective cells. We use values of α and σ_w^2 from REF fields to derive thresholds (α_t, σ_{wt}^2) to classify gray zone fields into under-resolved, transitional, and fully developed structures (S):

$$S = \begin{cases} \text{Fully developed}, & \text{if } \alpha \le \alpha_t \\ \text{Transitional}, & \text{if } \alpha > \alpha_t \text{ and } \sigma_w^2 \ge \sigma_{wt}^2 \\ \text{Under–resolved}, & \text{if } \alpha > \alpha_t \text{ and } \sigma_w^2 < \sigma_{wt}^2 \end{cases}. \tag{12}$$

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

A reasonable choice for the values of α_t and σ^2_{wt} is the maximum and minimum of the filtered REF fields (0.3140 and 0.3706 $m^2\ s^{-2}$, respectively). The rationale for the proposed classification is the following: (i) if structures are highly anisotropic (i.e., α larger than the maximum anisotropy in REF α_t) and have very low energy (lower than the minimum energy in REF σ^2_{wt}) they will not resemble the REF fields (*under-resolved*) and (ii) if they are defined as anisotropic but comparable in energy they still entail substantial mixing and therefore they are classified as *transitional*, as they exhibit an intermediate behavior. Although the choice of the thresholds is somewhat arbitrary, the automatic classification based on Eq. (12) agrees very well with a visual inspection of the vertical velocity fields. An example can be found in section 4e.

e. Partitioning of turbulent fluxes

For each instance in time characterized by $-z_i/L > 50$ at hourly frequency, we calculate the partitioning of vertical fluxes of horizontal momentum $\left(\overline{w'u'^2} + \overline{w'v'^2}\right)^{0.5}$ and virtual temperature fluxes ($\overline{w'\theta'_v}$) between SGS and resolved scales. u and v represent the zonal and meridional velocity components, respectively. It should be noted that in our real simulations the total fluxes (i.e., SGS + resolved) are not guaranteed to be the same for all simulations, as opposed to the coarse-graining method where the total amount of energy (or fluxes) is assumed to be known (Honnert et al. 2011).

SGS fluxes are computed with Eq. (2) for local schemes and with Eqs. (3) and (4) for nonlocal schemes. Although WRF internally computes eddy viscosities and diffusivities, some of the PBL schemes do not output eddy viscosities by default, and therefore small modifications need to be introduced in module_pbl_driver.F, module_physics_init.F, and the specific PBL modules (e.g., module_bl_myjpbl.F) to fetch the right coefficients.

Resolved vertical fluxes for a generic quantity C can be computed by taking the covariance between vertical velocity fluctuations (w') and the fluctuations of the quantity itself (C'), at a specific instant in time. To calculate the turbulent fluctuations of the generic quantity C for gray zone domains, we remove the mean and the mesoscale contributions from each instantaneous spatial field, as illustrated in Fig. S14. Removing mesoscale features from the field is necessary as the simulations are forced by realistic mesoscale boundary conditions, and it is accomplished via spectral filtering, as recommended by recent experimental work (Donateo et al. 2017). Given a certain linearly detrended instantaneous spatial field, we calculate the azimuthally averaged spectrum with Eq. (7) and apply a high-pass filter to remove all the low mesoscale wavenumbers from the resolved field. Based on Larsén et al. (2016), the cutoff wavenumber (k_{HP}) used in the high-pass filter is the spectral gap position of each field (details in Fig. S14) as it separates mesoscale and turbulent motions (Larsén et al. 2016). In other words, we consider all the wavenumbers above k_{HP} to produce turbulent fluctuations, whereas all the wavenumbers below k_{HP} to be associated with mesoscale variability.

4. Results

a. The effect of suppressing resolved convection (P1 simulations)

Figure 2 shows the average vertical profiles of potential temperature and wind speed, calculated at 1200 UTC (1500 LST) at the radiosonde site, according to the measurements and REF and P1 numerical solutions, both at mesoscale and gray zone resolutions. Observations show the typical structure of the atmospheric boundary layer during strong convection. The surface layer is superadiabatic and is characterized by a steep decrease in temperature and quasi-logarithmically increasing wind speed. Above the surface layer, a deep mixed layer is formed by vigorous and large turbulent eddies driven by buoyancy that efficiently homogenize the inner part of the boundary layer. The capping inversion on top of the mixed layer marks the boundary layer height and the beginning of the free atmosphere, where the effect of surface friction and surface heat fluxes becomes negligible.

A traditional 12-km mesoscale simulation performs reasonably well in calculating the mean structure of the boundary layer, despite missing some important features. The first noticeable discrepancy is the wind speed profile in the surface layer, with observations and REF showing an increasing profile and the model calculating an almost well-mixed profile. A more subtle, yet important aspect is that mesoscale MYJ produces a slightly superadiabatic layer in the mixed layer, which is in contrast to the more neutral layer in the observations and in REF. This result is somewhat expected, as local PBL schemes can only represent downgradient fluxes [Eq. (2)], whereas countergradient fluxes are typically observed in the upper part of the CBL (LeMone et al. 2013). The gray zone simulation with MYJ closure produces a slightly more neutral temperature profile that better resembles the well-resolved LES solution in the mixed layer. Resolved convection on the grid seems to effectively add some nonlocal mixing as heat and momentum are transported upward by turbulent fluctuations, thereby achieving a more neutral vertical temperature profile even with a local scheme. The net effect of increasing horizontal diffusion via the Smagorinsky constant is to nudge the gray zone solution to the mesoscale one, as shown in both temperature and wind speed profiles. As resolved convection is considerably suppressed from the grid, most of the mixing is done by the PBL scheme as in the 12-km simulation. As noted before, the temperature profile remains slightly superadiabatic when vertical fluxes are mostly calculated by the local PBL scheme.

The partitioning between subgrid and resolved temperature and momentum fluxes is presented in Fig. 3. The REF solution shows almost linearly decreasing heat fluxes in the boundary layer, as typical during strong convection. The two P1 gray zone simulations produce total temperature fluxes that are comparable to REF, despite showing some differences in the surface layer and in the entrainment zone. As the model resolution increases, subgrid fluxes are progressively dampened even if no scale-awareness is present in the PBL scheme. Resolved turbulent fluctuations in d04 diffuse heat and momentum in the boundary layer (i.e., mix the boundary

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

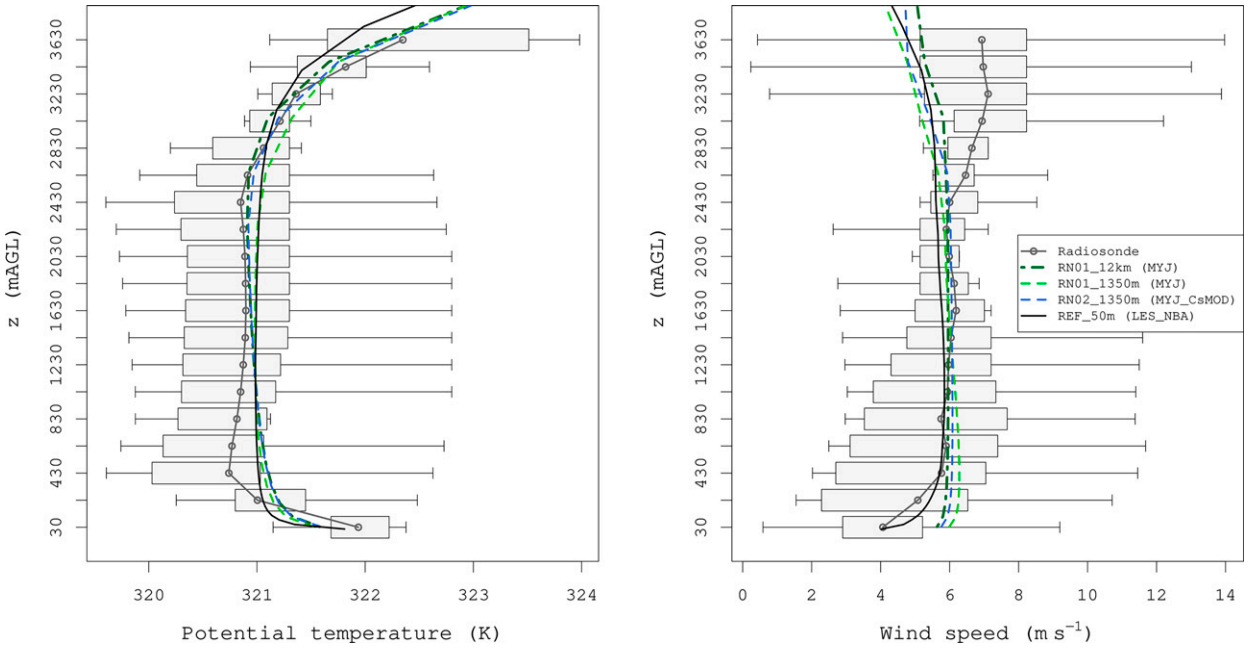

FIG. 2. Mean characteristics of the boundary layer: (left) potential temperature and (right) wind speed vs altitude (z; m AGL), according to radiosonde data, REF, and P1 model simulations. Both modeled and observed data are averaged across nine time points (i.e., at 1500 LST 23 Jul–1 Aug). Boxes indicate the first and third quartile of radiosonde data for each linearly spaced height range, whereas whiskers extend to the minimum and maximum across the 9 days of observations.

layer), thereby reducing shear and buoyant production of TKE owing to lower mean gradients of temperature and velocities. Lower subgrid TKE entails lower eddy viscosities (Fig. S15), which explains the *implicit scale-awareness* of MYJ observed in Figs. 3 and S15. As expected, significantly suppressing resolved convection from the grid considerably reduces the resolved fluxes and brings the solution closer to the mesoscale one. All simulations underestimate REF momentum fluxes, although the MYJ gray zone partly mitigates the difference. Resolved fluxes in RN01 overcompensate the decrease in eddy viscosities related to higher resolution, thus producing larger fluxes that in total become more comparable to REF. In the entrainment zone, the MYJ scheme underestimates the magnitude of turbulent fluxes for both the mesoscale and the gray zone simulation, given the local nature of MYJ calculations. However, the problem is partially mitigated at gray zone resolutions, as heat and momentum can be entrained at resolved scales (Fig. 3). It appears that underresolved convective structures at gray zone resolutions effectively complement local schemes by adding nonlocal transport

