CrossMark

# The Onset of Resolved Boundary-Layer Turbulence at Grey-Zone Resolutions

**John C. Kealy[1]** · **Georgios A. Efstathiou[1]** · **Robert J. Beare[1]**

## Abstract

Numerical weather prediction (NWP) models are now capable of operating at horizontal resolutions in the 100-m to 1-km range, a grid spacing similar in scale to that of the turbulent eddies present in the atmospheric convective boundary layer (CBL). Known as the 'grey zone' of turbulence, this regime is characterized by significant contributions from both the resolved and subgrid components to represent the dominant motions of the system. This study examines how the initiation of resolved turbulence – a concept commonly referred to as 'spin-up'– can be delayed during the evolution of a simulated CBL in the grey zone. We identify the importance of imposed pseudo-random perturbations of potential temperature ($\theta$) for the development of the resolved fields showing that without such perturbations, resolved turbulence does not become established at all. When the perturbations are organized, spin-up can develop more rapidly, and we find that the earliest spin-up times can be achieved by applying an idealized profile of variance to derive the $\theta$ perturbation values. The perturbation structures are shown to be most effective when applied at intervals following the mixed-layer time scale, $t_*$, rather than perturbing only at the initial time. We also propose a modification to the three-dimensional Smagorinsky turbulence closure, in which the Smagorinsky constant is replaced by a scale-dependent coefficient. Both the approaches of: (1) applying structured $\theta$ perturbations, and (2) using a dynamically-evolving Smagorinsky coefficient are shown to encourage faster spin-up independently of each other, but the best results clearly emerge when the two methods are applied concurrently.

**Keywords** Convective boundary layer · Grey zone · Morning transition · Spin-up · Turbulence parametrization

## 1 Introduction

Since its inception, numerical weather prediction (NWP) has always been governed by a tight balance; on the one hand using the most up-to-date science in representing atmospheric processes, and on the other, the best use of available computational resources. Ever since the earliest days of operational NWP in the 1950s, parametrization has been key in

---

✉ John C. Kealy
j.kealy@exeter.ac.uk

[1] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

 Springer

maintaining this balance, such that subgrid-scale physical processes could be represented and their influence be passed back to the grid. Increasing model resolution through the years has allowed for more advanced dynamics (e.g. non-hydrostatic modelling) and finer detail (e.g. mesoscale systems and orography), with a marked increase in forecast accuracy (Carpenter 1979; Simmons et al. 1989; Mass et al. 2002; Lean et al. 2008). However, with many NWP centres now running limited-area models (LAMs) at grid spacings on the kilometric or even sub-kilometric scales, the time has come to once again ask the question – is higher resolution always better?

The grey zone of turbulence (or terra incognita, Wyngaard 2004) refers to resolutions at which turbulent eddies in the atmospheric boundary layer are partially resolved and partially parametrized, a regime that is now emerging in the highest resolution mesoscale models. Similarity functions derived by Honnert et al. (2011) and a length scale analysis from Beare (2014) imply that the limits of this regime can vary for different model fields and configurations, with a strong sensitivity to the scale of the features being modelled. Eddies that are represented within the grey zone often appear to evolve in a grid-dependent way (Ching et al. 2014; Zhou et al. 2014), and since neither one-dimensional (1D) planetary boundary-layer (PBL) schemes (e.g. Lock et al. 2000) nor three-dimensional (3D) large-eddy simulation (LES, Lilly 1967) formulations are designed to work within this regime, a necessity is growing to find new ways to represent these processes.

It is not self-evident that resolved turbulent motions should be permitted within the grey zone (Zhou et al. 2014). Ching et al. (2014) investigate possible methods for suppressing this resolved motion, so that the model physics relies solely on the 1D PBL scheme throughout. This would negate some of the issues arising from representing each turbulent eddy with as little as one or two grid points. However, as the authors note, a drawback of this approach is that the assumptions used in Reynolds-averaged Navier–Stokes (RANS) modelling may no longer be valid at such small grid spacings. Additionally, there is evidence to suggest that resolving turbulent motions in the PBL might have succeeding effects on how larger-scale features develop, such as deep convection (Petch et al. 2002; Efstathiou et al. 2016).

A key consideration in grey-zone modelling is the scale of the atmospheric processes with respect to the grid spacing (along with the effective resolution needed to represent them). Many atmospheric processes may change in scale as a consequence of diurnal variation, large-scale forcing, thermal advection, etc., whereas the grid size tends to remain static. This implies that processes beginning their life cycle within a PBL scheme can enter the grey zone as they develop.

One such feature is the convective boundary layer (CBL) over land. The CBL has been the focus of many recent grey-zone studies (Beare 2014; Zhou et al. 2014; Ito et al. 2015; Shin and Hong 2015; Efstathiou et al. 2016; Honnert et al. 2016), and is of particular interest because the scale of the boundary-layer depth ($z_i$) changes during the transition from the shallow, stable, and fully-parametrized boundary layer ($O[10^1$ m$]$) of the night-time (Beare et al. 2006a, b) to the deeper ($O[10^3$ m$]$), thermally-driven CBL in the afternoon. Efstathiou et al. (2016) investigated the morning transition of the CBL for the Wangara experiment in Australia (Clarke et al. 1971) and also for a case from the Met Office site in Cardington, UK. Using horizontal resolutions in the range 50–800 m, they showed that mean profiles of potential temperature ($\theta$) had a tendency to become overly superadiabatic, with vertical mixing being increasingly delayed as resolution decreased. The time series of horizontally-averaged turbulence kinetic energy (TKE, Sect. 2.2) for these runs also showed an associated delay in the transition into resolved motion, with the coarser resolution runs being the last to exhibit resolved convection. (The onset of resolved convection in this context is often referred to as 'spin-up'.)

Spin-up has also been linked to Rayleigh–Bénard thermal instability theory (Zhou et al. 2014). Simulations within the grey zone appear to need larger values of the Rayleigh number ($Ra$) as grid spacing coarsens before spin-up occurs, but there does exist a critical Rayleigh number ($Ra_c$) that is theoretically achievable in the grey zone, marking the onset of resolved turbulence. It follows from this that the $\theta$ profiles might become more superadiabatic before $Ra_c$ is reached, as with the cases in Efstathiou et al. (2016). However, the question still remains: what exactly governs the timing of spin-up, and can the transition be accelerated?

In this paper, we investigate the factors that influence the spin-up of CBL turbulence within the grey zone. The primary aim is to categorize and test model sensitivities, but we also offer some practical techniques that, as will be shown, have a marked influence on how early in the simulation the resolved turbulence can be encouraged to appear. Introducing pseudo-random perturbations into the $\theta$ field is the key to this approach, and the impacts of arranging such perturbations in both time and space are investigated. In LES modelling, the choices of perturbation structure can affect certain fields, such as the initiation of deep convection (Stirling and Petch 2004). We investigate here how such perturbations affect the grey-zone boundary layer, and what implications this has. An important aspect of this approach is that the modifications are applied only to the resolved fields, allowing use of the method without forcing constraints on the choice of subgrid scheme.

We also make our own adjustments to the parametrization itself, in this case the 3D Smagorinsky scheme, highlighting the need to modify standard LES or PBL parametrizations when attempting to create an optimal grey-zone model configuration. The methods employed in designing the $\theta$ perturbations and the subgrid-scheme modifications are described in Sect. 2. In Sect. 3, we investigate several approaches for applying the perturbations, and also the sensitivity of our modified parametrization. We make the argument that the most promising avenue towards finding the optimal method of inducing grey-zone spin-up, while preserving or even improving the overall evolution of the simulated CBL, is to customize both the resolved fields and the subgrid scheme in combination. We discuss our findings in Sect. 4.

## 2 Methods

### 2.1 Model Set-Up

We make use of the UK Met Office's large-eddy simulation model, known as the **M**et-**O**ffice **N**ERC **C**loud-resolving model (MONC, Brown et al. 2015). The MONC model employs an identical scientific formulation to its predecessor, the Large-eddy model (Gray et al. 2001), with improved parallel capability. We have here configured the MONC model to simulate day 33 of the Wangara experiment, and since this is a widely studied CBL case (Yamada and Mellor 1975; André et al. 1978; Nakanishi et al. 2014), we do not repeat the evaluation of model performance at LES resolutions against observations taken during the campaign.

As a benchmark simulation, we use a LES run with a horizontal grid spacing of $\Delta x = \Delta y = 50$ m and a vertical grid spacing of $\Delta z = 20$ m across a $9.6 \times 9.6$ km$^2$ domain. The simulation employs a Boussinesq approximation with the advection of momentum following the scheme of Piacsek and Williams (1970), and advection of heat by the total variation diminishing scheme of Leonard et al. (1993). The lower boundary is forced using time-varying sensible-heat-flux data from the campaign, with periodic flow in the horizontal, and an upper boundary set at a height of $z = 2500$ m. A relaxation term is applied above 2000 m to inhibit gravity-wave formation. All simulations begin at 0900 local time (LT) and

run until 1800 LT, with output taken at intervals of approximately 100 s. The initial profile of $\theta$ is prescribed, and indicates the presence of an inversion at $z_i = 100$ m, which rises throughout the simulation. Horizontal wind-speed profiles have also been prescribed as per the observations, but wind speeds are generally low ($< 5$ m s$^{-1}$), and so we concentrate here on thermally-driven circulations.

