Evaluation of Modeled Stratocumulus-Capped Boundary Layer Turbulence with Shipborne Data

TAKANOBU YAMAGUCHI
Cooperative Institute for Research in Environmental Sciences, University of Colorado, and NOAA/Earth System Research Laboratory, Boulder, Colorado

W. ALAN BREWER AND GRAHAM FEINGOLD
NOAA/Earth System Research Laboratory, Boulder, Colorado

(Manuscript received 11 February 2013, in final form 19 July 2013)

ABSTRACT

Numerically modeled turbulence simulated by the Advanced Research Weather Research and Forecasting Model (ARW) is evaluated with turbulence measurements from NOAA’s high-resolution Doppler lidar on the NOAA Research Vessel Ronald H. Brown during the Variability of the American Monsoon Systems (VAMOS) Ocean–Cloud–Atmosphere–Land Study—Regional Experiment (VOCALS-Rex) field program. A nonprecipitating nocturnal marine stratocumulus case is examined, and a nudging technique is applied to allow turbulence to spin up and come into a statistically stationary state with the initial observed cloud field. This “stationary” state is then used as the initial condition for the subsequent free-running simulation. The comparison shows that the modeled turbulence is consistently weaker than that observed. For the same resolution, the turbulence becomes stronger, especially for the horizontal component, as the length of the horizontal domain increases from 6.4 to 25.6 km. Analysis of the power spectral density shows that, even for the largest domain, the horizontal component of the turbulence is limited by the upper limit of the domain size; supporting evidence from past studies is provided. Results suggest that convergence is expected for (i) energy spectra of turbulence with a sufficiently large domain and (ii) liquid water path with an adequately large domain and fine resolution. Additional tests are performed by changing momentum advection and turning off subgrid-scale diffusion. These exhibit more significant changes in turbulence characteristics compared to the sensitivity to domain size and resolution, suggesting that the model behavior is essentially established by the configuration of the model dynamics and physics and that the simulation only gradually improves when domain size and resolution are increased.

1. Introduction

The southeastern Pacific atmospheric boundary layer has been the focus of intensive study over recent years because of its importance in Earth’s climate system (de Szoeke and Xie 2008). The region is characterized by strong ocean upwelling, cool sea surface temperatures (SSTs), and an extensive stratocumulus cloud sheet that reflects much more solar energy to space than the underlying ocean. The coastal region is also a strong source of anthropogenic particulate matter so there exists the potential for aerosol–cloud interactions that could modify cloud microphysics and precipitation formation, with further implications for cloud reflectance.

A recent field experiment, Variability of the American Monsoon Systems (VAMOS) Ocean–Cloud–Atmosphere–Land Study—Regional Experiment (VOCALS-Rex; Wood et al. 2011) deployed multiple aircraft and ships to study the coupled ocean–atmospheric boundary layer system. The experiment was designed to address the key science issues of ocean–atmosphere coupling and the role of the aerosol in modifying cloud microphysics, precipitation, and cloud forcing. During VOACLS-Rex, the National Oceanic and Atmospheric Administration (NOAA) Research Vessel (R/V) Ronald H. Brown (RHB) was outfitted with a suite of measurements to characterize, inter alia, the thermodynamics, dynamics, and microphysics of the cloudy marine boundary layer. The RHB spent a total of 63 days in the southeastern Pacific,
focusing primarily on latitude 20°S (±1°), between 70° and 85°W. It was equipped with a large number of in situ and remote sensing instruments to measure SST, surface fluxes, atmospheric state, aerosol size, composition and optical properties, and cloud and drizzle [for a full list of instruments, see Wood et al. (2011, their Table 3)]. Radiosondes were launched at 4-h intervals. The focus of this paper will be on the scanning, NOAA high-resolution Doppler lidar (HRDL; Tucker et al. 2009) that provided continuous measurements of the dynamics of the boundary layer through measurement of the radial wind velocity, as well as aerosol-backscatter signal intensity along the lidar’s line of sight (LOS). Together with some of the other key instruments such as the NOAA C-band precipitation radar (Comstock et al. 2004; Mechem et al. 2012); the microwave radiometer, which provided measurements of cloud liquid water path (LWP; Zuidema et al. 2005); and the NOAA W-band radar (Moran et al. 2012), the RHB cruise yields an unprecedented view of the dynamics of the cloudy marine boundary layer over a range of SSTs, boundary layer depths, aerosol conditions, and over many diurnal cycles. Key to this is the fact that the instruments sampled the boundary layer continuously and at high temporal frequency.

