Scale Awareness, Resolved Circulations, and Practical Limits in the MYNN–EDMF Boundary Layer and Shallow Cumulus Scheme

WAYNE M. ANGEVINE,[a,b] JOSEPH OLSON,[c] JAKE J. GRISTEY,[a,b] IAN GLENN,[a,b] GRAHAM FEINGOLD,[b] AND DAVID D. TURNER[c]

[a] *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado*
[b] *NOAA/Chemical Sciences Laboratory, Boulder, Colorado*
[c] *NOAA/Global Systems Laboratory, Boulder, Colorado*

(Manuscript received 25 February 2020, in final form 10 September 2020)

ABSTRACT: Proper behavior of physics parameterizations in numerical models at grid sizes of order 1 km is a topic of current research. Modifications to parameterization schemes to accommodate varying grid sizes are termed ''scale aware.'' The general problem of grids on which a physical process is partially resolved is called the ''gray zone'' or ''terra incognita.'' Here we examine features of the Mellor–Yamada–Nakanishi–Niino (MYNN) boundary layer scheme with eddy diffusivity and mass flux (EDMF) that were intended to provide scale awareness, as implemented in WRF, version 4.1. Scale awareness is provided by reducing the intensity of nonlocal components of the vertical mixing in the scheme as the grid size decreases. However, we find that the scale-aware features cause poorer performance in our tests on a 600-m grid. The resolved circulations on the 600-m grid have different temporal and spatial scales than are found in large-eddy simulations of the same cases, for reasons that are well understood theoretically and are described in the literature. The circulations [model convectively induced secondary circulations (M-CISCs)] depend on the grid size and on details of the model numerics. We conclude that scale awareness should be based on effective resolution, and not on grid size, and that the gray-zone problem for boundary layer turbulence and shallow cumulus cannot be solved simply by reducing the intensity of the parameterization. Parameterizations with different characteristics may lead to different conclusions.

KEYWORDS: Boundary layer; Cumulus clouds; Numerical weather prediction/forecasting; Parameterization; Single column models; Subgrid-scale processes

1. Introduction

Numerical models for weather prediction and research are run on grids with spacing decreasing toward 1 km and smaller. On these grids, some physical processes are partially resolved. Traditionally, on coarser grids, these ''subgrid'' processes have been handled entirely by parameterizations. A fundamental assumption of parameterization development has been that an ensemble of the relevant entities (e.g., turbulent eddies or clouds) exists within the grid cell, and that the properties of that ensemble can be described by deterministic physical or statistical formulations (Arakawa 2004). Different processes have different characteristic scales. For example, the most important scales of boundary layer turbulence vary from ~100 m (or less) in stable or neutral (Honnert 2019) conditions to ~1000 m (or more) in convective boundary layers (Honnert et al. 2011), roughly proportional to the boundary layer depth. For clouds, the important scales are a few hundred meters for shallow cumulus in the boundary layer, up to several km or more for deep moist convective clouds. As grid sizes decrease below ~10 km, more and more processes enter the partially resolved regime. This work is concerned with boundary layer turbulence and shallow cumulus under convective conditions, where the dominant scales are between 100–200 m and 1–2 km.

The range of model resolutions in which some process of interest is partially resolved is called the ''gray zone'' or, for turbulence, ''terra incognita'' (Wyngaard 2004). It is important to keep in mind that the gray-zone scales are different for different processes.

Model resolution and grid spacing are not equal. This will become a key point later in the paper. In a perfect model, the smallest sine wave that can be crudely resolved has wavelength $2\Delta x$, where Δx is the grid spacing (assumed to be the same in both horizontal directions). This is simply the Nyquist criterion. However, practical numerics for the advection of quantities in numerical models of the atmosphere never achieve this resolution, as they introduce numerical diffusion for reasons having to do with numerical stability among other considerations. Studies have shown that the effective resolution is $6\Delta x$–$8\Delta x$, or even coarser (Beare 2014; Lean et al. 2019; Skamarock 2004). Note also that we use the terms ''fine'' and ''coarse'' throughout the paper to describe both grid spacing and resolution, in preference to ''high'' and ''low,'' which can be confusing.

Our working postulate is that the goal of scale awareness is to produce approximately the same model solution regardless of the grid spacing. In this case, by ''model solution,'' we mean the state variables. The benefits of finer resolution are expected to come from other aspects of the model system, for example the terrain. In situations where a process can be considered fully resolved, it is hoped that removing the uncertainties associated with parameterization will produce better results. This is the philosophy behind large-eddy simulation (LES), which resolves the important and variable scales of turbulence, leaving only small scales that are assumed to be easily handled

Corresponding author: Wayne Angevine, wayne.m.angevine@noaa.gov

DOI: 10.1175/MWR-D-20-0066.1

by a simple parameterization based on the theory of isotropic turbulence in the inertial subrange. The reduction of uncertainty is also commonly cited as a reason to use grids fine enough to avoid the need for parameterization of deep convection. Here we are dealing with gray-zone situations where some parameterization is still needed, hence the above postulate. In the terms defined by Panosetti et al. (2019), we are looking for bulk convergence. To paraphrase those terms, bulk convergence is achieved when the state variables of the model are the same regardless of the grid spacing, within some range of interest. Structural convergence requires that the physical processes (e.g., "sizes and properties of individual clouds and updrafts") are the same. However, although we are not looking for structural convergence, we will use the structures that the model produces to explain why bulk convergence is better without some scale-aware features.

Most of the published development of scale-aware parameterizations for the boundary layer and shallow cloud has been based on LES studies (Honnert et al. 2011; Shin and Hong 2013). These studies apply averaging via top-hat filters to LES output in order to diagnose what fraction of vertical mixing is resolved versus subgrid at various grid spacings. The top-hat filtered LES outputs have resolution $2\Delta x$. Since the LES contain essentially all relevant scales of motion, the filtered output is influenced by all scales. That is, even though the filtered output does not contain actual motions on scales less than $2\Delta x$, the influence of those scales on larger scales is still present (Piotrowski et al. 2009). These LES studies do not address the structure or appropriateness of the resolved motions. One study that used mesoscale numerics showed bulk but not structural convergence between grid spacings of 100 and 50 m (Lean et al. 2019).

Resolved motions on gray-zone grids differ from real atmospheric motions. They depend on the grid spacing as well as atmospheric variables. They develop more slowly and produce incorrect patterns. These motions are described as model convectively induced secondary circulations (M-CISCs). The differences between real and modeled CISCs are described by Ching et al. (2014), Piotrowski et al. (2009), Poll et al. (2017), and Zhou et al. (2014). Similar phenomena arise in partially resolved simulations of deep convection (e.g., Bryan and Morrison 2012).

One form of scale awareness is present in all models. Mixing of heat and moisture by both parameterized and partially resolved processes takes place roughly simultaneously (within each model time step). As grid spacing decreases and partially resolved circulations become more intense, more of the excess heat and moisture (buoyancy) at the surface are transported into the bulk of the boundary layer by the partially resolved processes. This means that less excess buoyancy is available to drive the parameterized fluxes, thus reducing the intensity of parameterized mixing, independent of any explicit grid dependency in the code.