## 2.2 Grey-Zone Simulations

For the grey-zone runs, we use horizontal resolutions ranging from $\Delta x = 200$ m through to $\Delta x = 800$ m, with at least $24 \times 24$ grid points in the horizontal and a grid spacing of approximately 40 m in the vertical (63 levels). The vertical levels remain fixed, with slightly higher vertical resolution near the surface boundary, in common with many NWP model configurations.

Our grey-zone simulations make use of the 3D static Smagorinsky scheme (see Sect. 2.4). Countergradient correction terms are sometimes added in the grey zone to provide extra non-local transport (e.g. Boutle et al. 2014), but we do not make use of such terms here because we wish to measure the effect our methods have on the non-local transport provided to the system by the resolved fields. Another key limitation when using a 3D turbulence scheme in the grey zone is the absence of a cascade of turbulent energy through the inertial subrange. These limitations should always be kept in mind when studying grey-zone modelling.

Honnert et al. (2011) describe a method of 'coarse-graining' the LES model solution to approximate the desired grey-zone variances, leading to a useful benchmark in each TKE time series to compare with the grey-zone runs. The horizontally-averaged TKE, denoted here as $\langle e \rangle$, is defined by:

$$\langle e \rangle = \frac{1}{2} \left( \overline{u'^2} + \overline{v'^2} + \overline{w'^2} \right), \tag{1}$$

where $\left( u', v', w' \right)$ are deviations from the mean state of the velocity components. The coarse-graining approach is based on taking a horizontal mean across bi-dimensional cells of several grid points in width, such that the variances in the LES model at 50-m resolution are reduced to match the desired grey-zone resolution. Although this approach is not without flaw, it is shown by Honnert et al. (2011) to be very useful in estimating the expected magnitude of resolved TKE in grey-zone runs. We therefore allow the coarse-grained fields to serve as an idealized result for such comparisons. Here we have created four coarse-grained time series of $\langle e \rangle$ from the 50-m LES model: $\Delta x = 200$ m, $\Delta x = 400$ m, $\Delta x = 600$ m and $\Delta x = 800$ m.

## 2.3 Structuring Perturbations in the Initial Fields

The traditional method of encouraging resolved turbulence to spin-up in a LES model is to impose artificial pseudo-random perturbations to the initial potential temperature and/or momentum fields. This is generally done at the lowest model levels, though some disparity exists in the literature as to the choices of perturbation. The consequences of selecting certain perturbation amplitudes and structures begin to have more relevance in the grey zone, and we will show that the effects of choosing such perturbations carefully have significant impacts on the resolved fields, at minimal computational expense. While many studies focus on the search for a parametrization scheme appropriate for the grey zone, here we investigate whether or not sensitivities exist in the resolved fields that may be useful in combination with an appropriate parametrization.

Some authors (Shin and Hong 2015; Shin and Dudhia 2016) choose the initially-imposed $\theta$ perturbations ($\theta'$) from the range $-0.05\,\mathrm{K} < \theta' < 0.05\,\mathrm{K}$ in the lowest four model levels. Others (Petch et al. 2002; Beare 2008) choose the range $-0.1\,\mathrm{K} < \theta' < 0.1\,\mathrm{K}$ in the lowest 100–200 m of the domain. The largest range we find in the literature, $-0.5\,\mathrm{K} < \theta' < 0.5\,\mathrm{K}$, has been applied up to various heights above ground level, including 50 m (Nakanishi et al. 2014), 200 m (Stirling and Petch 2004), 500 m (Mirocha et al. 2014), the lowest model level only (Mason and Brown 1999; Brown et al. 2000), and $2z_i/3$ (Muñoz-Esparza et al. 2014). So far, these authors have not explicitly explained their choice of perturbation structure, primarily because it is generally accepted that this choice does not seem to have much significance in the LES regime. These numbers apply to the CBL for the most part, though the perturbations are typically included for simulations of the stable boundary layer (SBL) as well (Beare et al. 2006b).

Muñoz-Esparza et al. (2014) have developed methods to impose these pseudo-random $\theta'$ values in a structured way. Their best performing method, which they refer to as 'cell perturbation', was shown to decrease the necessary fetches from the edges of the domain for spin-up to occur when nesting high-resolution domains that pass through the grey zone to reach the inner nests. We have applied a similar principle to this approach by adding organization in the horizontal by means of cells, in which each cell contains a unique perturbation common to all grid points that lie within that cell. In testing the method, we focus on the time series of $\langle e \rangle$ rather than the spatial approach taken by Muñoz-Esparza et al. (2014), and stress that our method differs from that of Muñoz-Esparza et al. (2014) in many respects.

We also test three other hypotheses regarding the structuring of the initial perturbations, to show their effect (if any) on spin-up in the grey zone. The four methods are:

(1) Increasing the amplitude of the initial perturbation of potential temperature (Fig. 1a).
(2) Our variation of the 'cell perturbation' method (hereafter denoted as CELL) using cells of $4 \times 4$ (CELL-4) and $8 \times 8$ (CELL-8) grid points (Fig. 1b).
(3) Applying coherence to the vertical dimension, such that each grid-point perturbation is unique in the horizontal, but identical in the vertical (Fig. 1c).
(4) Using mixed-layer scaling theory to structure the perturbation range (Fig. 1d).

Method 4 above is the most complex and requires further calculation, but we show below that this method may well be worth the extra computation. According to mixed-layer scaling theory, the values of $\theta'$ in the CBL are thought to be largest at the surface and at the inversion, and Garratt (1994) explains how this can be quantified using the empirical relationship

$$\sigma_\theta^2 = T_*^2 \left[ 2 \left( \frac{z}{z_i} \right)^{-2/3} \left( 1 - \frac{z}{z_i} \right)^{4/3} + 0.94 \left( \frac{z}{z_i} \right)^{4/3} \left( 1 - \frac{z}{z_i} \right)^{-2/3} \right] \tag{2}$$

for $z < z_i$, where $\sigma_\theta^2$ is the naturally occurring variance of $\theta'$ as a function of $z/z_i$. Here, $T_*$ is defined by

$$T_* = \left[ \frac{\overline{\theta} \left( \overline{w'\theta'} \right)_0^2}{g z_i} \right]^{1/3}, \tag{3}$$

where $\overline{\theta}$ is a reference temperature, $\left( \overline{w'\theta'} \right)_0$ is the kinematic heat flux at the surface, and $g$ is the acceleration due to gravity. Knowledge of $\sigma^2$ in a uniform distribution allows for the calculation of its upper and lower limits $[a, b]$, since $\sigma^2 = (b - a)^2/12$, and therefore if we assume that $\overline{\theta'} = 0$, we can calculate the limits $[a, b]$ from which to source the random numbers at each vertical model level,
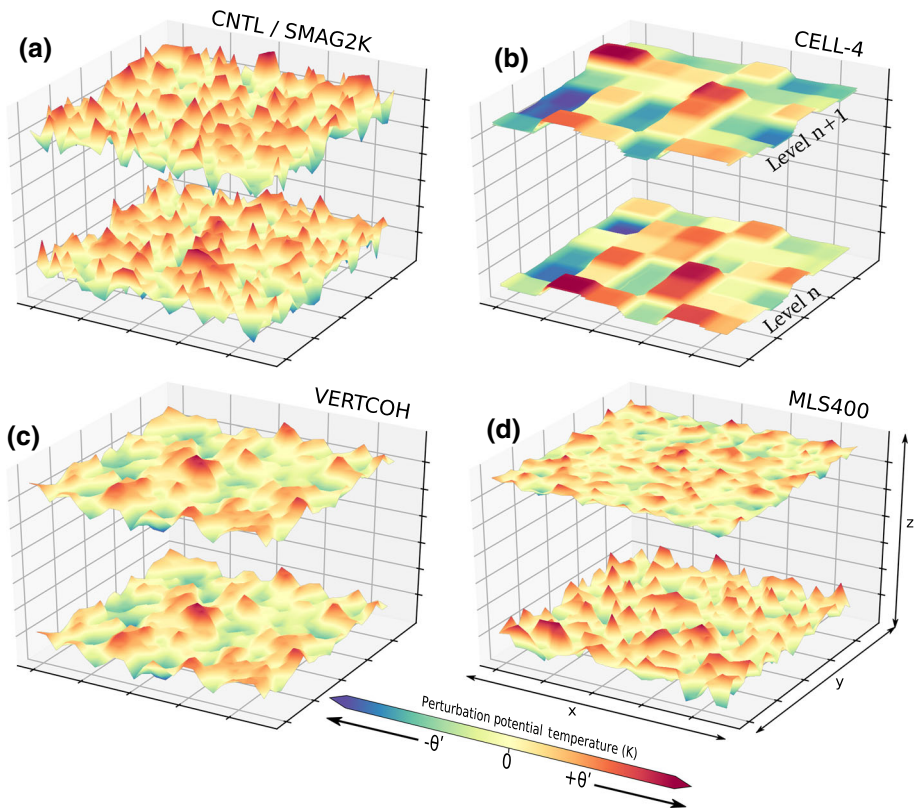
**Fig. 1** 3D visualization of the various methods of perturbing $\theta$, representing adjacent model levels (at some arbitrary level $n$) in the lower boundary layer at $t + 0$. The methods are: (**a**) the default set-up (the SMAG2K and CNTL simulations differ only in the amplitude of $\theta'$); (**b**) the CELL method with cell sizes of $4 \times 4$ grid boxes; (**c**) the vertical coherence method; (**d**) the mixed-layer scaling method

$$a = -\sqrt{3\sigma_\theta^2} \qquad (4a)$$

$$b = \sqrt{3\sigma_\theta^2}. \qquad (4b)$$

We have elected to assume a uniform distribution for $\theta'$, rather than a Gaussian distribution, based on tests performed by Muñoz-Esparza et al. (2014). Preliminary runs have been performed to test how the system responds to perturbations in the vertical-velocity field ($w$), rather than $\theta$, and these have shown comparable results. We have therefore chosen to perturb $\theta$ only, given the dependence of $w$ on $\theta$ in the basic equation set.