Models of different levels of complexity have been developed and applied to the stratocumulus-capped marine boundary layer. These include mixed-layer models (e.g., Lilly 1968; Schubert 1976), one-dimensional turbulence closure models (e.g., Mellor and Yamada 1974; Bougeault 1985; Chen and Cotton 1987), and large-eddy simulation (LES; e.g., Deardorff 1972; Moeng 1986; Mason 1989). LES is probably the most appropriate numerical tool to address processes associated with the turbulent cloudy boundary layer but the computational challenge is to do so with fine-enough resolution and large-enough domain size. For small horizontal domains on the order of $3 \times 3$ km$^2$, horizontal grid spacing $\Delta x$ of 5 m is feasible (Yamaguchi and Randall 2012, hereafter YR12). Advances in computing power have made it possible to perform LES of a stratocumulus-topped boundary layer on a $23 \times 23$ km$^2$ domain with $\Delta x = 12$ m (Yamaguchi and Feingold 2013). Vertical grid spacing $\Delta z$ for stratocumulus LES is generally less than or equal to $\Delta x$. For long term and/or large domain simulations, the “near LES” or very-LES framework (Mechem et al. 2012) is an economical although deficient alternative, but has yielded important results in a number of studies (Wang and Feingold 2009a,b; Feingold et al. 2010; Kazil et al. 2011; Solomon et al. 2011; Soloman et al. 2011; Blossey et al. 2013) and therefore it behooves the community to evaluate its performance, not only against other LES models, but equally importantly, against observations. ARW has been tested on several LES intercomparison cases of the Global Atmospheric System Studies (GASS, formerly GCSS) (Stevens et al. 2005, hereafter S05; Ackerman et al. 2009; vanZanten et al. 2011) and the results have been shown to compare reasonably well with other models (Wang et al. 2009; Yamaguchi and Feingold 2012, hereafter YF12). Comparison with observations is thus timely.

The robustness of LES-generated turbulence can be assessed within the modeling framework alone (e.g., Matheou et al. 2011), which is an essential step, especially when testing newly developed model and/or numerical schemes. Another approach is model evaluation with well-defined turbulence measurements such as those made by lidar (e.g., Mayor et al. 2003; Lenschow et al. 2012). In the current study, the dataset produced
during VOCALS-REx represents significant opportunity for model comparison with observations.

The objectives of this study can be described as 1) simulation of a nonprecipitating stratocumulus case with the ARW based on ship measurements during VOCALS-REx, 2) development of a framework for comparison between model and observation, 3) evaluation of model-simulated turbulence with lidar observations, and 4) evaluation of the influence of the ARW configuration on the quality of the comparison. For the modeled turbulence evaluation, we focus primarily on Doppler lidar observations of (i) \( w \) statistics, including \( w^2 \) and \( w^3 \), where the overbar (prime) denotes mean (perturbation), and (ii) the horizontal velocity \([V_H = (u, v)]\), including turbulence kinetic energy (TKE) and relative contributions from \( w^2 \) and \( V_H^2 \). The higher-order statistics yield important information on the physical mechanism driving the dynamics. For instance, \( w^3 < 0 \) near cloud top suggests a radiatively driven “top down” circulation, while \( w^3 > 0 \) near the surface indicates a turbulence driven by surface warming. Because of the limited in-cloud penetration depth of the lidar, the focus is on subcloud sampling.

To accomplish these objectives, we will present results from one LES and three near LESs. Three square domains with horizontal length \( L \) of 6.4, 12.8, and 25.6 km are used. For LES of stratocumulus, ARW tends to be more expensive compared to large-eddy models formulated with the anelastic system (YF12). For the worst case, ARW is about 20 times more expensive than an anelastic counterpart. We use \( \Delta x = 100 \text{ m} \) for our near LES, which is finer than past near LES \( (\Delta x \approx 125 \text{ m}) \). Because of the desire to use a \( \Delta z \) that is as small as possible to reduce overentrainment (e.g., Bretherton et al. 1999b; S05), the grid aspect ratio \( (\Delta x/\Delta z) \) is sacrificed and set to 10—a value that has often been used in past LES studies (e.g., Ackerman et al. 2009). Potential consequences include less accurate representation of entrainment and subgrid-scale (SGS) TKE, which would best be represented with a high-resolution isotropic grid (i.e., \( \Delta x/\Delta z = 1 \)). To minimize simulation bias, \( \Delta x/\Delta z \) is fixed for all simulations (Cheng et al. 2010; YR12). Sensitivities with various \( \Delta x/\Delta z \) are an interesting topic, especially in the near-LES framework (e.g., Wang and Feingold 2009a); however, we defer the subject for future work.