In this work, we use a multicolumn idealized model (MCM) framework to examine the behavior of the Weather Research and Forecasting (WRF) Model with the Mellor–Yamada–Nakanishi–Niino (MYNN) boundary layer scheme with eddy diffusivity and mass flux (EDMF) on a grid with 600-m spacing. We consider the developed resolved structures and how they interact with the parameterization. Cases of shallow cumulus developed by the LES ARM Symbiotic Simulation and Observation (LASSO) project are used, and LESs of those cases are used as the primary reference for correct behavior. We compare the multicolumn results with single-column model (SCM) results at a nominal grid spacing of 13 km, outside the gray zone (of turbulence and shallow convection). We find that the best correspondence between the 600-m and 13-km simulations occurs when the parameterization is at full strength, that is, without the model developer's enforced tapering of the parameterized mixing by the inclusion of scale-aware features. When the parameterization intensity is reduced, the resolved structures (M-CISCs) produce clouds that form too late and have too little liquid water path. The results also depend on whether we turn on an explicit numerical filter, which further coarsens the effective resolution.

All practical parameterization schemes contain code that implements limits to preserve numerical stability. These limits are not always described in the scheme documentation and can have important effects that may or may not have been intended by the developers. In this study, we look at a limit that turns off the mass flux when the vertical velocity on the grid is significant. Removal of that limit is necessary for resolved motions to develop. In essence, the "w limit" conflicts with the scale-aware features. Alternatively, it could be considered an additional scale-aware feature itself.

Some studies have shown improved performance of mesoscale models when parameterized mixing is reduced at gray-zone scales (Boutle et al. 2014; Lancz et al. 2018). Because different parameterization schemes have different behavior (Shin and Dudhia 2016), we caution that the results herein may not apply to other schemes. Furthermore, different kinds of cases (e.g., dry vs shallow vs deep convection vs stratocumulus) may call forth different behavior. Approaches involving blending of 3D diffusion and 1D vertical mixing may provide a way forward (Efstathiou and Plant 2019; Ito et al. 2015; Kurowski and Teixeira 2018). However, the basic theoretical and practical considerations that we describe should be considered in evaluation of all schemes and systems.

2. The MYNN–EDMF scheme in WRF

The MYNN–EDMF parameterization for the boundary layer and shallow cloud is derived from the earlier MYNN scheme by the addition of a mass flux representation of nonlocal mixing by thermals. MYNN–EDMF is described in detail by Olson et al. (2019a,b). The mass flux part of the scheme uses up to 10 plumes, representing discrete plume diameters from 100 to 1000 m. The number of plumes and the maximum plume diameter are controlled by the boundary layer height diagnosed within the scheme, and by the horizontal grid spacing. This work uses the scheme as implemented in version 4.1 of WRF.

For scale awareness, the maximum plume diameter is limited by the grid spacing Δx in addition to the boundary layer height. For example, if the boundary layer height is 1500 m but the grid spacing is 600 m, only the smallest six plumes (up to 600-m diameter) are used. Since the larger plumes entrain less and therefore reach higher and carry more flux, this effectively

reduces the amount of nonlocal mixing in the scheme at smaller grid spacing. In the eddy diffusivity part of the scheme, the mixing length formulation is blended from a purely mesoscale form valid at large grid sizes ($\Delta x > 1000$ m) to a purely local form valid at small (LES) grid sizes. The nonlocal attributes of the mesoscale form and local form are documented in Olson et al. (2019a).

The MYNN–EDMF contains several practical limits to promote numerical stability in an operational environment, since this scheme is used in the Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR) models (Benjamin et al. 2016) that are run hourly by the U.S. National Weather Service. Among these, the most important for the present study is a limit that shuts off the mass flux if the vertical velocity on the grid scale (the model state variable w) within the boundary layer is outside the range $-1 < w < 2$. The mass flux is linearly tapered within these limits. We refer to this as the "w limit." The runs presented here have this limit turned off to allow the other scale-aware features to express themselves. The results below are therefore not exactly what would be achieved using the released code.

The WRF numerics in this study uses the default settings—in particular, third-order Runge–Kutta time stepping, fifth-order horizontal advection, and third-order vertical advection. Recommended coefficients of 0.1 for divergence damping (smdiv) and 0.01 for the external mode filter (emdiv) are used. The base runs have the sixth-order filter (diff_6th_opt) turned on with settings as in the operational RAP/HRRR models; see the code that is provided in the data statement for the exact settings. This has the effect of adding some diffusion and slightly coarsening the effective resolution.

3. LASSO

The LASSO project (https://www.arm.gov/capabilities/modeling/lasso/) (Gustafson et al. 2020) has selected cases of shallow cumulus over the ARM SGP site in Oklahoma (Sisterson et al. 2016) for four years, 2015–18. The project provides initial and forcing data from several methods that can be used for LES or single-column modeling and provides LES output from several different setups for each case. Selected observations are also bundled with the forcing and LES, and evaluation plots and metrics can be produced. Here we make use of one set of forcing data, created by the ARM Variational Analysis system (VARANAL; Xie et al. 2004). For the 2015 case, we use the initial sounding from VARANAL. For 2016 and 2017 cases, the observed sounding at (nominally) 1200 UTC is used to initialize the models (LES, SCM, and MCM).

Four cases are shown here. They are chosen to be the most representative and canonical shallow cumulus cases. All have moderate cloud cover in the afternoon, and no cloud at the beginning or end of the day. The cloud tops are low enough not to contain ice, and no higher layers are present. The four cases are 27 June 2015, 25 June 2016, 14 June 2017, and 17 July 2017. These are the same cases selected and analyzed by (Gristey et al. 2020). Other cases that do not meet such stringent criteria will be presented in future work.

4. LES

For the 2015 case, we use the WRF LES produced by LASSO (simulation code 30). It is run at 100-m horizontal spacing on a 14.4-km domain with 226 vertical levels. Vertical grid spacing is 30 m from the surface to 5000 m, then increases gradually to 300 m at ~10 km, and remains at 300 m to the model top at 14.7 km. For the 2016 and 2017 cases, the LES (Glenn et al. 2020; Gristey et al. 2020) are produced with the System for Atmospheric Modeling (SAM), version 6.10.10, on a horizontal grid spacing of 100 m and a vertical grid stretched from 30 m below 5 km to 300 m near the top of the domain, which is at 15 km. The domain is 24 km × 24 km in the horizontal plane.

5. Single-column and multicolumn model setup

The usual SCM setup in WRF is actually a 2 × 2 horizontal grid with doubly periodic boundary conditions and strong horizontal diffusion (Angevine et al. 2018; Hacker and Angevine 2013). Here we use this setup with a 13-km nominal grid spacing, but without the extra horizontal diffusion. Negligible differences between columns develop. To test scale awareness, we build a 42 × 42 grid with 600-m spacing, so it covers the same area as the SCM grid. Doubly periodic boundary conditions are still used. Both setups use a coupled land surface model (Angevine et al. 2018). The initialization is uniform on the grid except that the 600-m simulations start with a tiny (0.1%) random perturbation in soil moisture to break the symmetry. Coupling to the land surface is necessary to allow resolved circulations to develop, but it introduces some differences from the LES used for comparison, which is driven by specified surface fluxes. The initial soil moisture for each case is tuned to match the specified surface fluxes as closely as possible. The vertical grid in the SCM and MCM is approximately that used in the operational RAP/HRRR models, described in (Angevine et al. 2018) and shown in figure A1 of that paper. In brief, there are 50 mass layers (51 interface levels in the WRF namelist). The midpoint of the first layer is at 7.5 m, the spacing is ~23 m near the surface, increasing to ~500 m at 3.3 km AGL. The model top is at 14.6 km. The multicolumn setup is similar to that used by Shin and Dudhia (2016) except that they use 3D TKE-based diffusion (km_opt=2 in the WRF namelist), suitable for LES, even on the gray-zone grids. We use 2D horizontal diffusion (km_opt=4), suitable for simulations using a PBL scheme.