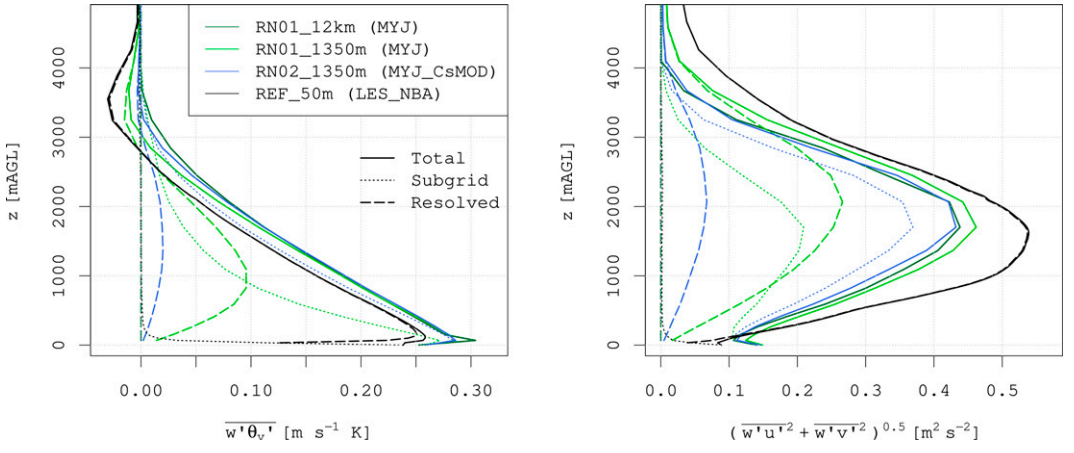

FIG. 3. Partitioning of vertical fluxes of (left) virtual potential temperature and (right) horizontal momentum between resolved (dashed) and subgrid (dotted) scales for P1 simulations. Values are averaged over all the free convection instances ($-z_i/L > 50$), and colors have the same meaning as in Fig. 2.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

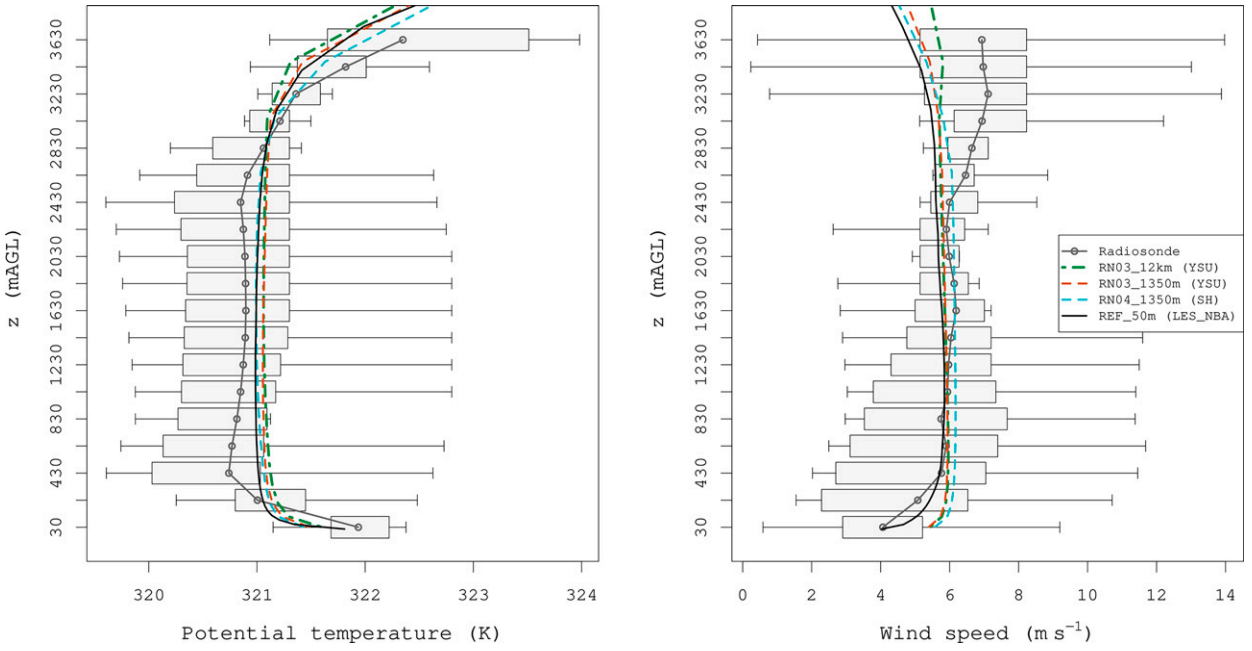

FIG. 4. Mean characteristics of the boundary layer: (left) potential temperature and (right) wind speed vs altitude (z; m AGL), according to radiosonde data, REF, and P2 model simulations. Both modeled and observed data are averaged across nine time points (i.e., at 1500 LST 23 Jul–1 Aug). Boxes indicate the first and third quartile of radiosonde data for each linearly spaced height range, whereas whiskers extend to the minimum and maximum across the 9 days of observations.

at resolved scales, which achieve a (i) a more neutral mixed layer, as previously discussed, and (ii) improved representation of entrainment fluxes.

b. *The effect of scale awareness in a nonlocal scheme (P2 simulations)*

In this section, we analyze P2 simulations to investigate the alternative approach of retaining the resolved convective structures while introducing scale-awareness in nonlocal PBL schemes.

P2 simulations produce very similar CBLs in terms of mean quantities (Fig. 4), with slight differences in the entrainment zone due to explicit parameterization of entrainment fluxes in YSU and SH [Eq. (3)]. Compared to P1, P2 simulations achieve a neutral temperature profile and show improvements in representing wind speed in the surface layer. This slight performance enhancement confirms that nonlocal schemes are typically better equipped than local schemes to deal with strong convection as they include the effect of nonlocal fluxes related to large eddies, also at gray zone resolutions. Introducing scale-awareness seems to contribute to modest improvements in representing the wind speed profile in the surface layer and in the mixed layer. Potential temperature profiles generated with and without scale-awareness are virtually indistinguishable, except for some minor differences in the entrainment zone.

Scale-awareness plays a larger role for turbulent quantities. The SH scheme develops stronger resolved updrafts and downdrafts than the traditional YSU scheme (Fig. 5), as it removes a significant fraction of subgrid fluxes, allowing for more vigorous resolved turbulent structures. In specific instances of time characterized by weaker convection ($-z_i/L \sim 50$, as exemplified in Fig. 5), YSU struggles to break down larger structures coming from the parent domain at gray zone resolutions (Δx = 450 m), whereas SH generates more realistic turbulence patterns. Damping subgrid fluxes and allowing three-dimensional structures seems therefore helpful at these resolutions, although a minor impact is observed on the vertical profiles of mean quantities.

The reason why small differences are noticed in the mean structure of the boundary layer between YSU and SH at gray zone resolutions can be investigated by looking at the temperature and momentum flux profiles and their partitioning between SGS and resolved scales (Fig. 6). For buoyancy, the total amount of turbulent mixing with the two parameterizations is very similar, which results in almost indistinguishable temperature profiles. In other words, the larger subgrid fluxes in YSU compensate for the lower resolved mixing from turbulent fluctuations, yielding a very similar profile of total virtual potential temperature flux. Larger differences are observed between SH and YSU total momentum mixing, with SH generating total momentum fluxes that more closely resemble REF. This result agrees with P1 simulations, which show that momentum fluxes are more similar to REF when the model is allowed to resolve more convection, as REF produces stronger fluxes than the mesoscale solution. As noted before, SH decreases subgrid diffusion allowing for larger resolved convection (i.e., the scheme brings the solution closer to LES than mesoscale), and therefore to a closer agreement with the REF momentum fluxes.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