Comparison between model and observation is not an easy task. A few of the key issues are outlined below. (i) While turbulence exists in nature at the simulation start time, LES is traditionally initialized in a quiescent state with pseudorandom perturbation to drive turbulence. It is well known that the LWP decreases in response to the first large-eddy turnover (e.g., Moeng et al. 1996; S05) so that the state when turbulence is fully spun up may have already diverged from the observed state. (ii) Frequently horizontal average LES quantities are used for comparison to observations, whereas the ship-based measurement is a two-dimensional “curtain” that advects in time owing to the relative speed of the ship and the mean wind, so that the statistics are derived based on a temporal analysis. Directly comparing these two different statistics has the potential to be misleading and will be investigated here. (iii) Although our objective is not a forecast of the system evolution, a subset of the model parameters should compare well to observations for turbulence comparison to be meaningful. Reasonable adjustments to the initial soundings and forcings have to be carefully considered. Keeping simplicity in mind, we develop a methodology for model comparison with observation, especially for turbulence, that we believe will be broadly beneficial for other model–observational comparisons.

The outline of this paper is as follows. The lidar and data characteristics are described in the next section. A brief description of ARW and our configuration of the dynamical core and physics is summarized in section 3. A case description of a nonprecipitating nocturnal stratocumulus based on ship measurements during VOCALS-REx is discussed in section 4. The details of our framework for model comparison with observation are presented in section 5. The model evaluation is carried out in section 6. We discuss additional sensitivity tests in section 7. A summary is given in section 8.

2. Lidar observation and data

NOAA’s HRDL was deployed on the RHB during VOCALS-REx to monitor \( V_H \), turbulence, and aerosol spatial distribution. HRDL utilizes a pulsed, 2-\( \mu \text{m} \), solid-state laser operated with a repetition frequency of 200 Hz. Multiple atmospheric returns were combined to measure the LOS wind speed and aerosol-backscatter signal strength twice per second with an along-beam resolution of 30 m. Depending on aerosol concentration, HRDL typically made these measurements horizontally out to distances of 6–8 km and vertically through cloud base.

The pointing of the lidar was controlled with a motion-stabilized hemispheric scanner. During its normal operation, HRDL performed both azimuthal and elevation scans and stared vertically. When operated from a moving platform, the pointing of the system was actively maintained in the world frame to within 0.5° of the predefined scan parameters. This allowed for lidar measurements within 1° of elevation off the ocean surface or stabilized at zenith under moderate sea-state conditions. Information about the ship’s translational and rotational

Unauthenticated | Downloaded 03/01/24 11:36 PM UTC
motion, supplied by the compensation system, was used to estimate and remove the effect of the platform motion from the LOS wind speed measurements. The precision of the LOS wind speed measurements depended in part on the strength of the atmospheric return but was typically 10–25 cm s\(^{-1}\) for the conditions encountered during VOCALS-REx.

HRDL operated continuously during the two research legs of the VOCALS-REx cruise, employing a 20-min repeating scan sequence that consisted of 10 min of azimuthal and elevation scans followed by 10 min of vertically staring measurements. The scanning data were combined to estimate profiles of \(\mathbf{V}_H\) and \(\mathbf{V}_i^2\) from within 15 m of the ocean surface through the depth of the subcloud layer and into the cloud as long as the lidar signal was not appreciably attenuated. The vertically staring data were used to estimate the \(w\) statistics from within 170 m of the ocean surface through the depth of the subcloud layer and up to the cloud. The quality of the statistics depends on the number of points sampled during the 10-min averaging period, both of which decrease in the vicinity of cloud base. This tends to result in a decrease in \(\bar{w}^2\) (e.g., Fig. 6). We have applied a sample number threshold that results in maximum lidar profile heights that are just above the lidar-detected cloud base.

From every 20-min sequence of data, single estimates of these profiles were calculated. These profiles form the basis for model evaluation. A caveat of the measured statistics is that they represent the variation over spatial scales up to approximately 3–4 km owing to the 10-min averaging period for \(w\) and the horizontal distance between the closest and farthest points at each level converted from the LOS for \(\mathbf{V}_H\).