6. Results

We first present one representative case in detail. Figure 1 shows cloud base and top, LWP, and cloud cover on 27 June 2015. The 13-km SCM simulation cloud base matches the LES well, given the coarse vertical grid (51 levels; see Angevine et al. 2018). Cloud top in the SCM is lower. LWP, which we emphasize as an integral measure of the cloud and a key component of the effect on radiation, matches the LES well. All of the LWP comes from the mass flux part of the scheme, there is no grid scale cloud in the 13-km SCM. Mass flux and total LWP are not identical because they are diagnosed with

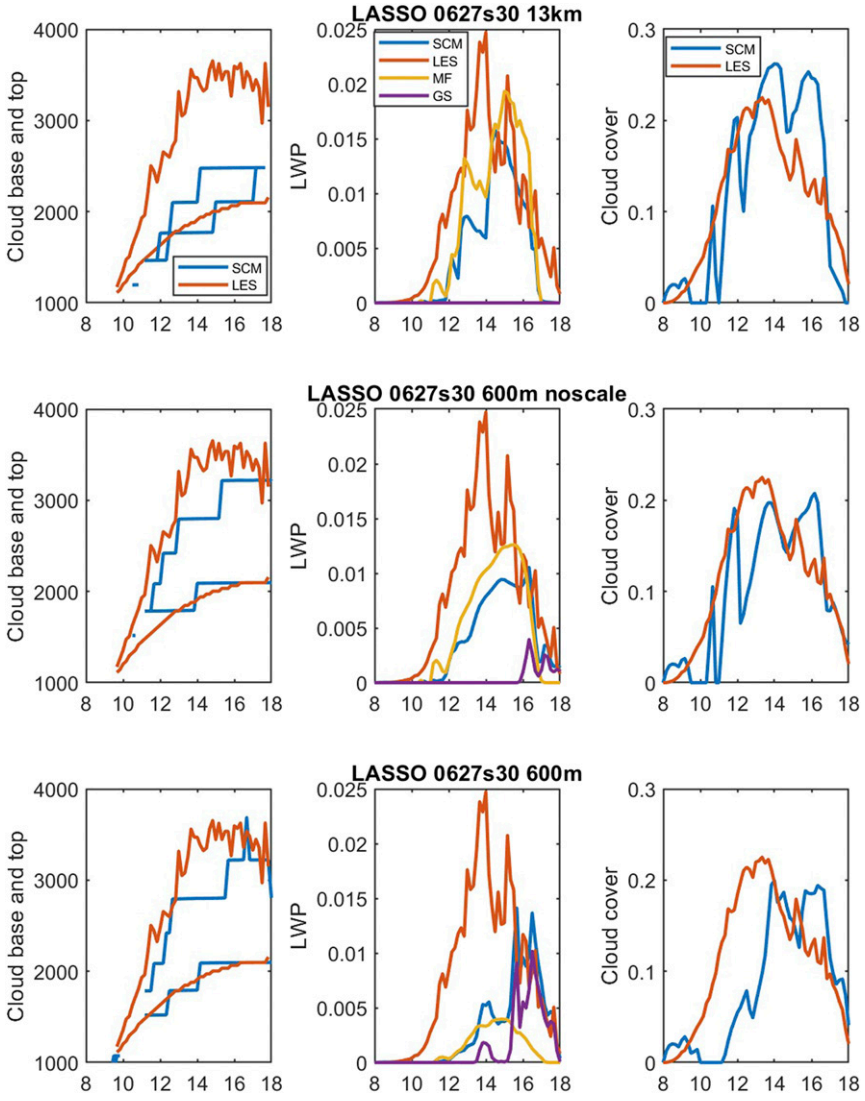

FIG. 1. (left) Cloud base and top (m MSL), (center) liquid water path (kg m^{-2}), and (right) cloud cover (fraction) on 27 Jun 2015 for (top) 13-km SCM simulations, (middle) 600-m multicolumn without scale awareness, and (bottom) 600-m MCM with scale awareness. Legends in the top row apply to each column. In the left column, LES cloud base and top are in red and SCM or MCM base and top are in blue. In the center column, LES LWP is in red, total LWP as sent to the radiation scheme is in blue, mass flux only LWP is in yellow, and gridscale LWP is in purple. In the right column, LES cloud cover is in red and SCM/MCM cloud cover as sent to the radiation scheme is in blue.

different area fractions. (The mass flux LWP is the vertical sum of the product of the updraft cloud liquid, updraft area fraction, and layer thickness; the total LWP is the vertical sum of the product of the diagnosed subgrid cloud liquid, subgrid cloud fraction, and layer thickness (see Angevine et al. 2018; Olson et al. 2019a,b for details). Cloud cover matches the LES very well. Cloud cover is calculated as the maximum cloud fraction in any layer below ~5.4 km, that is, a maximum overlap assumption. When we move to the non-scale-aware (''noscale'') 600-m MCM grid (middle row) we find a higher cloud base and cloud top. LWP is a little less than in the 13-km SCM. Some gridscale cloud forms late in the run, after 1500 LST. Cloud cover is also a little low-biased early and a little high-biased late. With scale aware features active (bottom row), cloud base and cloud top are similar. LWP is much reduced and ramps up about an hour later. As expected, mass flux cloud is less and gridscale cloud is more. Cloud cover is more significantly low-biased between 1000 and 1200 CST.

Plots of LWP for all four cases are shown in Figs. 2–5. The LES is shown for comparison. Total LWP as sent to the radiation scheme is shown in blue, and two of its components are also shown, mass flux cloud in yellow and gridscale cloud in