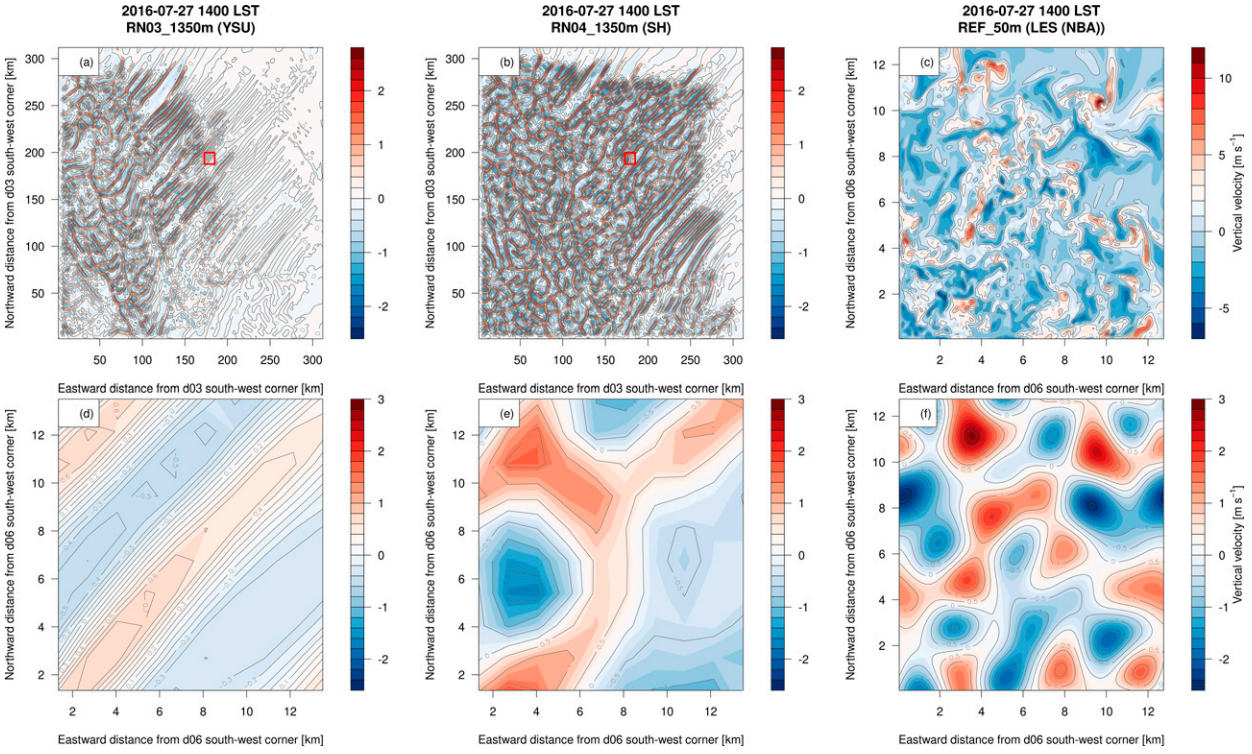

FIG. 5. Contours of vertical velocity (m s^{-1}) at 1400 LST 27 Jul 2016 according to P2 model simulations, approximately 600 m above the ground. (a)–(c) Vertical velocity on the native model grid [d03 for (a) and (b); d06 for (c), on a different color scale]. (d)–(f) The same contours clipped on d06 domain [which is indicated by the red box in (a)–(c)]. A top-hat filter is applied in (f) to allow for a proper comparison with (d) and (e) [i.e., (f) is the filtered version of (c)].

c. *The effect of scale awareness in a local scheme (P3 simulations)*

The impact of introducing scale-awareness in a local PBL scheme at gray zone resolutions is presented in this section. The mean structure of the boundary layer according to P3 simulations (both at mesoscale and gray zone resolution), REF and the radiosonde data are shown in Fig. 7. As noted for P1 simulations, the mesoscale solution produces a clear superadiabatic layer in the middle of the boundary layer with local schemes, but resolved convection at gray zone resolutions adds nonlocal mixing that makes the temperature profile more neutral. Differences between the MYNN2.5 gray zone

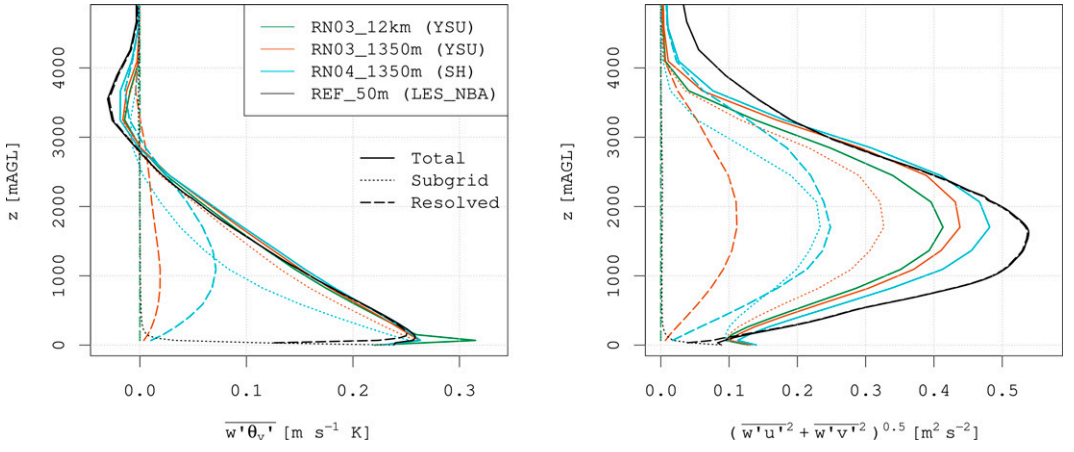

FIG. 6. Partitioning of vertical fluxes of (left) virtual potential temperature and (right) horizontal momentum between resolved (dashed) and subgrid (dotted) scales for P2 simulations. Values are averaged over all the free convection instances ($-z_i/L > 50$). Colors have the same meaning as in Fig. 4.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

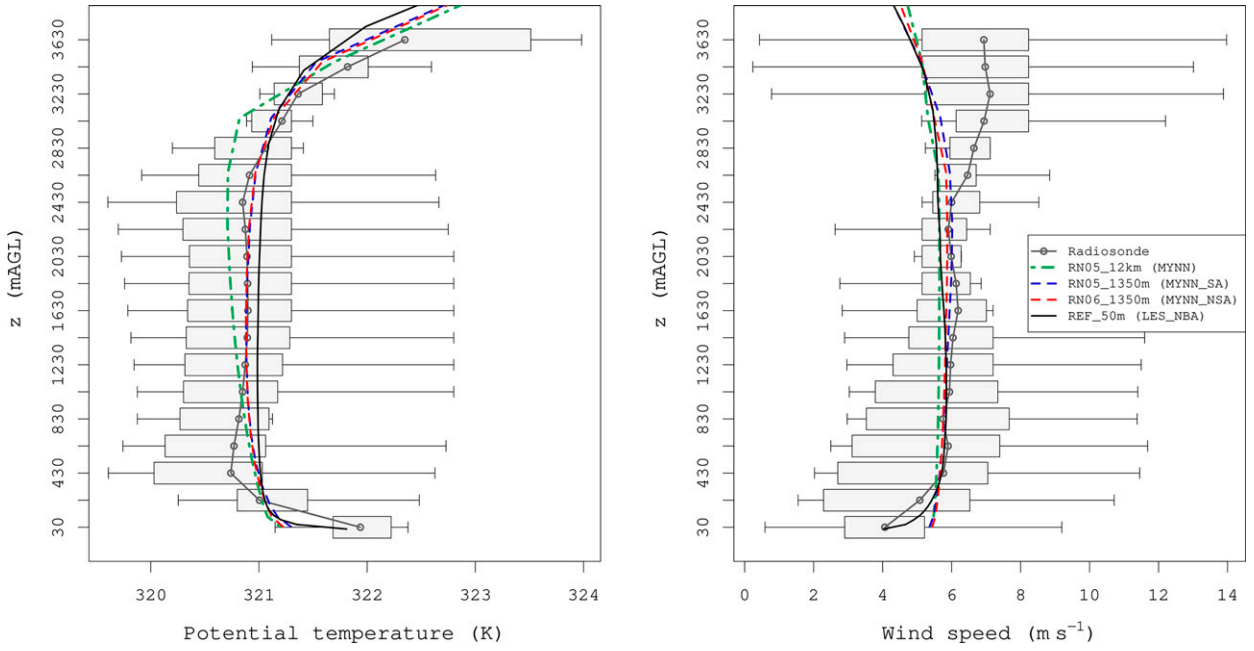

FIG. 7. Mean characteristics of the boundary layer: (left) potential temperature and (right) wind speed vs altitude (z; m AGL), according to radiosonde data, REF, and P3 model simulations. Both modeled and observed data are averaged across nine time points (i.e., at 1500 LST 23 Jul–1 Aug). Boxes indicate the first and third quartile for each linearly spaced height range and whiskers extend to the minimum and maximum across the 9 days of observations.

simulation and its scale-aware version are instead almost negligible for both potential temperature and wind speed. In the near–gray zone regime (Δx = 450 m), introducing scale-awareness improves the wind speed predictions (Fig. S8), as the applicability of a non-scale-aware scheme becomes more questionable for such high resolution. The improvement of the scale-aware version is related to better resolved three-dimensional structures as explained in more details in section 4d. Differences in turbulent structures generated when diffusion is calculated with 1D PBL schemes and 3D closures (for REF and P4 simulations) are shown in Figs. S16 and S17 by means of TKE contours.

Similarly to P2 simulations, the resemblance of scale-aware and not scale-aware MYNN simulations for first-order statistics can be explained by the vertical fluxes of virtual temperature and horizontal momentum (Fig. 8). Although the partitioning of turbulent fluxes between resolved and subgrid scales is different at gray zone resolutions (with differences becoming more pronounced at near–gray zone resolutions, Fig. S9), the total amount of mixing—for both buoyancy and

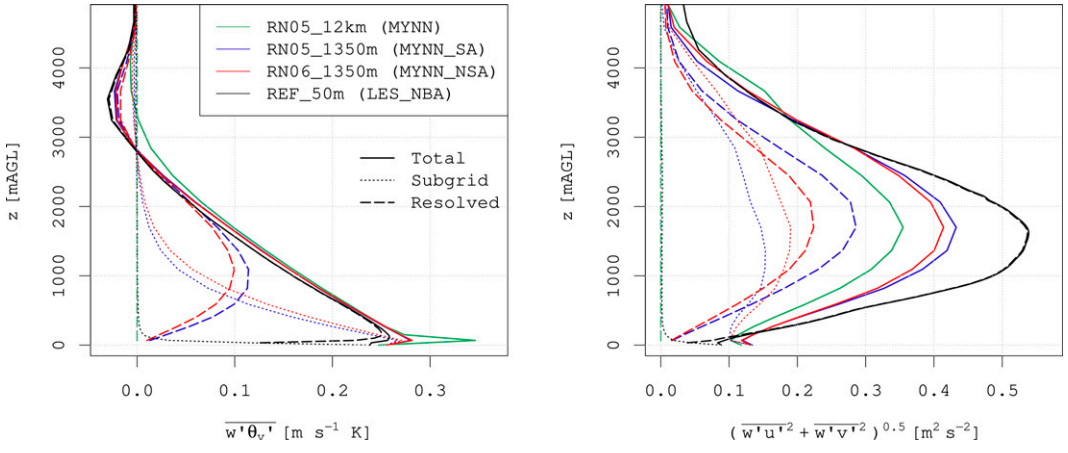

FIG. 8. Partitioning of vertical fluxes of (left) virtual potential temperature and (right) horizontal momentum between resolved (dashed) and subgrid (dotted) scales for P3 simulations. Values are averaged over all the free convection instances ($-z_i/L > 50$). Colors have the same meaning as in Fig. 7.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