We have also investigated the use of applying the perturbations in time. At pre-defined intervals, new random numbers are chosen and the perturbations are reapplied. The mixed-layer time scale ($t_*$) is our favoured choice of interval in which to inject the perturbations, because it is both physically tenable and responds well to testing. Here, $t_*$ represents the time scale for the life cycle of a single CBL eddy and follows the relation

$$t_* = \frac{z_i}{w_*}, \qquad (5)$$

where $w_*$ is the convective velocity scale, defined by

$$w_* = \left[ \frac{g z_i}{\overline{\theta}} \left( \overline{w'\theta'} \right)_0 \right]^{1/3}. \tag{6}$$

Hence $t_*$ increases as $z_i$ increases, and the perturbations are implied less frequently as the CBL develops; here, $t_*$ is employed as a domain-averaged value.

## 2.4 Diffusion in the Subgrid Scheme

The subgrid scheme in use with the MONC model is the 3D Smagorinsky scheme (Smagorinsky 1963; Brown et al. 1994), a parametrization that fundamentally relies on a constant to control the level of diffusion of the eddy energy. This Smagorinsky constant, usually denoted $C_S$, is used to calculate the basic mixing length ($\lambda$), such that

$$\lambda = C_S \max(\Delta x, \ \Delta y). \tag{7}$$

The basic mixing length undergoes a modification to accommodate near surface effects, which produces the neutral mixing length ($l$),

$$\frac{1}{l^2} = \frac{1}{(kz)^2} + \frac{1}{\lambda^2}, \tag{8}$$

where $k$ is the von Kármán constant. Eddy viscosity ($\nu$) and diffusivity ($\nu_h$) are then calculated using

$$\nu = l^2 \left| S_{ij} \right| f_M(Ri) \tag{9a}$$

$$\nu_h = l^2 \left| S_{ij} \right| f_H(Ri), \tag{9b}$$

where $f_M$ and $f_H$ are stability functions for momentum and heat, which depend on the Richardson number ($Ri$), and $\left| S_{ij} \right|$ is the modulus of the strain tensor defined by

$$S_{ij} = \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}. \tag{10}$$

Finally, the subgrid-stress and buoyancy-flux terms in the full transport equations are calculated using

$$\overline{u'_i u'_j} = -\nu S_{ij} \tag{11a}$$

$$\overline{u'_j \theta'} = -\nu_h \frac{\partial \overline{\theta}}{\partial x_j}, \tag{11b}$$

where the subscripts $[i, \ j]$ imply tensor notation.

Lilly (1967) originally defined the value of the Smagorinsky constant to be $C_S \approx 0.17$ (assuming a value of $\alpha = 1.5$ for the Kolmogorov constant). However, this value is often adjusted for use in atmospheric models, typically requiring some tuning to best balance the signal-to-noise ratio after considering dissipation effects from the numerical discretization scheme and other model components. Both the MONC model and the Met Office Unified Model (MetUM) use a value of $C_S = 0.23$ by default, while the LES configuration of the Weather Research and Forecast (WRF) model has a default of $C_S = 0.18$. Beare (2014) and Efstathiou and Beare (2015) investigate some of the effects of changing $C_S$ in grey-zone simulations, one of which was the dampening of resolved turbulence when $C_S$ is increased. Lower values of $C_S$ tended to allow for faster spin-up, but introduced noise in the resolved

fields that inhibited the development of coherent structures in the fully-developed CBL, as well as allowing for excessive resolved eddy energy.

The value of $C_S$ does not necessarily need to be constant throughout a simulation. Investigations into scale-dependence in LES modelling have lead to the development of dynamic formulations (Germano et al. 1991; Porté-Agel et al. 2000; Meyers and Sagaut 2006; Efstathiou et al. 2018). These dynamic models are based on the concept that, as a coefficient, $C_S$ should exhibit a dependency on the quantity $L/\Delta$, where $L$ is the dominant length scale and $\Delta$ is the filter width. For the case of the atmospheric CBL, this quantity is analogous to $z_i/\Delta x$, which also happens to be the reciprocal of the grey-zone similarity variable proposed by Honnert et al. (2011). Wyngaard (2004) also explains the importance of the variable $L/\Delta$, stressing how fundamental this ratio is to grey-zone modelling. Also related is the similarity variable $z_i/l_d$ from Beare (2014), where $l_d$ is the dissipation length scale. Recent work by Efstathiou et al. (2018) has shown the utility of dynamic models in simulating the CBL, although this approach can be somewhat demanding computationally.

With these considerations in mind, we propose a coefficient $C_S$ that adapts to the evolving CBL from the fully-parametrized state into that of partially-resolved turbulence during the morning transition. Our $C_S$ value is constant across the domain but evolves in time alongside $z_i$ such that

$$C_S = C_{S0} \tanh\left(\beta \frac{z_i}{\Delta x}\right), \qquad (12)$$

where $C_{S0}$ is an equilibrium value in the interior of the fully-developed CBL, and $\beta$ is some constant in the unit interval. The ideal value of $\beta$ is likely to be model dependent (i.e. requires some tuning), again due to varying dissipation sources from other model components. We justify the hyperbolic tangent form of Eq. 12 by comparing it with $C_S$ data taken from the Lagrangian-averaged scale-dependent dynamic (LASD) Smagorinsky model of Efstathiou et al. (2018) during the morning CBL development. Figure 2 shows how the vertical maximum of horizontally-averaged $C_S$ data from the LASD model approximately traces a tanh shape, and we find that a best fit of these data using the functional form of Eq. 12 employs the parameters $\beta = 0.29$ and $C_{S0} = 0.19$. Since we are seeking a single, optimal value of $C_S$ for the entire domain, the influence of the upper and lower boundaries must be minimized. This is the reason why the form of Eq. 12 has been based on the maximum $C_S$ values from the LASD model. The figure also shows the form of Eq. 12 with $\beta$ set to 1.0 and $C_{S0} = 0.23$, which we have found in our simulations to be a better choice when combining the method with those of Sect. 2.3. Also shown is the curve following $\beta = 0.6$, which we discuss in Sect. 3.3. In other models (e.g., the WRF model), we suggest that also using the default value of $C_S$ for setting $C_{S0}$ is probably the best starting point. The scope of our study includes the introduction of Eq. 12 as a proof of concept, but we acknowledge that more work needs to be done before claiming that the method is fully robust in its present form.

# 3 Results

## 3.1 Initial-Condition Perturbations

During the simulated development of a nocturnal boundary layer, the variances in the resolved momentum fields inevitably tend to zero. This is because negative thermodynamic fluxes at the lower boundary no longer promote convective activity, and the model tends to a more laminar flow overnight (in the absence of mechanically-driven turbulence). The following morning, when positive surface heat fluxes return, the resolved fields have little small-scale
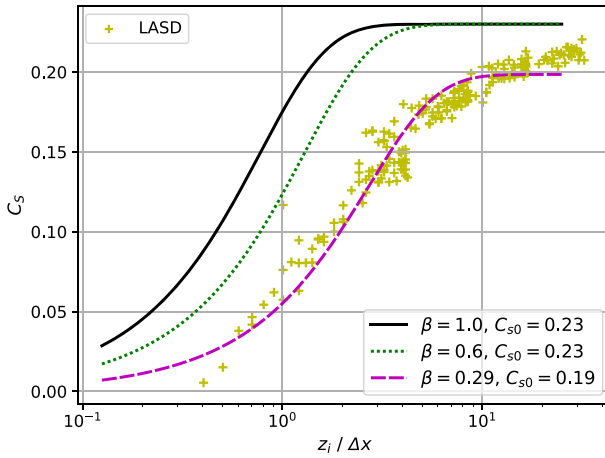
**Fig. 2** Prognostic values for $C_S$ as described by Eq. 12. Also shown are $C_S$ data taken from the LASD model (Efstathiou et al. 2018) as a basis for the tanh shape of the equation
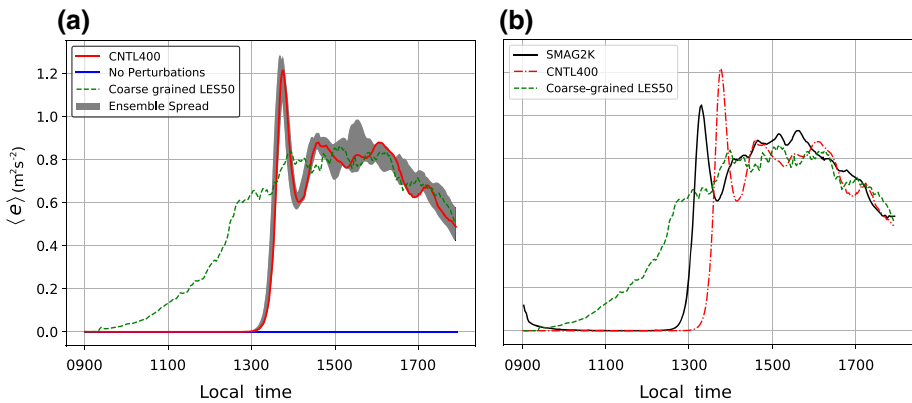


**Fig. 3** Time series of $\langle e \rangle$ at $\Delta x = 400$ m for (**a**) the CNTL400 run (solid red) with the spread of possibilities across a further 12 runs using altered seeds (grey shading). Also shown are the the coarse-grained fields (dashed green) and the outcome of not using any initial perturbations (solid blue); (**b**) SMAG2K run (solid black) showing the effect of imposing perturbations in the larger range of $-2\,\mathrm{K} < \theta' < 2\,\mathrm{K}$ at $t+0$. Note how the initially high energy value of the SMAG2K run is dissipated away almost immediately

variation. This lack of variation is not a problem in mesoscale models since all subsequent CBL turbulent transfer is represented by the parametrization scheme, as in the night-time SBL. In the grey zone, however, a lack of this variation produces difficulties in the transition from subgrid to partially-resolved flow. Figure 3 shows how this affects the evolution of the Wangara CBL. We will focus on a horizontal grid spacing of 400 m throughout this section, since this resolution lies firmly within the grey zone, and exhibits a significant resistance to spin-up. In the figure, $\langle e \rangle$ is calculated at the closest model level to $z/z_i = 0.5$.