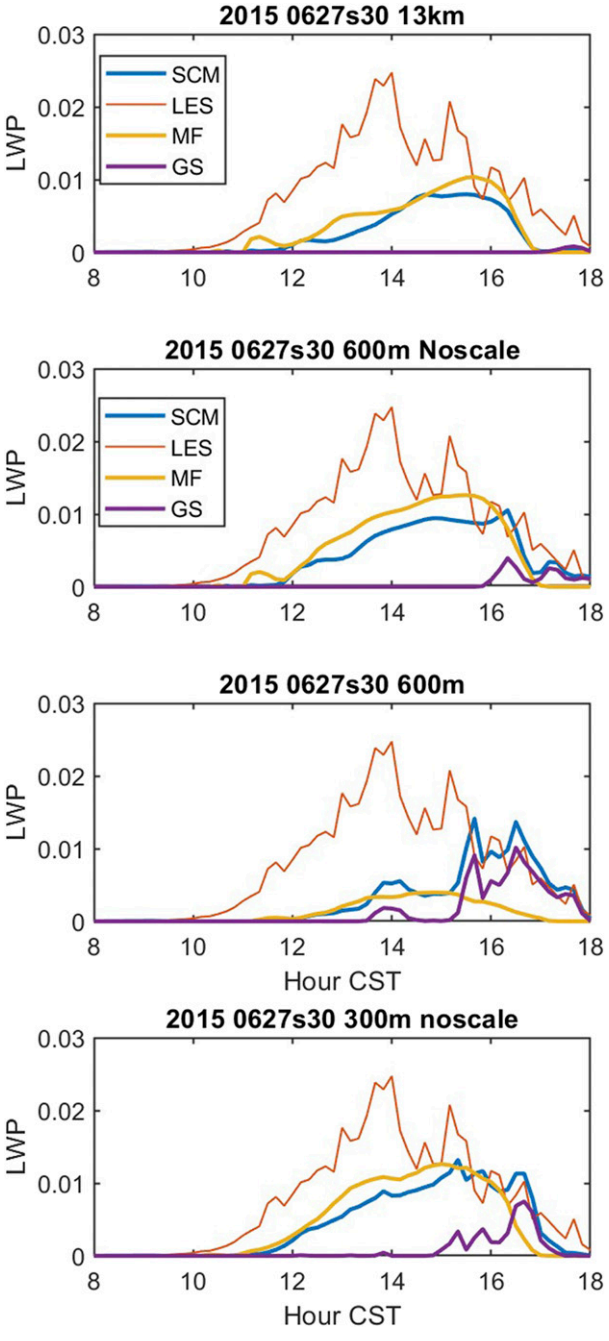

FIG. 2. Liquid water path in LES and SCM (total and components) for runs as labeled; MF is cloud from the mass flux part of the scheme, and GS is gridscale cloud. See the text for details.

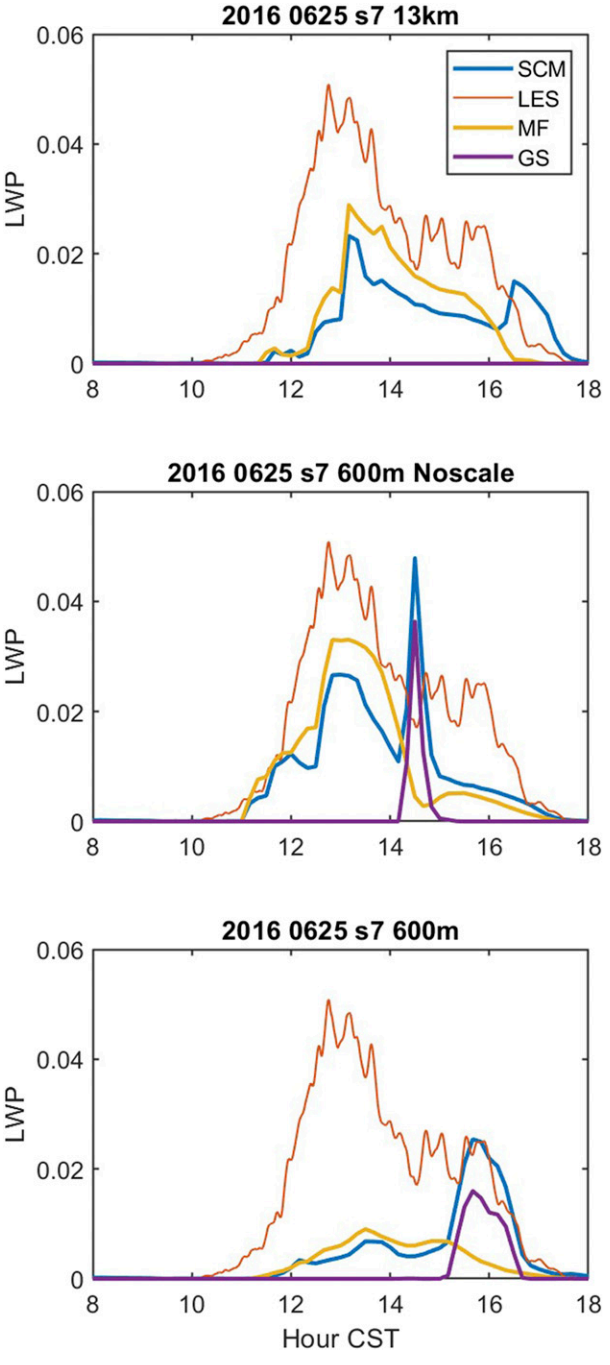

FIG. 3. Liquid water path in LES and SCM (total and components) for runs as labeled. See the text for details.

purple. For 27 June 2015 (Fig. 2), in addition to the main features already noted, we show the LWP from a run with 300-m grid spacing and no scale awareness. The domain for the 300-m run is one-half the size of the 600-m domains (same number of grid points). The LWP is very similar to what is seen in the 600-m noscale run.

On 25 June 2016 (Fig. 3), the SCM and MCM underestimate LWP with respect to the LES. The onset of cloud is delayed in all of the SCM and MCM runs, but more so in the scale-aware configuration. The scale-aware configuration produces some gridscale cloud late in the run that matches the LES rather well. On 14 June 2017 (Fig. 4), the 13-km SCM overestimates the LWP by a factor of two or more in the afternoon. The noscale MCM matches the LES well, but the scale-aware configuration underestimates LWP. All three runs start making cloud late relative to the LES, although the noscale run is a little better.

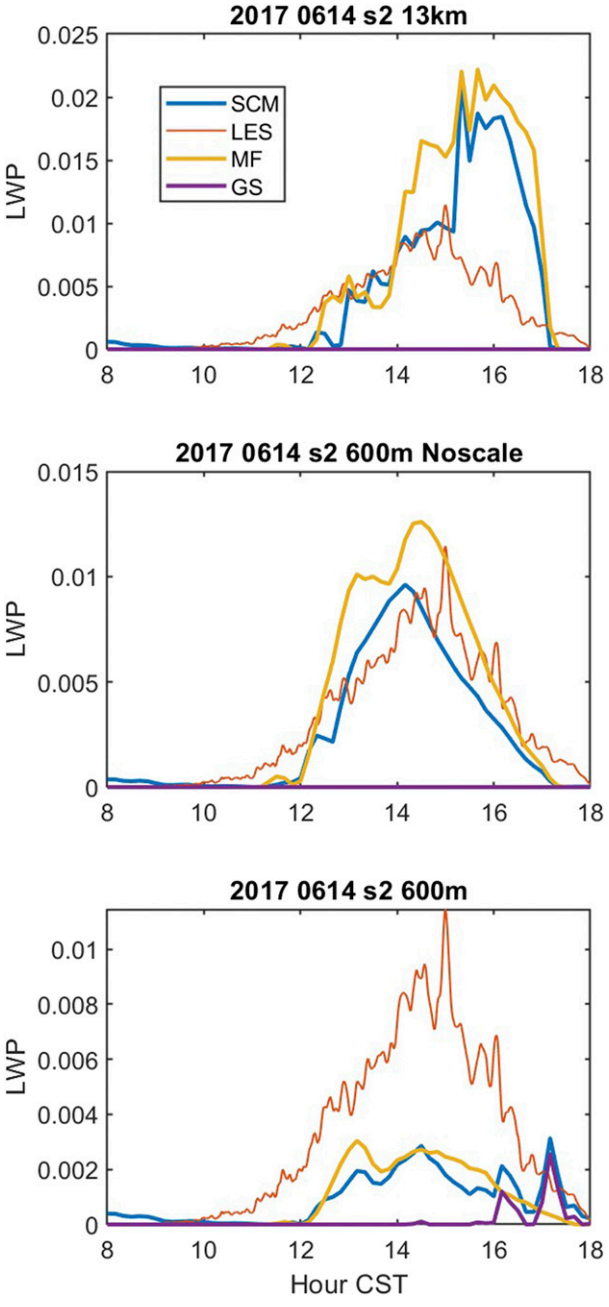

FIG. 4. Liquid water path in LES and SCM (total and components) for runs as labeled. Note the different vertical scales. See the text for details.