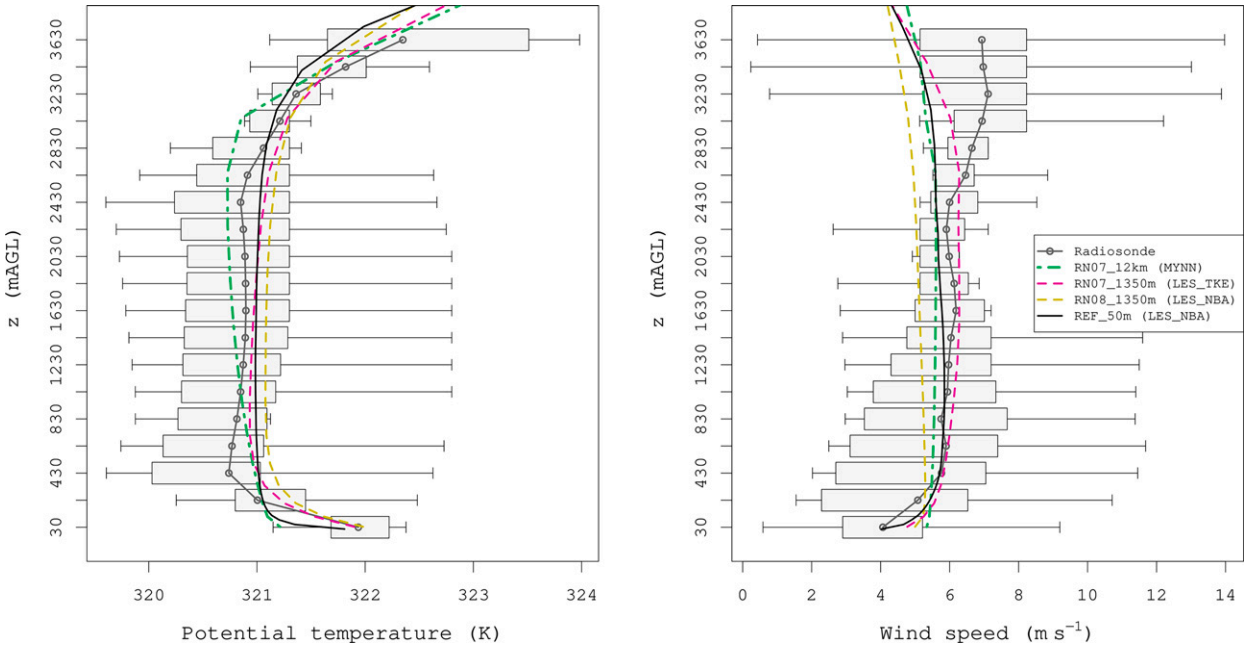

FIG. 9. Mean characteristics of the boundary layer: (left) potential temperature and (right) wind speed vs altitude (z; m AGL), according to radiosonde data, REF, and P4 model simulations. Both modeled and observed data are averaged across 9 time points (i.e., 1500 LST 23 Jul–1 Aug). Boxes indicate the first and third quartile for each linearly spaced height range and whiskers extend to the minimum and maximum across the 9 days of observations.

momentum—is very similar. As discussed in section 4a, more resolved convection entails lower subgrid fluxes, and vice versa. In the entrainment zone and above the boundary layer, turbulent fluxes are reasonably represented by P3 gray zone simulations compared to REF, although none of the simulations (including REF) agree well with the wind speed profile measured by radiosonde, suggesting a potential problem in the mesoscale forcing rather than in turbulence parameterizations above the boundary layer.

d. 3D turbulence closures

The TKE-based 3D turbulence closure at Δx = 1350 m considerably improves the model performance in the surface layer, for both temperature and wind speed, when compared to the mesoscale solutions and other gray zone simulations (Fig. 9). It is worth noting that the vertical discretization is the same for all gray zone and mesoscale simulations, which entails that the differences in the surface layer profiles are entirely related to turbulent diffusion. The reason for the different behavior of RN07 compared to the other gray zone simulations is that P1, P2, and P3 runs are all based on PBL schemes, and therefore they calculate mixing based on the mean vertical shear only, whereas the 3D approach prescribes stresses and diffusivities from three-dimensional turbulence at the resolved scales. Although the NBA approach (RN08) provides a slightly better representation of the vertical shear in the surface layer, the performance in RN08 deteriorates in the mixed layer and in the entrainment zone, hinting to a different representation of the entrainment processes that drive the boundary layer growth in the NBA model. This result differs from Mirocha et al. (2010), which implemented and tested the NBA approach for an idealized simulation over a symmetrical 2D hill in neutral conditions.

The partitioning of turbulent fluxes further clarifies the results shown in Fig. 9. Compared to REF, temperature and momentum fluxes are better reproduced with the TKE-based closure than with PBL schemes in the surface layer (Fig. 10). Owing to the different turbulent transport in the lower part of the atmosphere, lower values of friction velocities and surface temperature scales (i.e., surface momentum and heat fluxes) produce an improved bottom boundary condition for velocity and temperature, which is in closer agreement with radiosonde data. However, turbulent fluxes are underestimated in the mixed layer as 3D closures lack part of the nonlocal transport that in 3D schemes is assumed to be completely resolved, but that at gray zone resolutions is only partially resolved. Finally, momentum fluxes calculated with the NBA model are considerably underestimated compared to REF. The largest differences with REF and the 3D TKE model are observed in the middle of the boundary layer (Fig. 10), as subgrid stresses are calculated based on the strain rate tensor for the NBA model rather than TKE. Although a version of NBA based on TKE was implemented in WRF by Mirocha et al. (2010) (which could be more comparable to the 1.5 TKE subgrid model of Lilly 1967), for nested simulations the strain-rate formulation is recommended given the mismatch of prognostic TKE at nest interfaces (Mirocha et al. 2010).

e. Simulation performance summary

Figures 11 and 12 summarize the performance of the whole simulation set in a quantitative way for potential temperature,

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

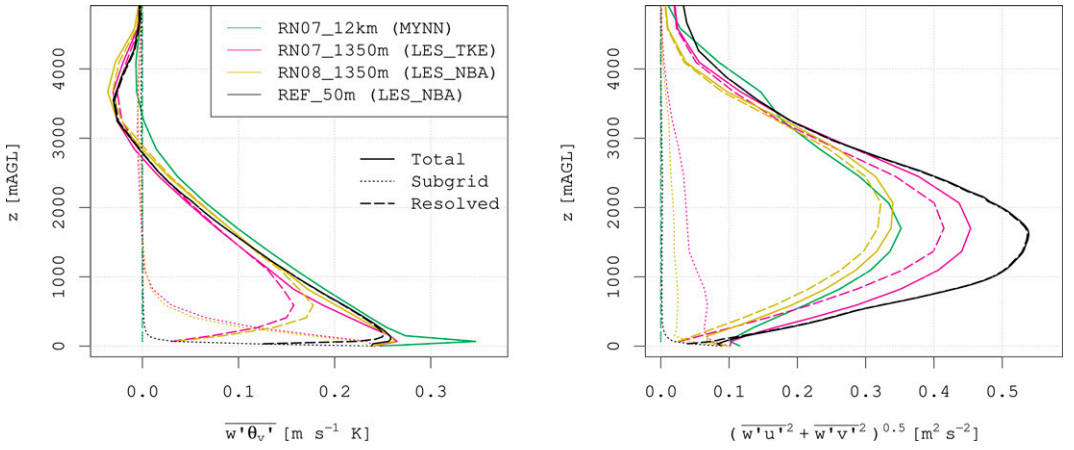

FIG. 10. Partitioning of vertical fluxes of (left) virtual potential temperature and (right) horizontal momentum between resolved (dashed) and subgrid (dotted) scales for P4 simulations. Values are averaged over all the free convection instances ($-z_i/L > 50$). Colors have the same meaning as in Fig. 9.

wind speed and turbulent fluxes (along with Tables S1 and S2, which report the actual values of RMSEs). Similar figures showing the near–gray zone regime can be found in the supplementary material (Figs. S18 and S19). We break down these results for the lower part of the boundary layer ($z/z_i \leq 0.2$), the mixed layer ($0.2 < z/z_i \leq 0.8$) and the entrainment zone ($0.8 < z/z_i \leq 1.0$).

Overall, errors related to first-order statistics of the flow at gray zone resolutions are of the same order of magnitude as mesoscale and microscale solutions, and they show a reasonable sensitivity to different turbulence parameterizations. This result suggests that simulations at gray zone resolutions do not influence the mean vertical structure of the fully developed boundary layer more than other important factors in real simulations (e.g., lateral and bottom boundary conditions, imperfect representation of other processes). It should be noted that modeled profiles are compared to radiosonde data at 1500 LST, when the boundary layer is fully developed, and that these results may change during the growth of the boundary layer, as shown later in this section. Moreover, Figs. 11 and 12 reveal that there