The difference between omitting and including random perturbations at time $t+0$ is immediately apparent, and it is clear from Fig. 3a that no resolved motion develops during the 9-h period in the unperturbed case. The perturbed case in Fig. 3a is based on a commonly-used LES set-up at the UK Met Office, in which the pseudo-random perturbations are drawn

from the range $-0.1\,\mathrm{K} < \theta' < 0.1\,\mathrm{K}$ up to a height of 250 m above ground level. We use this as the point of reference (a control, hereafter CNTL) for subsequent configurations.

Pseudo-random numbers are typically generated using a seeding function, which for the MONC model is loosely based on the grid-point location and the processor allocation. We have run 12 additional unique simulations using the CNTL400 configuration, altering the random number seed in each, to examine when and by how much each run diverges when the random perturbations are different (grey shading in Fig. 3a). This technique has a similar principle to ensemble modelling. The spin-up times for each of these members are within 5 min of one another, implying that the random numbers themselves have little effect on the actual timing of convective onset. The runs then diverge after resolved convection is initiated, showing a spread of approximately $0.2\,\mathrm{m}^2\,\mathrm{s}^{-2}$ in the TKE amplitude. Each run does, however, follow the general pattern of the coarse-grained fields once a steady state is reached.

A strong peak in $\langle e \rangle$ is apparent in the CNTL400 simulation, following the abrupt onset of resolved convection (at 1330 LT in Fig. 3a). As the surface heat flux increases at the lower boundary, the lack of resolved motion causes a build-up of energy preceding convective onset, which corresponds with the achievement of the critical Rayleigh number (Zhou et al. 2014). At the point of spin-up, the TKE values tend to overshoot so that this energy can be released, sometimes followed by a slight oscillation before settling into the same pattern as the coarse-grained fields. This effect has probably been present in previous studies, but is only apparent here because we use a very high temporal resolution in our output ($\approx 100\,\mathrm{s}$). For CNTL400, the model takes more than half of the simulation time before the resolved turbulence can reach a steady state. It is therefore desirable to make modifications to induce spin-up at an earlier time, allowing the energy entering the system to be properly transported and diffused.

### 3.2 Inducing Spin-Up

An important practical consideration is to determine whether or not spin-up can be accelerated, thereby encouraging non-local transport and relaxing the excessively superadiabatic profiles of the CNTL run during the mid- to late-morning. In the previous section, we discussed the effects of varying the seed of the initial pseudo-random numbers, noting a negligible change in spin-up time. We now apply the methods outlined in Sect. 2.3, and analyze their effects on the timing of spin-up. For the purpose of comparison, we shall here define 'spin-up' as the point in time at which the magnitude of $\langle e \rangle$ reaches $0.1\,\mathrm{m}^2\,\mathrm{s}^{-2}$. This is an arbitrary choice, but nonetheless useful for the comparison. A similar definition is used by Zhou et al. (2014), who define spin-up using $\langle u_i' u_i' \rangle / w_*^2 = 0.1$.

Many of our early tests with the methods of vertical coherence, CELL, mixed-layer scaling, and customization of perturbation altitude/amplitude yielded quite similar results. Although initially organized, the variations in the resolved fields would tend to dissipate away, returning the $\langle e \rangle$ field to near-zero despite the positive gradient in surface heat flux. This leads to our first key finding: perturbations imposed at the initial time have a tendency to be damped away by the subgrid scheme in the grey zone, in contrast with the LES model behaviour. We can visualize this by showing how even an unrealistically large perturbation range of $-2\,\mathrm{K} < \theta' < 2\,\mathrm{K}$ behaves in the grey zone (SMAG2K, Fig. 3b). The resolved TKE in the SMAG2K run is initially very large, but this energy is quickly removed by the subgrid scheme within the first hour of the simulation. It follows from this that specific structures such as those

of the CELL method lose their organization early on if the perturbations are only applied at the $t + 0$ timestep.

Although much of the organization is lost, the modification of the initial state of $\theta$ does affect the resolved fields to a certain extent. We have found differences in spin-up time of up to 40 min in our preliminary tests, implying that despite the $\langle e \rangle$ field tending to near-zero ($O[10^{-4}\,\text{m}^2\,\text{s}^{-2}]$ in Fig. 3b), there still exists a 'memory' between $t + 0$ and the time of spin-up. This is evident in Fig. 3b, in which the modified SMAG2K run exhibits a spin-up that is $\approx 30$ min earlier than the CNTL400 run, despite a negligible amount of resolved energy being present just before spin-up.

Because of this, we find that accelerating the transition to resolved turbulence in the grey zone requires the implementation of the pseudo-random perturbation structures at intervals of time. The perturbations are applied every $t_*$ and are then ceased entirely at $t = 15{,}000\,\text{s}$, when the CBL is fully developed. A preliminary run using the same settings as the CNTL400 run (but applied every $t_*$) produced an acceleration in spin-up of 44 min, and following on from this, we sought to improve the result by implementing the methods from Sect. 2.3. Table 1 shows how each of these methods affects the timing of spin-up. Spin-up times shown in the table reflect how much earlier the resolved motions appear with respect to the CNTL400 simulation.

The CELL method (Fig. 1b) was applied for bi-dimensional cells of $4 \times 4$ (CELL-4) and $8 \times 8$ (CELL-8) grid points, with an acceleration in spin-up of 37 min and 24 min respectively. Overall, we find that perturbing $\theta$ every $t_*$ shows no improvement over imposing a unique $\theta'$ value to each grid point every $t_*$ (CELL-1). It is possible that numerical considerations are working against the physical basis of the method here, since the technique was originally designed for grid spacings of no more than $\Delta x = 100\,\text{m}$. The horizontal scale of the perturbations in the grey zone may simply be too large when multiple grid points are assigned to each cell, but using multiple grid points is generally necessary for satisfying the effective-resolution requirements of the model. On balance, it seems that in the grey zone, perturbing by cells is not effective; however, we have not been able to test every combination of cell size, domain size, and resolution, and therefore we cannot make this claim explicitly. We stress that this result does not call into question the underlying viability of Muñoz-Esparza's approach, but for now, it does appear that perturbing $\theta$ by cells is not a defensible approach at grey-zone resolutions.

The vertical coherence method (hereafter denoted VERTCOH, Fig. 1c) is based on the idea of creating disturbances in $\theta'$ that are complementary along each vertical column, preventing perturbations at adjacent levels from cancelling each other, and establishing organization in the vertical. This method has also been employed by Stirling and Petch (2004). When the technique is applied up to a height of 250 m (for comparison with the CNTL400 simulation), spin-up occurs 1.4 h before the CNTL400 run, which is very favourable. We see no major drawback to this method, particularly since the amplitude of perturbations is so small (0.1 K), and the method is computationally inexpensive.

The earliest to spin-up of all the methods tested was the mixed-layer scaling method (hereafter MLS400, Fig. 1d), which allowed spin-up to occur 1.8 h ahead of the CNTL400 simulation. Although this method is slightly more complex than for the VERTCOH simulation, it is still relatively inexpensive since $\sigma_\theta^2$ is calculated as a horizontal-average value at each level. MLS400 is also arguably more physical, since it follows mixed-layer theory closely, and uses $z_i$ for the depth through which the perturbations are applied (rather than up to a fixed height, as with the other methods). We seek to generalize this approach in future work, but when doing so, we must stress the importance of the limits $z \to 0$ and $z \to z_i$ in

**Table 1** Key simulations performed in this study

| | $\Delta x$ (m) | Max. $\theta'$ (K) | $\theta'$ interval | $\theta'$ depth (m) | $\beta$ | $SU_{Control} - SU$ (s) | r.m.s. error (m$^2$ s$^{-2}$) |
|---|---|---|---|---|---|---|---|
| LES50 | 50 | $\pm 0.1$ | Only $t+0$ | 250 | – | – | – |
| CNTL | 200, 400, 600, 800 | $\pm 0.1$ | Only $t+0$ | 250 | – | 0 | 0.184, 0.301, 0.198, 0.131 |
| SMAG2K | 400 | $\pm 2.0$ | Only $t+0$ | 250 | – | 1756 (29 min) | 0.233 |
| CELL-1 | 400 | $\pm 0.1$ | Every $t_*$ | 250 | – | 2644 (44 min) | 0.123 |
| CELL-4 | 400 | $\pm 0.1$ | Every $t_*$ | 250 | – | 2250 (37 min) | 0.205 |
| CELL-8 | 400 | $\pm 0.1$ | Every $t_*$ | 250 | – | 1468 (24 min) | 0.241 |
| VERTCOH | 400 | $\pm 0.1$ | Every $t_*$ | 250 | – | 4897 (1.4 h) | 0.102 |
| MLS400 | 400 | $\pm\sqrt{3\sigma_\theta^2}$ | Every $t_*$ | $z_i$ | – | 6443 (1.8 h) | 0.063 |
| CsCo400 | 400 | $\pm 0.1$ | Only $t+0$ | 250 | 0.6 | 4750 (1.3 h) | 0.115 |
| CsCoMLS | 200, 400, 600, 800 | $\pm\sqrt{3\sigma_\theta^2}$ | Every $t_*$ | $z_i$ | 1.0 | Various (see Fig. 7) | 0.075, 0.052, 0.039, 0.061 |