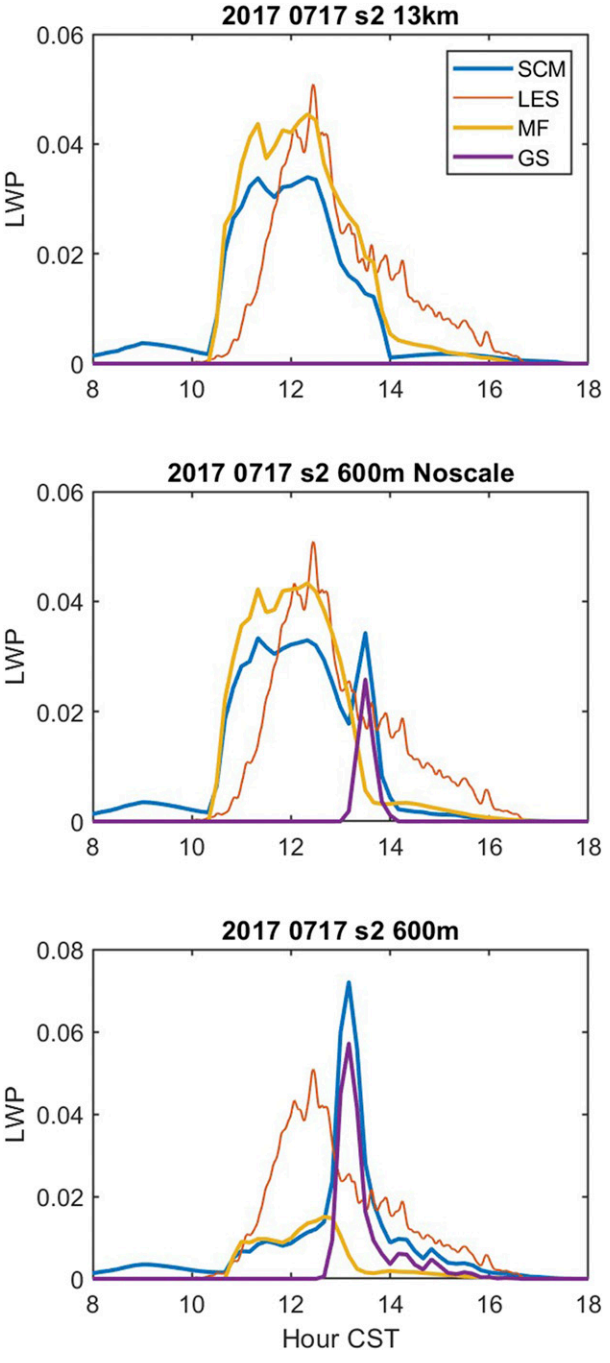

FIG. 5. Liquid water path in LES and SCM (total and components) for runs as labeled. See the text for details.

In contrast to the other cases with delayed cloud onset, on 17 July 2017 (Fig. 5), all of the runs have a small amount of subgrid cloud early. The SCM and noscale runs ramp up the LWP more quickly than the LES, with the amounts matching well by noon. The scale-aware configuration ramps the LWP up much more slowly, although it reaches about the same peak amount by 1300 CST.

Downwelling shortwave radiation at the surface is shown in Fig. 6. The observations are taken from the Radiative Flux Analysis (RFA) product (ARM 2018), averaged over the central facility and several extended facilities of the ARM SGP complex, as in (Gristey et al. 2020), and time averaged to 60 min. Nine sites are used for 25 June 2016 and 14 June 2017; only seven sites are available for 17 July 2017. On 25 June 2016, all of the models have too little attenuation before 1200 CST. Between 1200 and 1500 CST, the 600-m run with scale awareness continues to have too little attenuation, while the

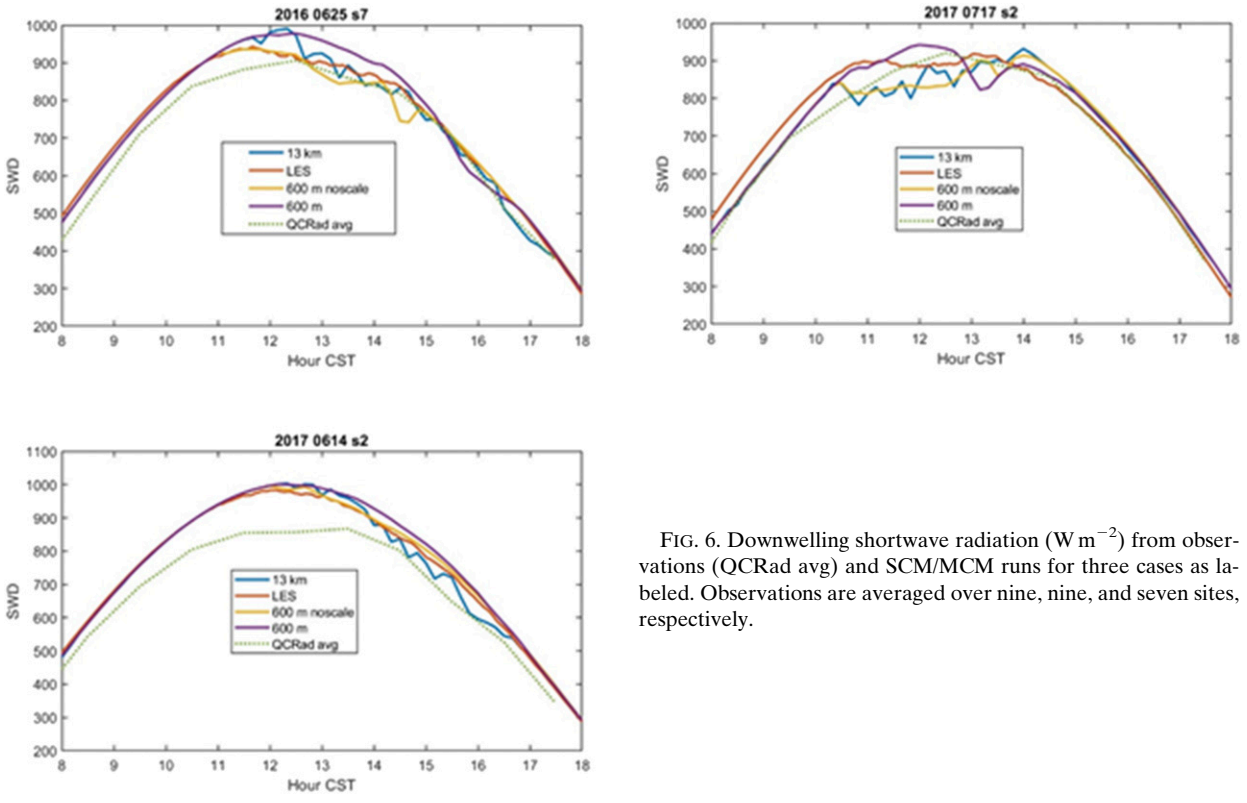

FIG. 6. Downwelling shortwave radiation (W m^{-2}) from observations (QCRad avg) and SCM/MCM runs for three cases as labeled. Observations are averaged over nine, nine, and seven sites, respectively.