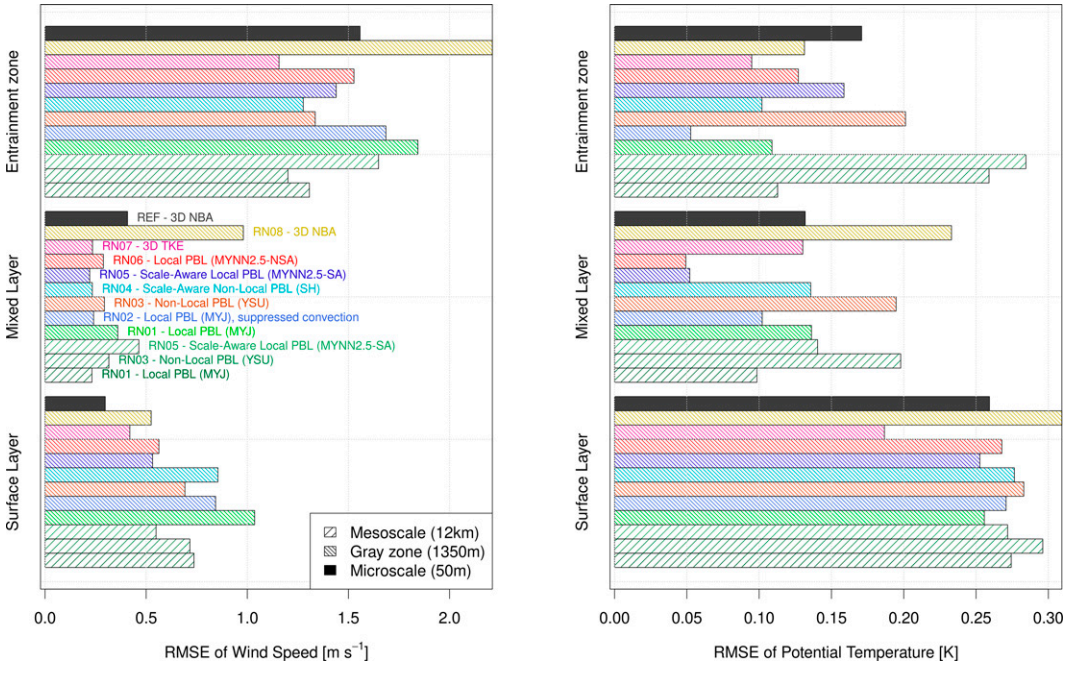

FIG. 11. Root-mean-square error (RMSE) of wind speed and potential temperature in the surface layer ($z/z_i \leq 0.2$), in the mixed layer ($0.2 < z/z_i \leq 0.8$), and in the entrainment zone ($0.8 < z/z_i \leq 1.0$), for P1, P2, P3, P4, and REF simulations. RMSE is calculated against radiosonde data.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

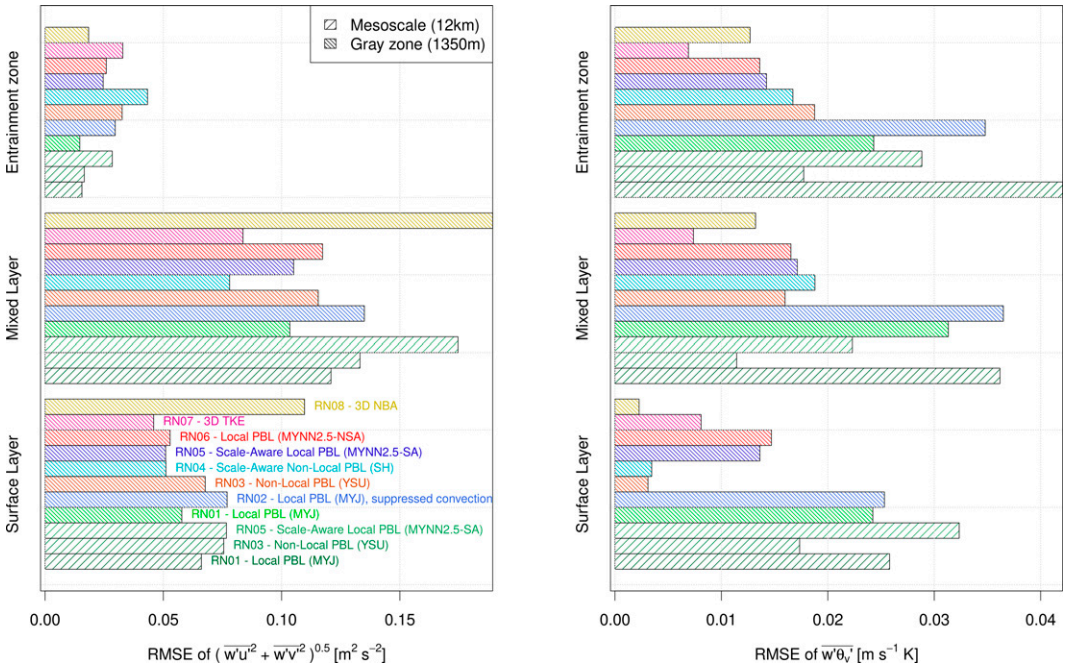

FIG. 12. Root-mean-square error (RMSE) of horizontal momentum and virtual potential temperature ($\overline{w'\theta_v'}$) fluxes in the surface layer ($z/z_i \leq 0.2$), in the mixed layer ($0.2 < z/z_i \leq 0.8$), and in the entrainment zone ($0.8 < z/z_i \leq 1.0$), for all the simulations presented in the results section. RMSE is calculated against REF simulation.

is no universal best setup at gray zone resolutions, and results may vary depending on the specific quantity of interest. In the surface layer, the microscale LES solution shows the lowest RMSE for wind speed, which represents approximately a 60% improvement compared to the mesoscale solutions. Among gray zone simulations, the MYNN local scheme produces the most realistic boundary layer structure at gray zone resolutions, with mild improvements given by its scale-aware version. However, the TKE-based 3D closure outperforms all other modeling strategies in the surface layer (Figs. S20 and S21), as also discussed in the previous sections. The explicit treatment of entrainment processes in the YSU scheme (and in its scale-aware version, SH) generates the best agreement in the entrainment zone, even at gray zone resolutions. Despite the presence of grid-dependent convective structures in the gray zone solutions (Fig. 13), total turbulent fluxes are generally better represented at gray zone resolutions compared to the mesoscale ones, where all turbulent motions are calculated with subgrid parameterizations. At near–gray zone resolutions, the benefits of the 3D-TKE closure over the parameterized schemes become more evident in the surface layer (Figs. S20 and S21), as $\Delta x/z_i$ becomes closer to the LES regime.

The grid dependency problem for all simulations at gray zone and near–gray zone resolutions is illustrated in Fig. 13, which shows the characteristic length scales of vertical velocity for all simulations, except for RN02 (since resolved convection is suppressed). At the top of the surface layer, the characteristic wavelength of vertical velocity fluctuations (Λ_c) produced by gray zone simulations and near–gray zone simulations approximately scales as $2.80z_i$ and $1.58z_i$, respectively, regardless of the turbulence parameterizations, whereas well-resolved LES solutions (both at 150- and 50-m resolution) show Λ_c of the order of the boundary layer height (Fig. 13a). Thermal structures become larger with altitude up to the middle of the boundary layer, which agrees with Kaimal et al. (1976) observations that revealed that peaks in the vertical velocity spectrum occur at larger wavelengths aloft, for the lower part of the boundary layer. While this aspect is well captured by all simulations, length scales for the Δx = 1350 m solution are significantly larger than the Δx = 450 m and REF (Δx = 150, 50 m), which instead show a good agreement between each other. This result confirms that Δx = 150 m can be already considered well resolved in our case study, whereas Δx = 450 m can be classified as near–gray zone resolution.

To understand the impact of horizontal structures on the mean vertical profiles, we further analyze the degree of anisotropy (α) and the resolved energy (σ_w^2) of the convective structures produced by each simulation, according to the methodology outlined in section 3d. We analyze all instances in time characterized by $-z_i/L > 50$, condition for which the LES solution shows visually isotropic convective structures (example in Fig. 5c). Figure 14a reveals that more than 50% of the convective structures are not classified as fully developed (despite being characterized by $-z_i/L > 50$) for parameterized simulations (RN01, RN03, RN06), whereas the frequency of fully developed structures is higher for 3D closures (RN07, RN08) and scale-aware schemes (RN04, RN05). In addition, Fig. 14b shows that the difference in potential temperature vertical profiles between REF and each simulation

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

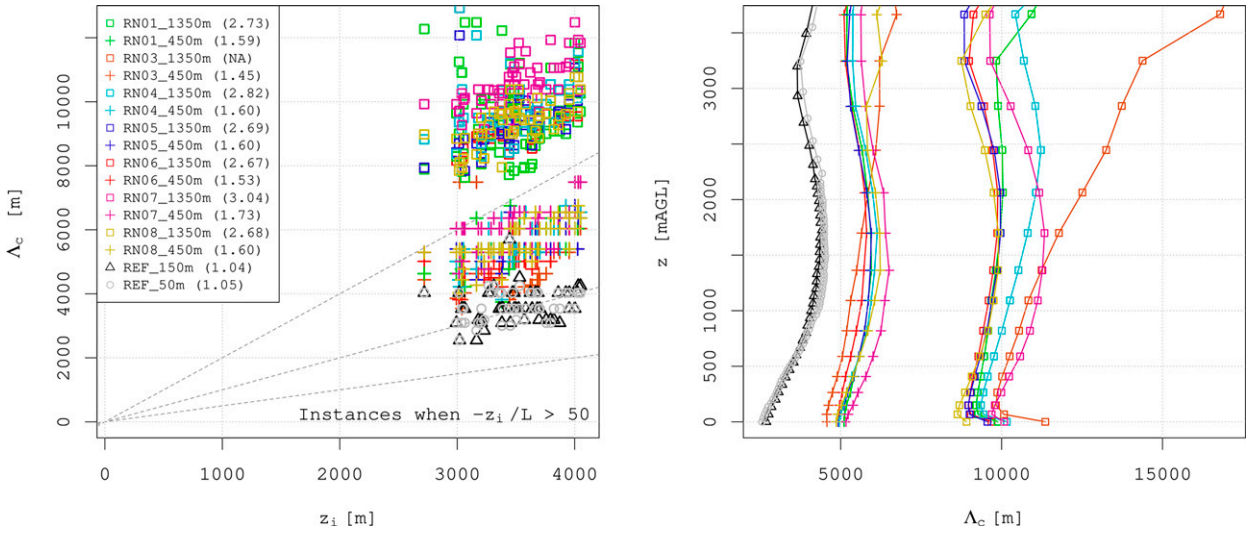

FIG. 13. Vertical velocity spectral peak Λ_c as a function of (left) boundary layer height z_i at $z \sim 500$ m and (right) average vertical profiles of Λ_c for P1, P2, P3, P4, and REF simulations. In the left panel, dotted lines represent the 1:2, 1:1, and 2:1 lines and numbers in parentheses indicate the mean Λ_c/z_i ratio. Only time instances characterized by $-z_i/L > 50$ and by convective structures larger than the effective resolution of the model ($6\Delta x$, as in Zhou et al. 2014) are considered.

is typically larger when the convective structures are under-resolved or transitional. The mean RMSE difference between fully developed and non-fully developed structures (i.e., transitional and under-resolved) is statistically significant at a 5% significant level with the Welch two sample *t*-test for RN01, RN03 and RN04. The statistical test is performed by pooling together the RMSE values from the transitional and under-resolved categories to ensure that all the simulations have enough data points to perform a meaningful statistical test. This result suggests that (i) there is a connection between vertical and horizontal structures and (ii) errors are larger during the growth of the boundary layer especially for simulations that struggle to develop convective cells under strong convection (i.e., RN01, RN03 and RN04).