$\sigma_\theta^2$ is the variance in $\theta'$ as described in Sect. 2.3. Spin-up times (denoted SU and corresponding to $\langle e \rangle \approx 0.1 \, \mathrm{m}^2 \, \mathrm{s}^{-2}$) are with respect to the CNTL run. Also shown are the root-mean-square errors (r.m.s. error) between each run's TKE time series and its corresponding coarse-grained field TKE time series for the first half of the simulation

Eq. 2, since $\sigma_\theta^2 \to \infty$ at these limits. The elevations of the grid points closest to $z = 0$ and $z = z_i$ therefore become important with respect to the perturbation amplitude.

The question of the depth through which the perturbations should be applied is still an open one. We use a height of 250 m in most of our simulations because this is a commonly used value in the MONC model (e.g. Efstathiou et al. 2016), but this height is somewhat arbitrary. As we have discussed in Sect. 2.3, it is clear that a consensus does not exist as to the choice of optimal depth, and it is unclear whether this depth should remain fixed or should evolve alongside the CBL. Nakanishi et al. (2014) show that there does exist TKE in the residual layer above $z_i$ during Wangara ($O[10^{-2} \, \text{m}^2 \, \text{s}^{-2}]$). Some of our testing has revealed that adding perturbations up to 2000 m above the surface (well above the inversion) has a significant impact on the time series of $\langle e \rangle$. It is the authors' opinion, however, that such an approach does not have enough physical grounds. This is another reason why we advocate the use of the mixed-layer scaling method; because the depth below which the perturbations should be applied is clear and physically based. We believe that the grey zone is inherently a problem that will always rely to some extent on numerical considerations and tuning, and that is exactly why we wish to take a physical approach wherever possible.

### 3.3 Modification of the Smagorinsky Scheme

We now show the results of applying Eq. 12 to the Wangara grey-zone simulations. These runs maintain the perturbation structure of the CNTL400 run, while introducing the new domain-wide $C_S$ coefficient at each timestep. Overall, the runs appear to be strongly sensitive to, (i) the initial perturbation amplitude at $t + 0$, and (ii) the constant $\beta$ in Eq. 12. Figure 4a shows the model's reaction to using the values $\beta = 0.6$ and $\beta = 0.7$. Using $\beta = 0.7$ encourages spin-up to appear $\approx 40$ min ahead of the CNTL400 simulation, with a further increase of $\approx 40$ min for $\beta = 0.6$ (the $\beta = 0.6$ run is denoted by CsCo400 in Fig. 4a and Table 1). Tests at $\beta < 0.6$ (not shown) reveal an evolution that is more energetic than the coarse-grained fields, while $\beta > 0.7$ bears a strong similarity to the CNTL400 run. These values of $\beta$ are probably specific to Wangara, and are unlikely to apply directly to the general case. However, the results do show the strong sensitivity of the model to this parameter.

Sensitivity tests of our simulations show that our $C_S$ coefficient requires higher values of both $\beta$ and $C_{S0}$ than do the LASD model data shown in Fig. 2. This identifies a key difference between using a single value for $C_S$ throughout the domain, rather than a dynamically calculated $C_S$ at each grid point. We have performed a simulation using the parameters $\beta = 0.29$ and $C_{S0} = 0.19$ as in the LASD model data in Fig. 2 (not shown), but find the evolution to be noisy and overly energetic. Although it may be possible to generalize the calculation of $\beta$ and $C_{S0}$ in a future study, it is also very possible that these constants must simply be tuned to certain model configurations. A comprehensive investigation into the effects of such tuning is beyond the scope of this paper; and at this stage we offer these results as a proof of concept, rather than as a fully generalized technique in itself. However, dynamically adjusting $C_S$ using our simple tanh relationship may be preferable from a pragmatic viewpoint because of the simplicity and faster run time the method affords.

Figure 4b presents a similar result to using $\beta = 0.6$ if the initial perturbation amplitude of the $t + 0$ initial state is increased to $-0.5 \, \text{K} < \theta' < 0.5 \, \text{K}$ and $\beta$ is set to 0.7. This highlights the interplay between the state of the resolved fields and the new $C_S$ coefficient, even when $\theta$ is modified only at the first timestep. We explore such combinations further in Sect. 3.4.

Figure 5b shows how $w$ responds to the changes in $C_S$, compared to the CNTL400 run (Fig. 5a), visualized as a horizontal cross-section in the mid-CBL at $t = 3.5$ h. It is clear
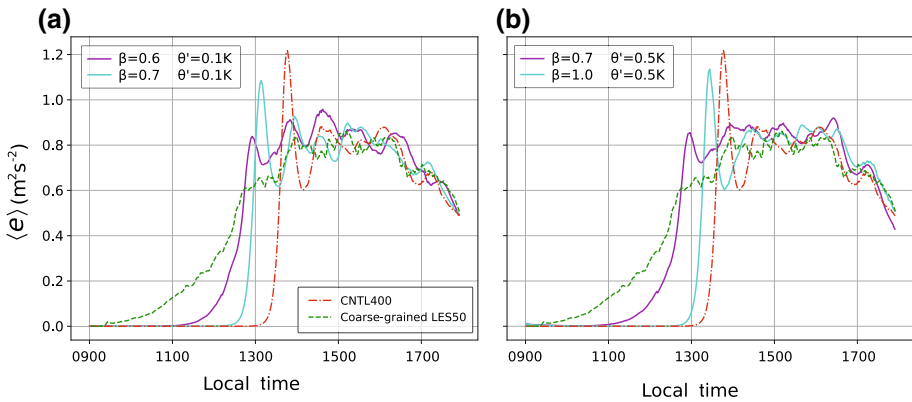
**(a)**



**(b)**



**Fig. 4** Time series of $\langle e \rangle$ evolution for the new $C_S$ coefficient, showing the effects of altering both the limits of $\theta'$ at $t + 0$ and $\beta$. The solid magenta in (**a**) is the CsCo400 run

from Fig. 5b that although spin-up has occurred, the resolved fields are somewhat noisy and lacking in organized structures (Efstathiou and Beare 2015). Although these structures are somewhat unphysical, they do provide non-local transport to the system, improving the overall evolution of the mean $\theta$ profiles (Fig. 6a). The $\theta$ profiles in CsCo400 quickly become less superadiabatic after spin-up at around 1200 LT, which implies that the resolved fields are providing improved mixing than in the CNTL400 run. By 1315 LT, the CBL is well mixed and matches well with the LES model profile. Finally, in the fully-developed CBL at 1500 LT and 1730 LT, both the CNTL400 and CsCo400 simulations are well mixed, but it appears that the earlier spin-up time of the CsCo400 run has allowed a better match with the LES model, as the CBL in CsCo400 is slightly warmer in the mid-CBL than in the CNTL400 run, and has a slightly larger value of $z_i$. Overall, this highlights two important outcomes: firstly, resolved motions appear to be a very important component of the CBL system in the grey zone – indeed it is entirely necessary when used alongside a 3D scheme such as the static Smagorinsky – and we consider its presence to be preferable over damping away the resolved fields (Ching et al. 2014) and relying on a non-local countergradient term for non-local transport. Secondly, the benefits of encouraging spin-up are not limited to the hours preceding the resolved motion, but in fact, it appears that the entire system benefits, even in the later hours just ahead of the evening transition.