others track the observations closely. On 14 June 2017, all of the models have too much downwelling shortwave through the day. This is primarily due to a large difference in cloud fraction (~15% in the models vs ~35%–40% observed). During the afternoon, the 13-km SCM compares best to the observations, while the 600-m scale-aware run is the worst. On 17 July 2017, there is more temporal variability in the model output, making comparison with the observations difficult. In the early morning (before 1100 CST) all of the SCM and MCM runs have some cloud (fraction ~10%) that is not present in the LES. The 13-km and 600-m noscale runs attenuate too much between 1100 and 1400 CST, while the 600-m scale-aware and LES runs attenuate too little. In general, the shortwave observation comparisons are consistent with the LWP as shown in the previous figures.

Although some of the cases underestimate LWP and some overestimate it, and timing differs, the main common feature of all of the cases is that the scale-aware configuration generates LWP that is too little and too late relative to the 13-km SCM or the noscale 600-m grid. The following section explores and explains this result.

7. Modeled circulations

Ching et al. (2014) named the features that arise in modeled convective boundary layers at gray-zone spacings "model convectively induced secondary circulations" or M-CISCs. The M distinguishes the modeled circulations from those arising in the real atmosphere. M-CISCs depend on the grid spacing, are slower to arise and grow than real CISCs (Efstathiou and Plant 2019; Poll et al. 2017), and also depend on the numerics of the model.

Figure 7 shows the structures in vertical velocity from MCM simulations of the 27 June 2015 case. The snapshots are taken during the most active period of the day. At 1200-m grid spacing (without scale aware features, which are never active at this grid spacing), the structures are obviously linear. Convective roll structures like these do arise in the real atmosphere within certain ranges of stability, see e.g., (LeMone 1973). As summarized by Ching et al. (2014), in the range of stability parameter h/L between −25 and 0, rolls are most likely, with cells being more common for more negative h/L values (more unstable). In this case, using cloud base for h and computing the Obukhov length L from the heat flux and surface stress, h/L is from −16 to −20, within the roll range (not shown). The four runs shown in Fig. 7 all have very similar h/L (approximately −16) except the 600-m scale aware run, which is slightly more unstable (h/L of approximately −20). At 600-m spacing, with or without scale aware features, the structures remain substantially linear. The 600-m noscale run shows more linearity than the 600-m scale-aware run because the parameterized mixing removes more of the instability, leaving less for the partially resolved motions to work with, which is also why the motions are weaker (note the color scale). At 300-m spacing, without scale aware features, the structures are cellular. This dependence on grid spacing is a clear indication that we are looking at model circulations, not atmospheric ones. Looking at Fig. 8, we see that the LES for the same case and

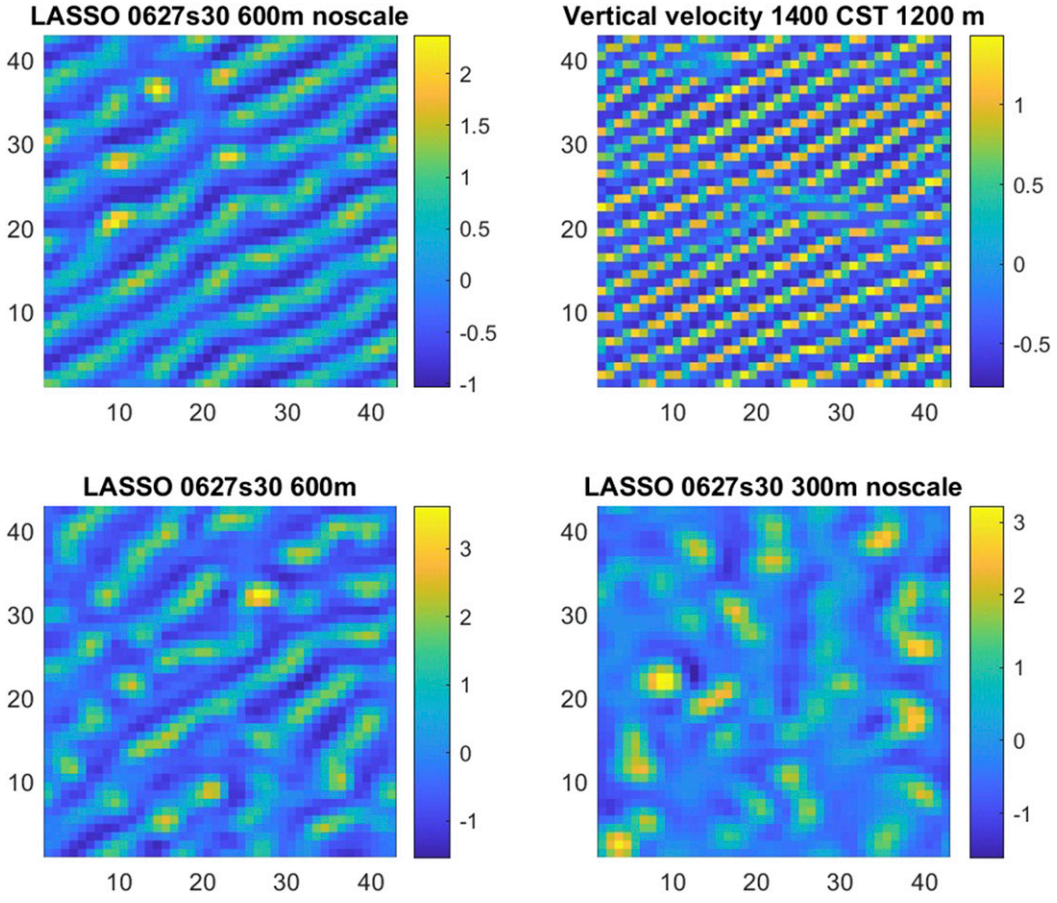

FIG. 7. Vertical velocity at 1200 m AGL 1400 CST 27 Jun 2015 in MCM simulations for grid spacings of (left) 600 m with and without scale-awareness as well as (right) 1200 and 300 m without scale-awareness. Note that the color scales differ. Axes are numbered by grid points. The same number of grid points is used in all simulations, so the 1200-m domain is 2 times the size of the 600-m domain, and the 300 m is one-half of the size.

time has yet a different structure, with no indication of linear organization and finer filaments. The LES and the 300-m MCM have the same domain size. Similar differences between MCM and LES structures can be seen in other cases [not shown here, but see Fig. 10 of Gristey et al. (2020)]. In particular, the MCM simulations are always more linear than the LES, but the no-scale simulations are not always more linear than the scale-aware simulations. One way to look at these results is to consider the underresolved model configurations to have larger effective viscosity than the real atmosphere (Piotrowski et al. 2009). This has the effect of shifting the stability ranges for rolls versus cells.