An illustrative example showing the comparison between RN01 and RN07 is presented in Fig. 15, whereas a comparison between YSU and its scale-aware version can be found in the supplementary material (Fig. S22).

As the boundary layer grows from late morning to midafternoon, convective structures in RN01 become more isotropic and approach the fully developed regime, whereas they remain fully developed in RN07 throughout the time period.

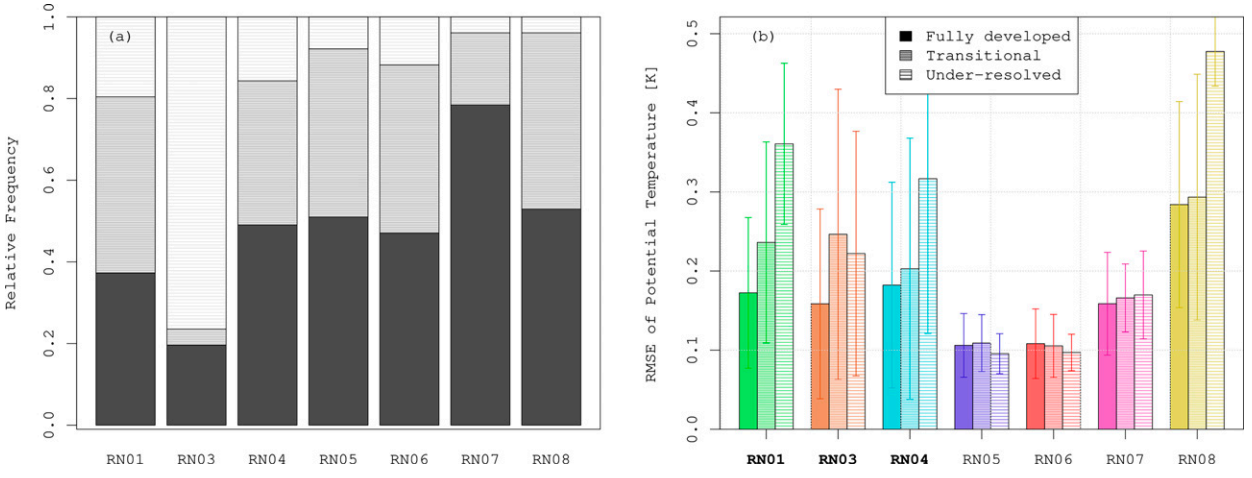

FIG. 14. (a) Classification of horizontal convective structures for each simulation, according to the methodology presented in section 3d, and (b) RMSE of potential temperature vertical profiles with respect to REF when horizontal structures are under-resolved, transitional, and fully developed. Error bars extend up to two standard deviations from the mean. All instances in time are characterized by $-z_i/L > 50$. Simulations characterized by statistically significant RMSE differences between fully developed and non-fully developed structures are highlighted in bold in (b).

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

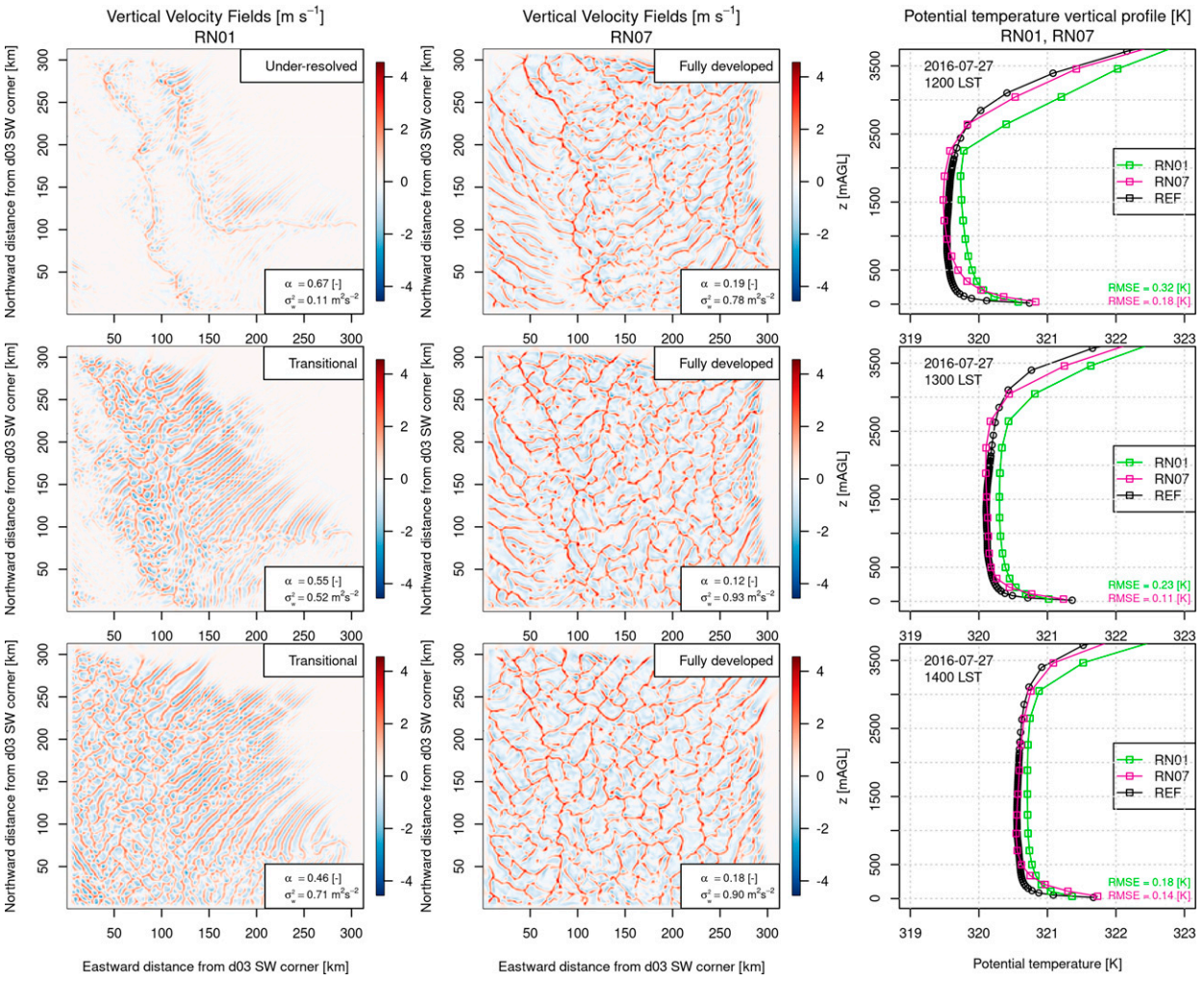

FIG. 15. (Left) RN01 vertical velocity fields in d03, (center) RN07 vertical velocity fields in d03, and (right) vertical profiles of potential temperature for selected times on 27 Jul 2016. Different rows represent the different time instances (1200–1400 LST, shown from top to bottom). The degree of anisotropy α and the resolved energy σ_w^2 are reported in the left and center panels.

Alongside the development of convective structures, the differences in vertical profiles between REF and the RN01 become smaller and the vertical structure of the boundary layer becomes comparable. At 1200 LST, the boundary layer in RN01 is still shallower compared to REF, despite developing up to the depth of REF later in the day. In contrast, the faster turbulence spinup in RN07 generates vertical structures that more closely resemble REF. A similar result is found when introducing scale-awareness in a nonlocal scheme, although the differences between RN03 and RN04 are more limited (Fig. S22). While scale-awareness has a noticeable impact on the horizontal structures (Fig. S22 and Fig. 5), it has a comparatively smaller effect on the vertical structures as nonlocal transport of heat and momentum can be achieved even with zero resolved turbulence (as also discussed in section 4b).

Overall, the illustrative example presented here (Figs. 15 and S22) points out the problem in the timing of convection initiation and shows that the 1D parameterization does not develop the boundary layer as quickly as the resolved convection in RN07, RN04, and REF, even in the presence of strong heat fluxes.