### 3.4 Implications Across the Grey Zone

Since the grey zone is a regime in which both subgrid schemes and resolved motions play a significant role, one might argue that finding an optimal grey-zone configuration should logically require a customization to both. This is indeed our finding. Combining the best performing of our perturbation structures, the mixed-layer scaling method, with the use of the new $C_S$ coefficient gives the $\langle e \rangle$ field that most closely matches (r.m.s. error $= 0.052 \, \text{m}^2 \, \text{s}^{-2}$, see Table 1) the coarse-grained fields at $\Delta x = 400$ m (designated CsCoMLS). Figure 7 shows the time series of $\langle e \rangle$ using this combination across the grey zone from $\Delta x = 200$ m to $\Delta x = 800$ m. At 200 m (CNTL200), spin-up is delayed until $t = 2.6$ h, while the new method (CsCoMLS200) gives a similar spin-up timing and structure as the coarse-grained fields.
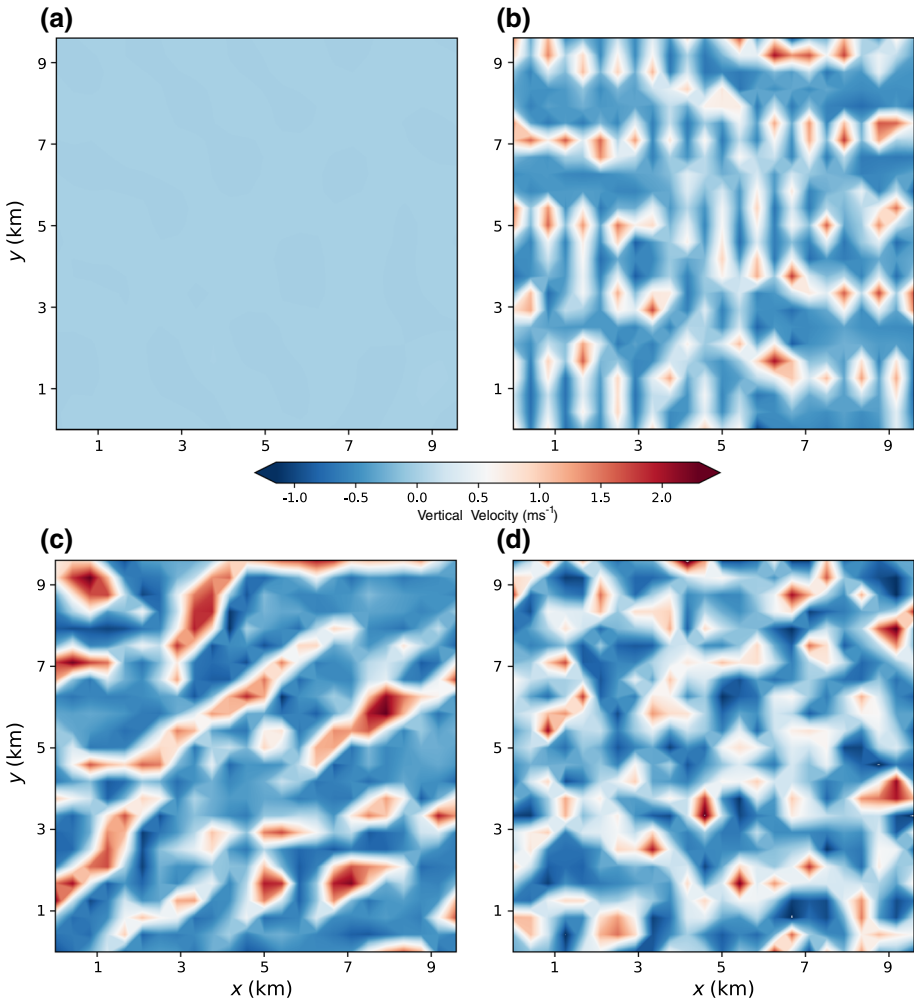
**Fig. 5** Horizontal cross-sections of $w$ at $z/z_i = 0.5$ for time $t = 3.5$ h: (**a**) CNTL400; (**b**) CsCo400; (**c**) CsCoMLS400; and (**d**) LES50 coarse-grained to 400 m

The 400 m run also follows the coarse-grained fields well, particularly at first, with an acceleration in spin-up of $\approx 2.6$ h. The large peak in $\langle e \rangle$ present in CNTL400 at $t \approx 4.5$ h is absent, although the eddies in CsCoMLS400 do become slightly over-energetic around this time. Two noteworthy features are present at this point in the time series: firstly, every simulation, including the LES model, shows a sharp increase in TKE at this time. This is coincident with the time at which the developing CBL penetrates the residual layer formed the night before. This may have implications for factors like convective inhibition, and we speculate that there may be an influence (either directly or indirectly) of this on the TKE at $z/z_i = 0.5$. Secondly, the larger vertical grid spacing of the grey-zone run (compared to the LES model) could also have implications for the behaviour of the TKE profile, since grey-zone runs have been shown to be somewhat unpredictable in these profiles, particularly near the inversion (Beare 2014). Some combination of these effects could reasonably explain
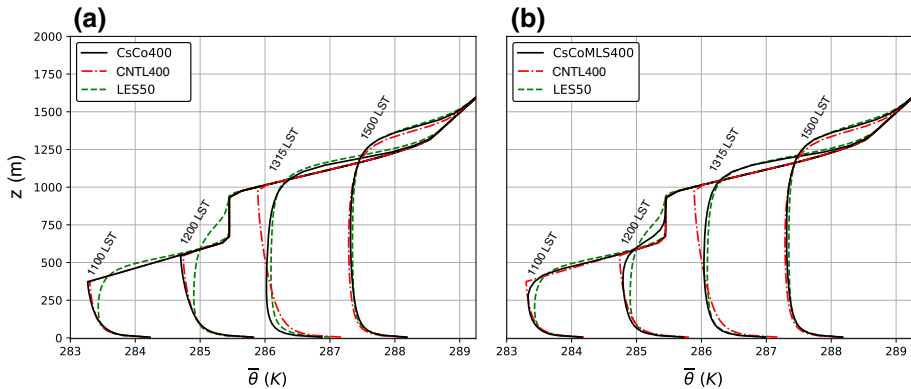
**Fig. 6** Mean vertical profiles of $\theta$: (**a**) CsCo400 before spin-up (1100 LT, 1200 LT), just after spin-up (1315 LT), and for the fully-developed CBL (1500 LT). Using the modified $C_S$ method (solid black) appears to accelerate mixing, allowing the inversion height to match the LES model profiles (dashed green) more closely than the CNTL400 run (dot-dashed red) for the fully-developed CBL. (**b**) Same as (**a**) but for CsCoMLS400, showing relaxation of the superadiabats/spin-up by 1100 LT, while remaining almost identical to CsCo400 by 1500 LT

the excessive TKE at this time in the CsCoMLS400 run. In addition to a well-matched $\langle e \rangle$ time series at $\Delta x = 400$ m, we also find that CsCoMLS400 exhibits organized structures, and their development and evolution appears physical from the beginning (Fig. 5c). The cross-section shown in Fig. 5c also compares well to the structures and amplitude of the coarse-grained fields shown in Fig. 5d. Mean $\theta$ profiles of CsCoMLS400 (Fig. 6b) are very similar to CsCo400, implying that the advantages of the modified $C_S$ coefficient are still present, but with the added advantage of earlier spin-up.

At $\Delta x = 600$ m and $\Delta x = 800$ m (the CsCoMLS600 and CsCoMLS800 runs), a very significant change occurs in the resolved TKE time series compared to the CNTL600 and CNTL800 simulations. From approximately 1330 LT onward, the $\langle e \rangle$ values match well with the coarse-grained fields, exhibiting spin-up well in advance of the CNTL runs. In contrast, spin-up in the CNTL runs is delayed for the majority of the simulations, producing an insufficient amount of non-local transport.

We do note that the CsCoMLS600 and CsCoMLS800 runs are overly energetic in the first 2–3 hours. This would appear to be a by-product of the forcing we have applied, and indeed some of this TKE may well be a direct result of the imposed perturbations themselves, rather than naturally occurring turbulence. This begs the question of whether the amplitude of the $\theta'$ perturbations should be scaled with the resolution. Although there is certainly a theoretical argument for such scaling, our testing of this has shown that the resolved fields do not evolve as favourably as in the CsCoMLS runs if the perturbations are scaled down. We conclude that allowing the full range of $\theta'$ from Eq. 2 is a pragmatically preferable solution for achieving an optimal fit to the coarse-grained fields later in the simulation.

As a final note, we have recorded some run times to test the computational burden of the new methods, and find an increase in run time of $\approx 10\%$ for the CsCoMLS400 run compared to the CNTL400 run. However, this increase also includes the calculation of $z_i$ at each timestep, which is not calculated by default in the MONC model (unlike other models, e.g. the WRF model).
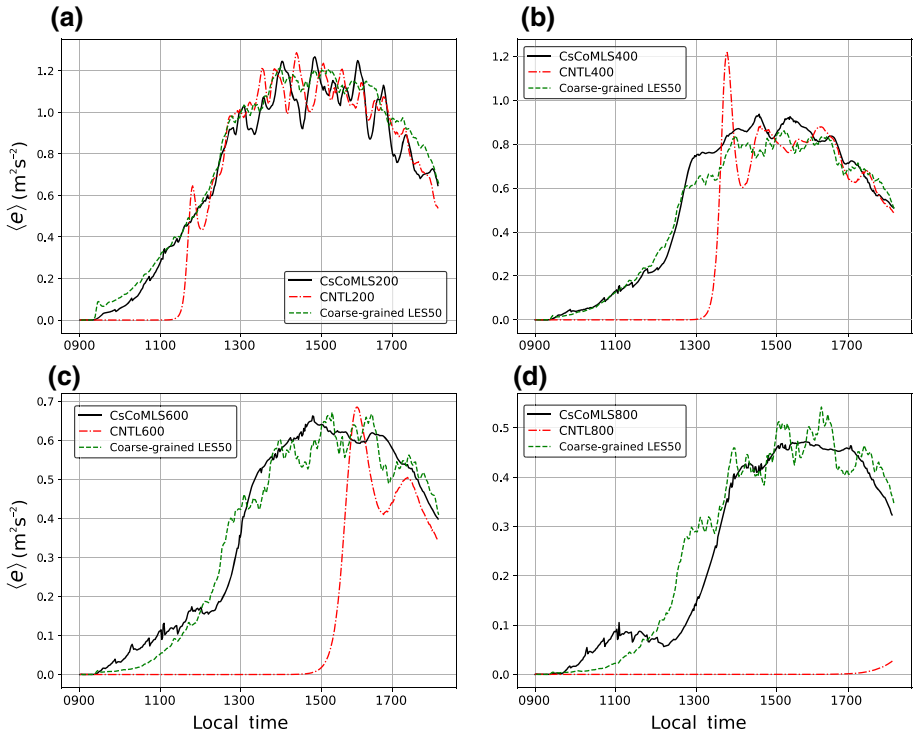
**Fig. 7** Time series of $\langle e \rangle$ for the resolutions: (**a**) $\Delta x = 200$ m, (**b**) $\Delta x = 400$ m, (**c**) $\Delta x = 600$ m, and (**d**) $\Delta x = 800$ m, using a combination of the mixed-layer scaling method and our modified $C_S$ coefficient. In each plot, $\beta = 1.0$. Also shown are the 50-m LES run (coarse-grained to each of the resolutions) and the time series for each CNTL simulation. Note the differences in scale for each y-axis

### 3.5 Further Analysis of CsCoMLS400

The resolution $\Delta x = 400$ m is a useful test-bed for the techniques presented here since it lies firmly within the grey zone, which is why it has been a focus in previous sections. This subsection presents a deeper analysis of the best performing of our methods in terms of improved spin-up time and agreement with the coarse-grained fields. CsCoMLS400 combines the use of our modified Smagorinsky coefficient with the mixed-layer scaling method of perturbing $\theta$.