Numerical settings in the model also change the M-CISCs. WRF has the option of a sixth-order filter, designed to eliminate noise that can occur at $2\Delta x$. This has the effect of increasing horizontal diffusion and reducing effective resolution. The runs shown here have the sixth-order filter on, with the settings used in the operational HRRR model. The exact settings used are: diff_6th_factor = 0.12, moist_mix6_off = .true., chem_mix6_off = .true., tracer_mix6_off = .true., scalar_mix6_off = .true., and tke_mix6_off = .true., which results in only filtering the horizontal wind components, temperature, and water vapor (U, V, T, and Qv). When the filter is turned off, the 600-m patterns become less linear (more quasi cellular) (not shown). Again, this dependence on numerics and effective resolution indicates that we are seeing M-CISCs.

As mentioned in the introduction, there is some built-in scale awareness because the availability of instability at the surface varies with grid spacing. As the partially resolved circulations mix the surface-layer, they remove some of the instability and moisture gradient that would otherwise be available to drive the parameterization. Conversely, the strength of the parameterization affects how much instability is available to induce partially resolved circulations. Figure 9 shows the vertical gradient of potential temperature and specific humidity in the surface-layer for the 27 June 2015 case. The scale-aware 600-m run (purple line) has the largest (magnitude) gradients, that is, the most instability. This indicates that the partially resolved circulations are not removing instability (or moisture) as efficiently as the parameterization does. The 600-m noscale run (orange line) does the best at

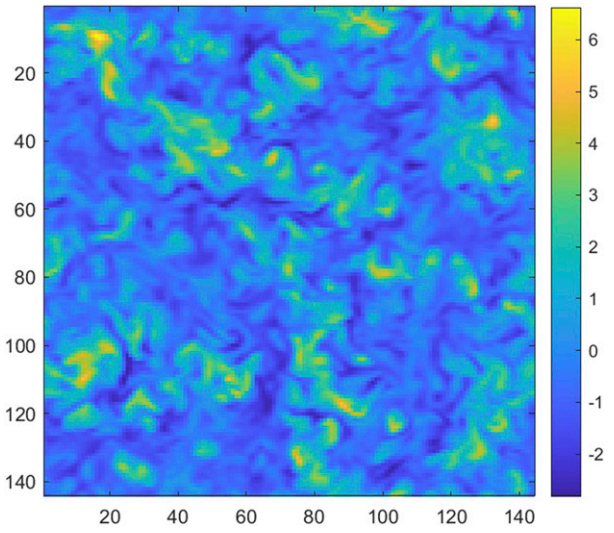

FIG. 8. Vertical velocity at 1106 m AGL from WRF LES at 1400 CST 27 Jun 2015 for comparison with Fig. 7. The LES domain is about one-half the size of the 600-m MCM domain.

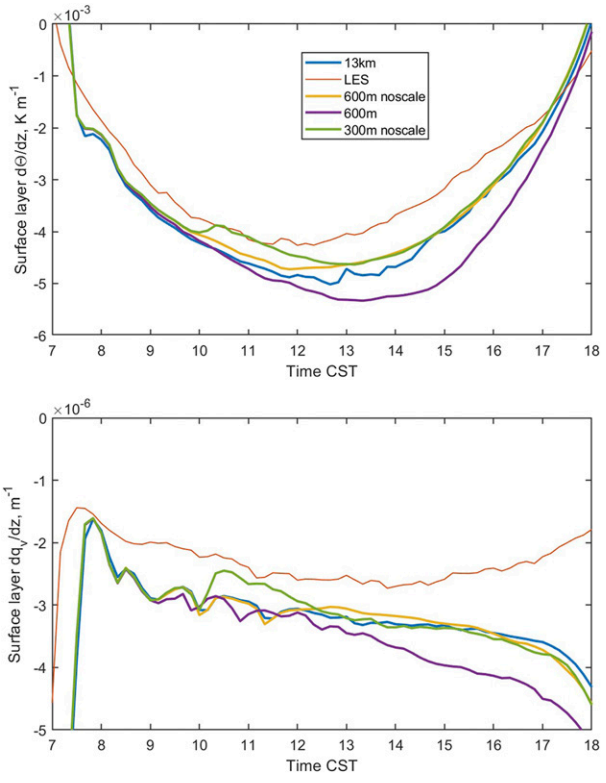

FIG. 9. Surface-layer gradients of (top) potential temperature and (bottom) specific humidity for LES and MCM runs of 27 Jun 2015. For this plot, the surface-layer is 144 m deep in the MCM and 137 m deep in the LES.

duplicating the 13-km benchmark. Between 1000 and 1300 CST, the 300-m noscale run removes more instability and moisture than the 13-km run. The LES run is shown for reference, although we do not expect perfect agreement because the SCM and MCM are running with interactive land surface. Profiles of eddy diffusivity and mass flux are shown in Fig. 10. The scale-aware features have the intended effect of reducing the mass flux, but the diffusivity is also reduced. Because these runs have a coupled land surface, the surface fluxes are not identical. The built-in scale awareness may be sufficient to produce good results on the 600-m grid without other (explicit) scale-aware features.

8. Conclusions

Multicolumn simulations of shallow cumulus over land were conducted to test whether current scale aware formulations in MYNN–EDMF, which reduce the intensity of mixing in the boundary layer and shallow cumulus scheme at finer grid spacings, are beneficial. We find that they are not. The reason is that the effective resolution of the model is considerably coarser than the grid spacing. The best results on a 600-m grid are obtained with the mass flux part of the scheme at full strength, that is, without model-developer-prescribed scale awareness. When the intensity of the mass flux is reduced, the resolved motions do not fully compensate. Clouds form too late and with too little liquid water. Even at 300-m spacing, there is little indication that the parameterized mixing is too strong.

To connect these findings to previous literature, we show that the resolved circulations depend on the grid and the numerics of the model. The pattern of vertical velocity or LWP tends to be more linear in the MCM than in LES. The best bulk convergence (Panosetti et al. 2019), that is, agreement in state variables between coarse grid SCM, fine grid MCM, and LES, is achieved by leaving the parameterization at full strength. Structural convergence between MCM and LES is not achieved in these simulations.

The use of small grid spacings (from 100 m to 1 km) in convective conditions with models formulated for mesoscale or LES applications is questionable. The assumptions underlying both mesoscale and LES models begin to lose validity in this regime and are most indefensible between 250 and 750 m, which we could call the ''dark-gray zone.'' Reasonable results may be achieved by limiting the use of poor model assumptions, but most of the important processes (i.e., thermal plumes) are neither adequately resolved or realistically parameterized in most PBL schemes. Users should try to bypass the dark-gray zone in nested model configurations as much as possible (Muñoz-Esparza et al. 2017). This has also been suggested for the gray zone of deep convection by Palmer (2019). In stable boundary layer regimes, use of small grid spacings is defensible because the important scales of motion are smaller. Blending of 3D diffusion and 1D vertical mixing may provide a way forward (Efstathiou and Plant 2019; Ito et al. 2015; Kurowski and Teixeira 2018), but this approach needs to be carefully evaluated in the context of the full modeling system. Similar approaches are under development for WRF, but at this time are not mature (Olson et al. 2019b).