5. Discussion and conclusions

In this study, we run a set of eight numerical experiments aimed at exploring different avenues to address the gray zone of turbulence for real and full-physics simulations, during strongly convective conditions. Overall, first-order statistics of the flow at gray zone resolutions compare well with mesoscale and LES fully developed CBLs, and they show a modest sensitivity to different turbulence parameterizations for the analyzed case. Turbulent quantities vary more considerably among the eight experiments. If setup properly, gray zone simulations and microscale LES simulations driven by realistic online boundary conditions can effectively improve the

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

model performance, despite the presence of under-resolved and grid-dependent convective structures in the resolved fields. This is especially true for model results in the surface layer, which are critical for a range of practical purposes (e.g., wind energy and air pollution dispersion). RMSE of wind speed in the surface layer shows a twofold decrease in the best-performing gray zone simulations compared to traditional mesoscale solutions. Combining the results presented in the previous section and the findings of other recent works, here we summarize the main evidence and recommendations for real gray zone simulations, with respect to the objectives outlined in section 1.

a. PBL schemes and turbulence parameterizations

We provide further evidence that under-resolved convective structures at gray zone resolutions in real simulations are grid-dependent. The size of these structures at the top of surface layer approximately scales as $\sim 1 z_i$ in well-resolved LES simulations ($\Delta x = 50$ m), whereas simulations with $\Delta x = 1350$ and 450 m generate larger structures ($2.80 z_i$ and $1.58 z_i$, respectively). Although the size and strength of the turbulent structures at gray zone resolutions can be quite different from a filtered LES solution, we argue that allowing grid-dependent resolved convection entails minor and possibly positive impacts on mean quantities and total turbulent fluxes. We stress that resolved convection should be looked at in combination with the subgrid fluxes produced by SGS parameterizations, because the nature of the modeled convective structures is intrinsically tied with the 1D PBL (or 3D) parameterization adopted. Considering only one of the two components (i.e., resolved or subgrid) would overlook the total mixing done in the numerical model, which is what drives the boundary layer dynamics. Heat fluxes are well reproduced when allowing resolved convection, owing to the *implicit scale-awareness* found for all PBL schemes analyzed in this work. For local schemes, another cobenefit of having resolved convection is to add some nonlocal mixing at the resolved scales, achieving more neutral temperature profiles in the mixed-layer and better representation of the entrainment processes. Introducing scale-awareness in 1D PBL schemes proves useful in improving momentum fluxes predictions as well as the partitioning between resolved and subgrid scales, although the total fluxes are similar with and without scale-awareness. The positive impact of scale-awareness schemes becomes more evident for near–gray zone simulations, where the validity of the assumptions behind standard PBL schemes become more questionable.

For deep CBLs, we illustrate how a TKE-based 3D turbulence closure outperforms 1D PBL schemes both at gray zone and near–gray zone resolutions in the surface layer, as it allows for 3D structures to develop and to better represent mixing. The performance of the recently proposed 3D NBA model is instead less satisfactory for the mixed layer and the entrainment zone, although the sensitivity to different 3D closures is not thoroughly investigated in this work. Mounting evidence from other works suggests that these findings (i.e., 3D closures outperforming 1D PBL schemes) are true also for moist convection (Fiori et al. 2010), in complex terrain (Liu et al. 2020) and for weaker convective conditions (Johnson and Wang 2019), although Doubrawa and Muñoz-Esparza (2020) found different results over the Mountain West of the United States.

Finally, we show that artificially increasing horizontal diffusion via the Smagorinsky coefficient in WRF is an efficient way of considerably damping resolved convection. However, in our case study characterized by small land heterogeneity in the finest domain, attenuating the strength of under-resolved structures from the solution effectively nudges the gray zone solution to the mesoscale one, which is faster and cheaper as highlighted by Zhou et al. (2014), thereby questioning the need for such approach.

b. Domains setup

Real high-resolution simulations—both at gray zone and LES resolutions—require careful considerations to properly set up the numerical domains. As resolution increases, high-resolution topography characteristics begin to be resolved, including steep slopes that would be otherwise averaged out in traditional mesoscale grids. Most mesoscale models (including WRF) use terrain-following coordinates, which suffer from coordinate transformation errors in the presence of steep elevation gradients that can significantly degrade the performance of the numerical solution (Lundquist et al. 2010).

A thorough sensitivity analysis (not shown here) reveals that avoiding large topographic features and steep elevation gradients along the boundaries of each domain is key to ensure the numerical stability of the solutions. At gray zone resolutions, another important aspect is to leave necessary fetches from the edges of the domain for turbulence spin-up in the area of interest (Muñoz- Esparza et al. 2014), with special care to be taken when moving from mesoscale to gray zone domains. Given then large heat fluxes that drive our simulations during convective hours, turbulent motions are triggered even close to the boundaries, especially with scale-aware schemes and 3D closures. Simulations in mountainous or urban environments at microscale resolutions ($\Delta x \ll l$), where bypassing the problem of abrupt topography changes becomes impractical, may require the application of different methodologies, such as the immersed boundary method (IBM) that eliminates the coordinate transformation (Arthur et al. 2020; Bao et al. 2018).

Some important aspects need to be considered when selecting vertical levels. At gray zone scales, we find that 49 vertical levels, with power-law decreasing Δz approaching the surface and near constant Δz above the boundary layer, guarantees numerical stability. Numerical instabilities and unphysical numerical waves arise with significant jumps in vertical layer thicknesses (as in Shin and Dudhia 2016), or with an excessively large or small number of vertical levels close to the surface (Rai et al. 2017). For microscale LES simulations at $\Delta x = 150$ and 50 m, a nearly constant value of $\Delta z = 64$ and 32 m in the boundary layer is found to be stable and to allow explicit resolution of three-dimensional turbulence features. In the presence of relatively large elevation gradients (as in

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

our d04, with terrain slopes greater than 0.1), increasing time off-centering (epssm in WRF) to 0.8 is found to stabilize vertically propagating high-frequency waves that may cause numerical instabilities. This approach entails no extra cost and has shown no significant effect on nonacoustic modes (Durran and Klemp 1983; Skamarock et al. 2008). More details about domains setup and vertical levels are available in the namelist settings of our simulations, which are made publicly available through Github (see data availability).

c. *Open knowledge gaps and future work*

Although the main focus of this work is to compare the performance of different modeling strategies in calculating the mean vertical structure of the fully developed boundary layer at gray zone resolutions, future work should target other equally important aspects of gray zone modeling. For instance, further understanding the link between horizontal structures and errors in vertical profiles during the growth of the boundary layer, as briefly exemplified in this work (Figs. 14 and 15), will be paramount to improve predictions during the morning transition to a fully developed CBL. Previous work in idealized settings showed that, during early stages of convection, total heat fluxes at gray zone resolutions can greatly differ from LES solutions despite converging after sufficiently long time (Efstathiou et al. 2016; Kealy et al. 2019). For this purpose, additional simulations using the cell perturbation method developed by Muñoz-Esparza et al. (2014) or the model proposed by Kealy et al. (2019) should be included. To allow for a seamless coupling between mesoscale and LES simulations, a full sensitivity analysis on the influence of gray zone boundary conditions on microscale LES performance should be also performed in real settings (Mazzaro et al. 2017).

For further model developments, our findings suggest that 3D closures are more appropriate at gray zone resolutions as they produce more realistic convective structures (despite being grid-dependent in size) and more accurate results in the surface layer. Nonetheless, 3D closures typically assume that the largest eddies responsible for nonlocal transport are resolved on the grid, which may not be completely valid at gray zone resolutions. A comprehensive scheme that can integrate the right amount of vertical nonlocal transport with local 3D fluxes could be therefore a potential avenue to move forward in gray zone modeling. The direction first proposed by Wyngaard (2004) and further explored by Zhang et al. (2018) seems promising.

Finally, the methodology and results presented here provide a reference dataset of real simulations that can be extended to assess how new techniques for simulations at gray zone resolutions compare to the modeling strategies presented herein [e.g., the dynamic reconstruction model of Simon et al. (2019), the mass-flux MYNN-EDMF scheme of Olson et al. (2019), the 3D closure of Zhang et al. (2018), the scale-adaptive one of Kurowski and Teixeira (2018) or the new mixing length model of Honnert et al. (2021)].

Acknowledgments. This study is supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award OSR-CRG 7 2018-3742.2.

Data availability statement. HRES-ECMWF operational analysis data were downloaded from the ECMWF data portal (https://www.ecmwf.int/en/forecasts/datasets/set-i) through the KAUST ECMWF license. Radiosonde data are acquired from the University of Wyoming (http://weather.uwyo.edu/upperair/sounding.html) public repository. All the namelist settings of the simulation set presented in this work are made publicly available through Github (https://github.com/Env-an-Stat-group/22.Giani.MWR).

REFERENCES

Arthur, R. S., K. A. Lundquist, D. J. Wiersema, J. Bao, and F. K. Chow, 2020: Evaluating implementations of the immersed boundary method in the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **148**, 2087–2109, https://doi.org/10.1175/MWR-D-19-0219.1.

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.

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.

Boutle, I. A., J. E. J. Eyre, and A. P. Lock, 2014: Seamless stratocumulus simulation across the turbulent gray zone. *Mon. Wea. Rev.*, **142**, 1655–1668, https://doi.org/10.1175/MWR-D-13-00229.1.

Cécé, R., D. Bernard, J. Brioude, and N. Zahibo, 2016: Microscale anthropogenic pollution modelling in a small tropical island during weak trade winds: Lagrangian particle dispersion simulations using real nested LES meteorological fields. *Atmos. Environ.*, **139**, 98–112, https://doi.org/10.1016/j.atmosenv.2016.05.028.

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.

Chu, X., L. Xue, B. Geerts, R. Rasmussen, and D. Breed, 2014: A case study of radar observations and WRF LES simulations of the impact of ground-based glaciogenic seeding on orographic clouds and precipitation. Part I: Observations and model validations. *J. Appl. Meteor. Climatol.*, **53**, 2264–2286, https://doi.org/10.1175/JAMC-D-14-0017.1.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, and H. E. Brooks, 2015: A review of planetary boundary layer parameterization schemes and their sensitivity in simulating southeastern U.S. cold season severe weather environments. *Wea. Forecasting*, **30**, 591–612, https://doi.org/10.1175/WAF-D-14-00105.1.

Conry, P., A. Sharma, M. J. Potosnak, L. S. Leo, E. Bensman, J. J. Hellmann, and H. J. S. Fernando, 2015: Chicago's heat island and climate change: Bridging the scales via dynamical downscaling. *J. Appl. Meteor. Climatol.*, **54**, 1430–1448, https://doi.org/10.1175/JAMC-D-14-0241.1.

Daniels, M. H., K. A. Lundquist, J. D. Mirocha, D. J. Wiersema, and F. K. Chow, 2016: A new vertical grid nesting capability

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

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., 1972: Numerical investigation of neutral and unstable planetary boundary layers. *J. Atmos. Sci.*, **29**, 91–115, https://doi.org/10.1175/1520-0469(1972)029<0091:NIONAU>2.0.CO;2.

de Roode, S. R., P. G. Duynkerke, and H. J. J. Jonker, 2004: Large-eddy simulation: How large is large enough? *J. Atmos. Sci.*, **61**, 403–421, https://doi.org/10.1175/1520-0469(2004)061<0403:LSHLIL>2.0.CO;2.