Two-dimensional spectra comparing the CNTL400 simulation with CsCoMLS400 are shown in Fig. 8. Since the CNTL400 run is unable to spin-up any turbulence in the resolved fields until near the halfway point of the run, the result is near-negligible values of $S_{\overline{w'^2}}$ at 1230 LT (Fig. 8a), while the CsCoMLS400 simulation is capable of developing spectra that are much closer to the LES model in magnitude. Naturally, much of the inertial subrange present in the LES model is absent in the grey-zone runs, and the point of departure from the ideal Kolmogorov $k^{-5/3}$ law is quite close to the peak wavelength. Since the CsCoMLS400 simulation exhibits the correct shape and magnitude in the spectra, this strengthens our argument that the manner in which these fields are induced does indeed give rise to physically legitimate structures. Later, at $t = 6$ h (Fig. 8b), the two simulations show little difference in shape and magnitude.
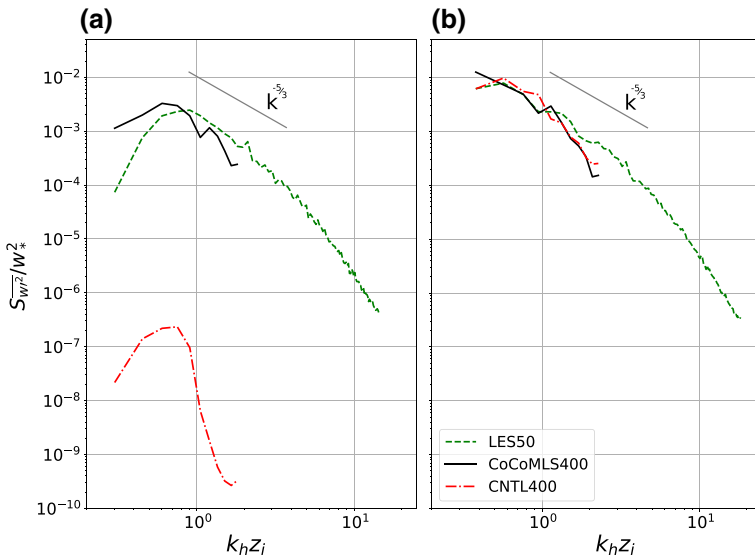
**Fig. 8** 2D normalized power spectra of vertical velocity as a function of the normalized horizontal wave number in the mid-CBL for (**a**) the CsCoMLS400 and CNTL400 simulations at $t = 3.5$ h (1230LT), and (**b**) the same at $t = 6$ h (1500LT). The LES50 run is also plotted in each frame. The $k^{-5/3}$ Kolmogorov power law is plotted in grey

Figure 9 shows the partitioning of kinematic heat flux $(\overline{w'\theta'})$ into resolved and subgrid components for the CNTL400 and CsCoMLS400 simulations, where the subgrid component is calculated using the buoyancy-flux term shown in Eq. 11b. Early in the simulation at $t = 3.5$ h, the Smagorinsky scheme in the CNTL400 run is not resolving any motion and there is no entrainment present, leading to a CBL depth that is $\approx 350$ m lower than that of LES50. The CsCoMLS400 run improves upon this by inducing resolved motion earlier, thus generating entrainment and non-local transport. The inversion height of CsCoMLS400 matches that of the LES50 run to within $\approx 30$ m. Later, the CNTL400 run adjusts towards the LES model solution, after the resolved fields have spun-up, again highlighting the importance of inducing resolved turbulence, and doing so as early as possible. By 1500LT, the LES model run (Fig. 9b), CNTL400 run (Fig. 9d) and CsCoMLS400 run (Fig. 9f) all agree well with each other, with the subgrid and resolved components of the grey-zone runs giving rise to a total-flux profile that matches well with the LES model.

## 4 Discussion and Conclusions

Several grey-zone simulations using the UK Met Office MONC model have been performed to better understand the spin-up of partially-resolved turbulent motions. Firstly, to address the question of whether or not this resolved turbulence should be allowed in the grey zone at all, our results show that allowing such resolved turbulence provides much of the necessary non-local transport to complement the 3D static Smagorinsky scheme, even in the absence of the countergradient correction term that is characteristic of 1D PBL schemes. This is true for a range of $z_i/\Delta x$, although as the model resolution approaches the mesoscale limit, a need for such a non-local transport term does begin to arise. It is here that one might benefit most
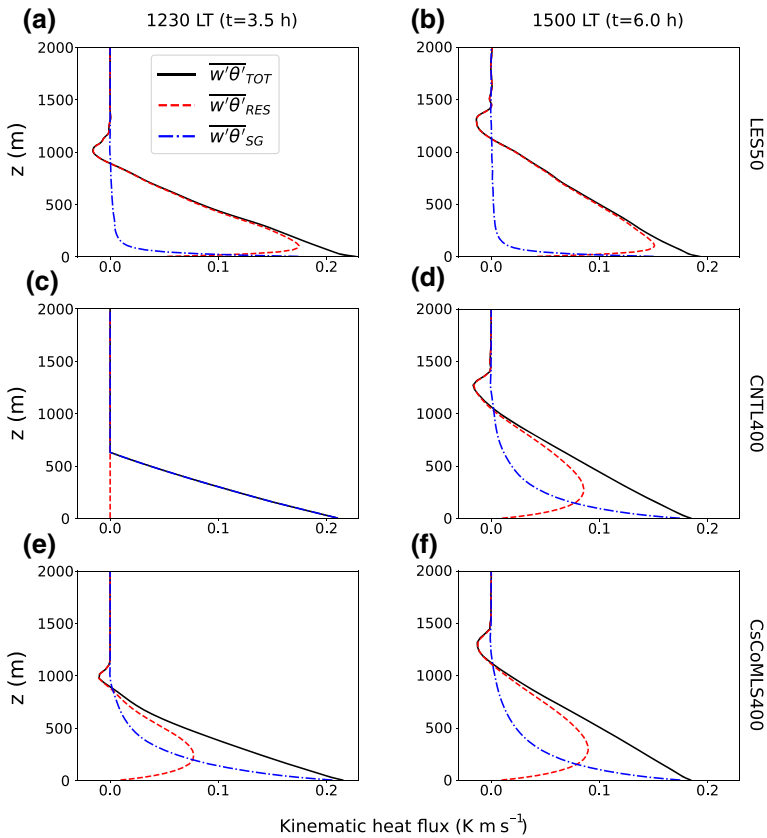
**Fig. 9** Vertical profiles of total kinematic heat flux (solid black), partitioned into subgrid (dot-dashed blue) and resolved (dashed red) components for, (**a**) LES50 at $t = 3.5$ h; (**b**) LES50 at $t = 6$ h; (**c**) CNTL400 at $t = 3.5$ h; (**d**) CNTL400 at $t = 6$ h; (**e**) CsCoMLS400 at $t = 3.5$ h; and (**f**) CsCoMLS400 at $t = 6$ h

from blending schemes such as that of Boutle et al. (2014). Nonetheless, we have shown that the sooner the resolved motion can be established, the more non-local transport is provided to the system, and the closer the results become to those of a coarse-grained LES model. Based on this, we have investigated why there exists a delay before resolved turbulence appears and whether this delay can be reduced.

Our results suggest that the grey-zone CBL is highly sensitive to structured, pseudo-random perturbations applied to the $\theta$ field. Without these perturbations, we observe no resolved turbulence in the model output at all. When applied to the initial state, the perturbations allow turbulence to spin-up with a delay that is proportional to the grid spacing. Since grey-zone grids are prone to producing grid-scale convection, the application of these perturbations (which is also done at the grid scale) becomes paramount to the development of a well-mixed and well-behaved CBL.

We suggest that the optimal grey-zone configuration consists not only of a customized parametrization, but in fact a customization to both the resolved and subgrid fields simultaneously. The logic in doing so lies in the fact that the grey zone itself can be defined by the coexistence of both resolved and subgrid components. We modify the Smagorinsky turbulence closure by reassigning the Smagorinsky constant as a coefficient with dependency

on the variable $z_i/\Delta x$, so that the eddy diffusivity in the newly formed CBL becomes more sensitive to the sensible heat fluxes at the lower boundary. The importance of the relationship between $z_i$ and $\Delta x$ is becoming increasingly apparent in grey-zone modelling, and here we base our modifications of the Smagorinsky scheme on this relationship.