The results herein apply specifically to the MYNN–EDMF scheme as implemented in WRF, version 4.1. In the future,

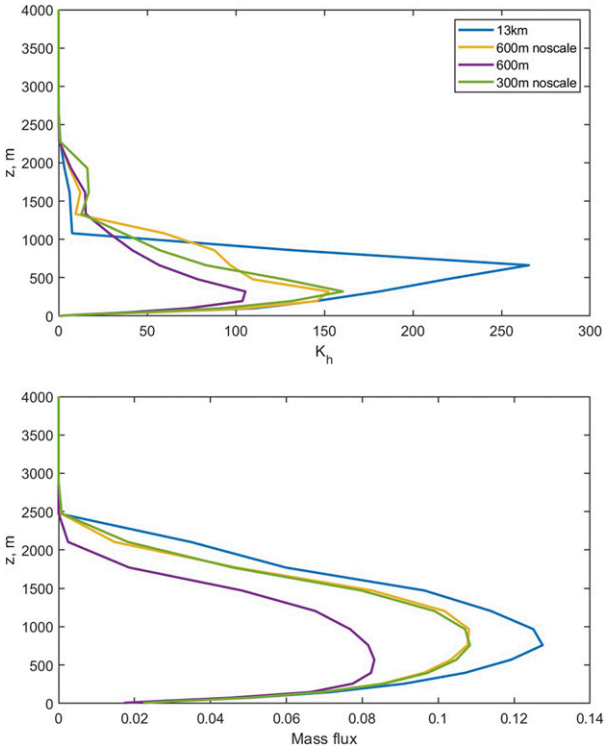

FIG. 10. (top) Eddy diffusivity and (bottom) mass flux for the 27 Jun 2015 case at 1400 CST. Runs are as shown in the legend.

we will reset the scale aware features to take effect only at smaller grid spacings, based on effective resolution. The practical limits are always under review and may be changed if they prove unnecessary.

Acknowledgments. This work was supported by the NOAA Atmospheric Science for Renewable Energy program.

Data availability statement. LASSO data are available at http://dx.doi.org/10.5439/1256454, and more information can be found at https://www.arm.gov/capabilities/modeling/lasso. The WRF SCM driving code, including scripts to convert forcings from LASSO, is available from https://esrl.noaa.gov/csd/groups/csd4/modeldata/. The version of the MYNN–EDMF scheme used here is available from the same URL. The provided code can be compiled with WRF, version 4.1. Modifications may need to be made to compile with earlier or later WRF versions.

REFERENCES

Angevine, W. M., J. Olson, J. S. Kenyon, W. I. Gustafson, S. Endo, K. Sušelj, and D. D. Turner, 2018: Shallow cumulus in WRF parameterizations evaluated against LASSO large-eddy simulations. *Mon. Wea. Rev.*, **146**, 4303–4322, https://doi.org/10.1175/MWR-D-18-0115.1.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *J. Climate*, **17**, 2493–2525, https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2.

ARM, 2018: Radiative flux analysis (RADFLUX1LONG). ARM Data Center, accessed 25 May 2018, https://doi.org/10.5439/1157585.

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.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Boutle, I., J. E. J. Eyre, and A. 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.

Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225, https://doi.org/10.1175/MWR-D-11-00046.1.

Ching, J., R. Rotunno, M. A. LeMone, A. Martilli, B. Kosovic, P. A. Jiménez, 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.

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

Glenn, I. B., G. Feingold, J. J. Gristey, and T. Yamaguchi, 2020: Quantification of the radiative effect of aerosol–cloud interactions in shallow continental cumulus clouds. *J. Atmos. Sci.*, **77**, 2905–2920, https://doi.org/10.1175/JAS-D-19-0269.1.

Gristey, J. J., G. Feingold, I. B. Glenn, K. S. Schmidt, and H. Chen, 2020: Surface solar irradiance in continental shallow cumulus fields: Observations and large-eddy simulation. *J. Atmos. Sci.*, **77**, 1065–1080, https://doi.org/10.1175/JAS-D-19-0261.1.

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

Hacker, J. P., and W. M. Angevine, 2013: Ensemble data assimilation to characterize surface-layer errors in numerical weather prediction models. *Mon. Wea. Rev.*, **141**, 1804–1821, https://doi.org/10.1175/MWR-D-12-00280.1.

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.

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.

Kurowski, M. J., and J. Teixeira, 2018: A scale-adaptive turbulent kinetic energy closure for the dry convective boundary layer. *J. Atmos. Sci.*, **75**, 675–690, https://doi.org/10.1175/JAS-D-16-0296.1.

Lancz, D., B. Szintai, and R. Honnert, 2018: Modification of a parameterization of shallow convection in the grey zone using a mesoscale model. *Bound.-Layer Meteor.*, **169**, 483–503, https://doi.org/10.1007/s10546-018-0375-1.

Lean, H. W., J. F. Barlow, and C. H. Halios, 2019: The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length

versions of the Met Office Unified Model. *Quart. J. Roy. Meteor. Soc.*, **145**, 1674–1689, https://doi.org/10.1002/qj.3519.

LeMone, M. A., 1973: The structure and dynamics of horizontal roll vortices in the planetary boundary layer. *J. Atmos. Sci.*, **30**, 1077–1091, https://doi.org/10.1175/1520-0469(1973)030<1077:TSADOH>2.0.CO;2.

Muñoz-Esparza, D., J. K. Lundquist, J. A. Sauer, B. Kosovic, 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.

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

——, and Coauthors, 2019b: Improving wind energy forecasting through numerical weather prediction model development. *Bull. Amer. Meteor. Soc.*, **100**, 2201–2220, https://doi.org/10.1175/BAMS-D-18-0040.1.

Palmer, T., 2019: The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Quart. J. Roy. Meteor. Soc.*, **145**, 12–24, https://doi.org/10.1002/qj.3383.

Panosetti, D., L. Schlemmer, and C. Schaer, 2019: Bulk and structural convergence at convection-resolving scales in real-case simulations of summertime most convection over land. *Quart. J. Roy. Meteor. Soc.*, **145**, 1427–1443, https://doi.org/10.1002/qj.3502.

Piotrowski, Z. P., P. K. Smolarkiewicz, S. P. Malinowski, and A. A. Wyszogrodzki, 2009: On numerical realizability of thermal convection. *J. Comput. Phys.*, **228**, 6268–6290, https://doi.org/10.1016/j.jcp.2009.05.023.

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.

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

——, and 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.

Sisterson, D. L., R. A. Peppler, T. S. Cress, P. J. Lamb, and D. D. Turner, 2016: The ARM Southern Great Plains (SGP) site. *The Atmospheric Radiation Measurement (ARM) Program: The First 20 Years, Meteor. Monogr.*, No. 57, Amer. Meteor. Soc., 6.1–6.14, https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0004.1.

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032, https://doi.org/10.1175/MWR2830.1.

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.

Xie, S., R. T. Cederwall, and M. Zhang, 2004: Developing long-term single-column model/cloud system–resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations. *J. Geophys. Res.*, **109**, D01104, https://doi.org/10.1029/2003JD004045.

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