Dimitrova, R., Z. Silver, T. Zsedrovits, C. M. Hocut, L. S. Leo, S. Di Sabatino, and H. J. S. Fernando, 2016: Assessment of planetary boundary-layer schemes in the weather research and forecasting mesoscale model using MATERHORN field data. *Bound.-Layer Meteor.*, **159**, 589–609, https://doi.org/10.1007/s10546-015-0095-8.

Donateo, A., D. Cava, and D. Contini, 2017: A case study of the performance of different detrending methods in turbulent-flux estimation. *Bound.-Layer Meteor.*, **164**, 19–37, https://doi.org/10.1007/s10546-017-0243-4.

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.

Durran, D. R., and J. B. Klemp, 1983: A compressible model for the simulation of moist mountain waves. *Mon. Wea. Rev.*, **111**, 2341–2361, https://doi.org/10.1175/1520-0493(1983)111<2341:ACMFTS>2.0.CO;2.

Efstathiou, G. A., R. J. Beare, S. Osborne, and A. P. Lock, 2016: Grey zone simulations of the morning convective boundary layer development. *J. Geophys. Res. Atmos.*, **121**, 4769–4782, https://doi.org/10.1002/2016JD024860.

——, R. S. Plant, and M.-J. 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.

Fernando, H. J. S., D. Zajic, S. Di Sabatino, R. Dimitrova, B. Hedquist, and A. Dallman, 2010: Flow, turbulence, and pollutant dispersion in urban atmospheres. *Phys. Fluids*, **22**, 051301, https://doi.org/10.1063/1.3407662.

——, 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.

Fiori, E., A. Parodi, and F. Siccardi, 2010: Turbulence closure parameterization and grid spacing effects in simulated supercell storms. *J. Atmos. Sci.*, **67**, 3870–3890, https://doi.org/10.1175/2010JAS3359.1.

Gamo, M., 1996: Thickness of the dry convection and large-scale subsidence above deserts. *Bound.-Layer Meteor.*, **79**, 265–278, https://doi.org/10.1007/BF00119441.

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.

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.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339, https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

——, 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., 2019: Grey-zone turbulence in the neutral atmospheric boundary layer. *Bound.-Layer Meteor.*, **170**, 191–204, https://doi.org/10.1007/s10546-018-0394-y.

——, 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.

——, and Coauthors, 2020: The atmospheric boundary layer and the "gray zone" of turbulence: A critical review. *J. Geophys. Res. Atmos.*, **125**, e2019JD030317, https://doi.org/10.1029/2019JD030317.

——, V. Masson, C. Lac, and T. Nagel, 2021: A theoretical analysis of mixing length for atmospheric models from micro to large scales. *Front. Earth Sci.*, **8**, 582056, https://doi.org/10.3389/feart.2020.582056.

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.

Janjić, Z. I., 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.

Joe, D. K., H. Zhang, S. P. DeNero, H.-H. Lee, S.-H. Chen, B. C. McDonald, R. A. Harley, and M. J. Kleeman, 2014: Implementation of a high-resolution Source-Oriented WRF/Chem model at the Port of Oakland. *Atmos. Environ.*, **82**, 351–363, https://doi.org/10.1016/j.atmosenv.2013.09.055.

Johnson, A., and X. Wang, 2019: Multicase assessment of the impacts of horizontal and vertical grid spacing, and turbulence closure model, on subkilometer-scale simulations of atmospheric bores during PECAN. *Mon. Wea. Rev.*, **147**, 1533–1555, https://doi.org/10.1175/MWR-D-18-0322.1.

Kaimal, J. C., J. C. Wyngaard, D. A. Haugen, O. R. Coté, Y. Izumi, S. J. Caughey, and C. J. Readings, 1976: Turbulence structure in the convective boundary layer. *J. Atmos. Sci.*, **33**, 2152–2169, https://doi.org/10.1175/1520-0469(1976)033<2152:TSITCB>2.0.CO;2.

Kealy, J. C., G. A. Efstathiou, and R. J. Beare, 2019: The onset of resolved boundary-layer turbulence at grey-zone resolutions. *Bound.-Layer Meteor.*, **171**, 31–52, https://doi.org/10.1007/s10546-018-0420-0.

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.

Kurowski, M. J., and J. Teixeira, 2018: A scale-adaptive turbulent kinetic energy closure for the dry convective boundary layer.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

J. Atmos. Sci., **75**, 675–690, https://doi.org/10.1175/JAS-D-16-0296.1.

Larsén, X. G., S. E. Larsen, and E. L. Petersen, 2016: Full-scale spectrum of boundary-layer winds. *Bound.-Layer Meteor.*, **159**, 349–371, https://doi.org/10.1007/s10546-016-0129-x.

LeMone, M. A., M. Tewari, F. Chen, and J. Dudhia, 2013: Objectively determined fair-weather CBL depths in the ARW-WRF model and their comparison to CASES-97 observations. *Mon. Wea. Rev.*, **141**, 30–54, https://doi.org/10.1175/MWR-D-12-00106.1.

Leonard, A., 1975: Energy cascade in large-eddy simulations of turbulent fluid flows. *Advances in Geophysics*, Vol. 18, Academic Press, 237–248.

Lilly, D. K., 1967: The representation of small-scale turbulence in numerical simulation experiments. NCAR manuscript 281, NCAR, 24 pp.

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.

Mazzaro, L. J., D. Muñoz-Esparza, J. K. Lundquist, and R. R. Linn, 2017: Nested mesoscale-to-LES modeling of the atmospheric boundary layer in the presence of under-resolved convective structures. *J. Adv. Model. Earth Syst.*, **9**, 1795–1810, https://doi.org/10.1002/2017MS000912.

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**, 16 663–16 682, https://doi.org/10.1029/97JD00237.

Moeng, C.-H., 1984: A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *J. Atmos. Sci.*, **41**, 2052–2062, https://doi.org/10.1175/1520-0469(1984)041<2052:ALESMF>2.0.CO;2.

Muñoz-Esparza, D., B. Kosović, J. Mirocha, and J. van Beeck, 2014: Bridging the transition from mesoscale to microscale turbulence in numerical weather prediction models. *Bound.-Layer Meteor.*, **153**, 409–440, https://doi.org/10.1007/s10546-014-9956-9.

Muñoz-Esparza, D., J. K. Lundquist, J. A. Sauer, B. Kosović, and R. R. Linn, 2017: Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign: From weather to boundary-layer eddies. *J. Adv. Model. Earth Syst.*, **9**, 1572–1594, https://doi.org/10.1002/2017MS000960.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31, https://doi.org/10.1023/B:BOUN.0000020164.04146.98.

——, and ——, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, https://doi.org/10.2151/jmsj.87.895.

Olson, J. B., J. S. Kenyon, W. Angevine, J. M. Brown, M. Pagowski, and K. Sušelj, 2019: A description of the MYNN-EDMF scheme and the coupling to other components in WRF–ARW. NOAA Tech. Memo. OAR GSD-61, NOAA, 42 pp., https://doi.org/10.25923/n9wm-be49.

Poll, S., P. Shrestha, and C. Simmer, 2017: Modelling convectively induced secondary circulations in the terra incognita with TerrSysMP. *Quart. J. Roy. Meteor. Soc.*, **143**, 2352–2361, https://doi.org/10.1002/qj.3088.

Rai, R. K., L. K. Berg, B. Kosović, J. D. Mirocha, M. S. Pekour, and W. J. Shaw, 2017: Comparison of measured and numerically simulated turbulence statistics in a convective boundary layer over complex terrain. *Bound.-Layer Meteor.*, **163**, 69–89, https://doi.org/10.1007/s10546-016-0217-y.

——, ——, ——, S. E. Haupt, J. D. Mirocha, B. L. Ennis, and C. Draxl, 2019: Evaluation of the impact of horizontal grid spacing in terra incognita on coupled mesoscale–microscale simulations using the WRF framework. *Mon. Wea. Rev.*, **147**, 1007–1027, https://doi.org/10.1175/MWR-D-18-0282.1.

Sharma, A., H. J. S. Fernando, A. F. Hamlet, J. J. Hellmann, M. Barlage, and F. Chen, 2017: Urban meteorological modeling using WRF: A sensitivity study. *Int. J. Climatol.*, **37**, 1885–1900, https://doi.org/10.1002/joc.4819.

Shin, H. H., and S.-Y. Hong, 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 J. Dudhia, 2016: Evaluation of PBL parameterizations in WRF at subkilometer grid spacings: Turbulence statistics in the dry convective boundary layer. *Mon. Wea. Rev.*, **144**, 1161–1177, https://doi.org/10.1175/MWR-D-15-0208.1.

Simon, J. S., B. Zhou, J. D. Mirocha, and F. K. Chow, 2019: Explicit filtering and reconstruction to reduce grid dependence in convective boundary layer simulations using WRF-LES. *Mon. Wea. Rev.*, **147**, 1805–1821, https://doi.org/10.1175/MWR-D-18-0205.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

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.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Kluwer Academic, 666 pp.

Sullivan, P. P., and E. G. Patton, 2011: The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulation. *J. Atmos. Sci.*, **68**, 2395–2415, https://doi.org/10.1175/JAS-D-10-05010.1.

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.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC

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.

Williams, A. G., and J. M. Hacker, 1992: The composite shape and structure of coherent eddies in the convective boundary layer. *Bound.-Layer Meteor.*, **61**, 213–245, https://doi.org/10.1007/BF02042933.

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., Y. Wang, and M. Wang, 2018: The performance of a scale-aware nonlocal PBL scheme for the subkilometer simulation of a deep CBL over the Taklimakan desert. *Adv. Meteor.*, **2018**, 8759594, https://doi.org/10.1155/2018/8759594.

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.

Brought to you by KAUST | Unauthenticated | Downloaded 05/15/22 10:04 AM UTC