At the same time, we apply pseudo-random perturbations to the $\theta$ field at intervals of $t_*$, thereby encouraging the natural variation of $\theta$ within the CBL. We have also carried out trial simulations with the perturbations applied to $w$ instead of $\theta$, but preliminary results of these tests have shown that the impact of perturbing $w$ is similar to perturbing $\theta$. Various methods of organizing the perturbations have been tested, including increasing the perturbation amplitude, applying coherence in the vertical dimension, applying uniform perturbations to bi-dimensional cells in the horizontal, and employing mixed-layer scaling theory to selecting the perturbation amplitude at each vertical model level. The mixed-layer scaling method is shown to provide the best result, with spin-up greatly enhanced, even at resolutions approaching the mesoscale limit. Of the two modifications (i.e. to the parametrization and to the resolved fields), it is our opinion that modifying the resolved fields tends to play a more significant role in allowing faster spin-up and establishing a well-mixed CBL.

Although our method is shown to be useful at resolutions approaching the mesoscale limit, the question arises as to when one should assume that RANS modelling becomes valid, and switch from a 3D scheme to a 1D scheme. There do exist methods that blend 1D and 3D schemes (Boutle et al. 2014), but here we have not yet supplied a means of determining at what resolution the imposed $\theta$ perturbations should cease. For Wangara, this limit would be in the vicinity of $\Delta x \approx 1000$ m, but generalizing the approach presented here to apply to any situation is beyond the scope of our work. In fact, it is as yet unclear whether a working generalized method of inducing spin-up (which applies to any CBL) is possible. We have shown that the grey zone exhibits strong sensitivities in how the resolved fields behave, and documenting every one of these sensitivities is simply not possible. However, in future we hope to apply our methods to real cases, such that inciting spin-up at an earlier time might have a positive impact on the entire atmospheric system in high-resolution NWP simulations. One such impact might be to encourage deep convection (Stirling and Petch 2004), since this process is fundamentally driven by the CBL.

Although operational NWP is, in general, still at the point of partially-resolving deep convection only (and not boundary-layer turbulence), there are regions of the world where these approaches may be practically useful. One such example is the Saharan atmospheric boundary layer, where the CBL grey zone can exist even at grid spacings as large as 4 km. In the future, as the realm of sub-kilometric resolutions becomes more commonplace in operational NWP, it will be more important than ever to understand the grey zone, and to identify which modifications can be applied to help with adapting to this partially-resolved regime.

# References

André J, De Moor G, Lacarrere P, Du Vachat R (1978) Modeling the 24-hour evolution of the mean and turbulent structures of the planetary boundary layer. J Atmos Sci 35(10):1861–1883

Beare R, Edwards J, Lapworth A (2006a) Simulation of the observed evening transition and nocturnal boundary layers: Large-eddy simulation. Q J R Meteorol Soc 132(614):81–99

Beare RJ (2008) The role of shear in the morning transition boundary layer. Boundary-Layer Meteorol 129(3):395–410

Beare RJ (2014) A length scale defining partially-resolved boundary-layer turbulence simulations. Boundary-Layer Meteorol 151(1):39–55

Beare RJ, Macvean MK, Holtslag AA, Cuxart J, Esau I, Golaz JC, Jimenez MA, Khairoutdinov M, Kosovic B, Lewellen D et al (2006b) An intercomparison of large-eddy simulations of the stable boundary layer. Boundary-Layer Meteorol 118(2):247–272

Boutle I, Eyre J, Lock A (2014) Seamless stratocumulus simulation across the turbulent gray zone. Mon Weather Rev 142(4):1655–1668

Brown AR, Derbyshire S, Mason PJ (1994) Large-eddy simulation of stable atmospheric boundary layers with a revised stochastic subgrid model. Q J R Meteorol Soc 120(520):1485–1512

Brown AR, MacVean MK, Mason PJ (2000) The effects of numerical dissipation in large eddy simulations. J Atmos Sci 57(19):3337–3348

Brown N, Lepper A, Weiland M, Hill A, Shipway B, Maynard C (2015) A directive based hybrid met office nerc cloud model. In: Proceedings of the Second Workshop on Accelerator Programming using Directives, ACM

Carpenter K (1979) An experimental forecast using a non-hydrostatic mesoscale model. Q J R Meteorol Soc 105(445):629–655

Ching J, Rotunno R, LeMone M, Martilli A, Kosovic B, Jimenez P, Dudhia J (2014) Convectively induced secondary circulations in fine-grid mesoscale numerical weather prediction models. Mon Weather Rev 142(9):3284–3302

Clarke RH, Dyer AJ, Brook RR, Reid DG, Troup AJ (1971) Wangara experiment: boundary layer data. Tech Rep 340, CSIRO

Efstathiou G, Beare RJ (2015) Quantifying and improving sub-grid diffusion in the boundary-layer grey zone. Q J R Meteorol Soc 141(693):3006–3017

Efstathiou G, Beare R, Osborne S, Lock A (2016) Grey zone simulations of the morning convective boundary layer development. J Geophys Res Atmos 121(9):4769–4782

Efstathiou G, Plant R, Bopape MJ (2018) Simulation of an evolving convective boundary layer using a scale-dependent dynamic smagorinsky model at near-gray-zone resolutions. J Appl Meteorol Climatol 57(9):2197–2214

Garratt JR (1994) The atmospheric boundary layer. Cambridge University Press, Cambridge, UK

Germano M, Piomelli U, Moin P, Cabot WH (1991) A dynamic subgrid-scale eddy viscosity model. Phys Fluids 3(7):1760–1765

Gray M, Petch J, Derbyshire S, Brown A, Lock A, Swann H, Brown P (2001) Version 2.3 of the met office large eddy model: Part ii. Scientific documentation Met Office (APR) Turbulence and Diffusion Rep 276

Honnert R, Masson V, Couvreux F (2011) A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale. J Atmos Sci 68(12):3112–3131

Honnert R, Couvreux F, Masson V, Lancz D (2016) Sampling the structure of convective turbulence and implications for grey-zone parametrizations. Boundary-Layer Meteorol 160(1):133–156

Ito J, Niino H, Nakanishi M, Moeng CH (2015) An extension of the Mellor–Yamada model to the terra incognita zone for dry convective mixed layers in the free convection regime. Boundary-Layer Meteorol 157(1):23–43

Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, Halliwell C (2008) Characteristics of high-resolution versions of the met office unified model for forecasting convection over the united kingdom. Mon Weather Rev 136(9):3408–3424

Leonard B, MacVean M, Lock A (1993) Positivity-preserving numerical schemes for multidimensional advection. Tech Rep 62

Lilly D (1967) The representation of small-scale turbulence in numerical simulation experiments. In: proceedings of the ibm scientific computing symposium on environmental sciences. Yorktown Heights, NY

Lock A, Brown A, Bush M, Martin G, Smith R (2000) A new boundary layer mixing scheme. Part i: scheme description and single-column model tests. Mon Weather Rev 128(9):3187–3199

Mason PJ, Brown AR (1999) On subgrid models and filter operations in large eddy simulations. J Atmos Sci 56(13):2101–2114

Mass CF, Ovens D, Westrick K, Colle BA (2002) Does increasing horizontal resolution produce more skillful forecasts? Bull Am Meteorol Soc 83(3):407–430

Meyers J, Sagaut P (2006) On the model coefficients for the standard and the variational multi-scale Smagorinsky model. J Fluid Mech 569:287–319

Mirocha J, Kosović B, Kirkil G (2014) Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the weather research and forecasting model. Mon Weather Rev 142(2):806–831

Muñoz-Esparza D, Kosović B, Mirocha J, van Beeck J (2014) Bridging the transition from mesoscale to microscale turbulence in numerical weather prediction models. Boundary-Layer Meteorol 153(3):409–440

Nakanishi M, Shibuya R, Ito J, Niino H (2014) Large-eddy simulation of a residual layer: low-level jet, convective rolls, and Kelvin–Helmholtz instability. J Atmos Sci 71(12):4473–4491

Petch J, Brown A, Gray M (2002) The impact of horizontal resolution on the simulations of convective development over land. Q J R Meteorol Soc 128(584):2031–2044

Piacsek SA, Williams GP (1970) Conservation properties of convection difference schemes. J Comput Phys 6(3):392–405

Porté-Agel F, Meneveau C, Parlange MB (2000) A scale-dependent dynamic model for large-eddy simulation: application to a neutral atmospheric boundary layer. J Fluid Mech 415:261–284

Shin HH, Dudhia J (2016) Evaluation of pbl parameterizations in wrf at subkilometer grid spacings: turbulence statistics in the dry convective boundary layer. Mon Weather Rev 144(3):1161–1177

Shin HH, Hong SY (2015) Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions. Mon Weather Rev 143(1):250–271

Simmons A, Burridge D, Jarraud M, Girard C, Wergen W (1989) The ecmwf medium-range prediction models development of the numerical formulations and the impact of increased resolution. Meteorol Atmos Phys 40(1–3):28–60

Smagorinsky J (1963) General circulation experiments with the primitive equations: I the basic experiment. Mon Weather Rev 91(3):99–164

Stirling A, Petch J (2004) The impacts of spatial variability on the development of convection. Q J R Meteorol Soc 130(604):3189–3206

Wyngaard JC (2004) Toward numerical modeling in the terra incognita. J Atmos Sci 61(14):1816–1826

Yamada T, Mellor G (1975) A simulation of the Wangara atmospheric boundary layer data. J Atmos Sci 32(12):2309–2329

Zhou B, Simon JS, Chow FK (2014) The convective boundary layer in the terra incognita. J Atmos Sci 71(7):2545–2563