3. Model description

The ARW, version 3.3.1, with the LES and statistics packages developed by YF12 is used for this study. ARW is a nonhydrostatic model, which solves the compressible system of equations in a flux form governed by a mass vertical coordinate. The basic prognostic variables are the zonal, meridional, and vertical velocities \((u, v, w)\); dry air mass; geopotential; and potential temperature \(\theta\). Depending on the microphysics and SGS turbulence schemes, mass mixing ratios, number concentrations (per unit mass of air), and SGS TKE are also predicted.

A time-split integration scheme, which separates the high-frequency acoustic and gravity waves from the low-frequency physical mode, is used for temporal discretization (Klemp et al. 2007). ARW uses a third-order Runge–Kutta scheme, and during each Runge–Kutta step, the horizontally propagating high-frequency mode is integrated with a forward–backward scheme with the acoustic time step, while an implicit scheme is used for the vertically propagating high-frequency mode. YF12 reported that simulation results exhibit a dependency on the acoustic Courant number associated with the acoustic time step. Additional discussion for the sensitivity on the acoustic time step is presented in appendix B.

The staggered Arakawa C grid is employed for spatial discretization. Advection is computed with a scheme described in Wicker and Skamarock (2002). As recommended in the ARW user’s guide, we select the fifth-order scheme for the horizontal advection and the third-order scheme for the vertical advection. For scalar advection, except \(\theta\), the monotonic limiter described in Wang et al. (2009) is used.

The following physical parameterizations are selected among the available schemes. A 1.5-order TKE closure (Klemp and Wilhelmson 1978; Deardorff 1980) is used as an SGS turbulence parameterization. Longwave radiative flux is computed with the Rapid Radiative Transfer Model (RRTMG) scheme (Mlawer et al. 1997) and updated every 30 s. This update period is appropriate for nocturnal stratocumulus simulations with \(\Delta x \approx 50\) m based on the sensitivity tests performed by YR12. This period is kept the same for all simulations performed in this study. The fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5) scheme based on similarity theory (Monin and Obukhov 1954) is used for the simulations that interactively compute surface friction velocity and sensible and latent heat fluxes. The microphysics parameterization is a bulk two-moment scheme (Feingold et al. 1998). A particular attribute of the scheme is that it uses a hybrid bin–bulk approach that attempts to retain the accuracy of a bin scheme without tracking a large number of prognostic scalars. The scheme predicts mixing ratio of water vapor \(q_v\), cloud water \(q_c\), and rainwater and number concentrations of aerosol \(n_a\), cloud droplets \(n_d\), and raindrops. It also predicts cloud supersaturation following the semianalytical approach of Clark (1973), and activates the ambient aerosol accordingly.

4. Nonprecipitating nocturnal stratocumulus case

a. Physical setup

Examining the sonde and lidar data throughout the period of the RIIH deployment in the VOCALS-REx region, we sought a nocturnal, well-mixed boundary layer structure (Lilly 1968) without precipitation, and selected 18 November 2008, when the ship was located at 19°S, 80°W. The sonde launch at 0336 UTC (2216 local time) was selected as the initialization time. The duration
of simulations is 6 h and 24 min so that it finishes at 1000 UTC (0430 local time before sunrise). During this period, the ship travels 0.3 W and stays on 19°S, which translates to an average ship velocity of $V_H = (-1, 0) \text{ m s}^{-1}$.

The observed and initial model soundings are shown in Fig. 1. A detailed description of the procedure to create the input profiles and data is given in appendix A. The figure also shows the $V_H$ profiles at the subsequent sonde launch times at 0814 and 1217 UTC, which are used to update the background wind (i.e., as a geostrophic wind) with a linear time interpolation. The $V_H$ profiles exhibit significant shear near the PBL top. This was a commonly observed feature during VOCALS-REx and including it is important for entrainment at the PBL top (Wang et al. 2012) and turbulence through shear production of TKE inside the PBL. Another commonly observed feature, apparent in Fig. 1, is that the liquid water potential temperature $\theta_l$ and total water mixing ratio ($q_t = q_o + q_c$) are usually slightly less mixed.

For the large-scale subsidence, a constant divergence of $D = 1.7 \times 10^{-6} \text{ s}^{-1}$ is applied at all levels. Divergence is estimated from the daily mean of the average $D$ between 1.3- and 2-km levels (free atmosphere) of the European Centre for Medium-Range Weather Forecasts (ECMWF) operational data. Lower boundary conditions (i.e., SST, sensible and latent heat fluxes, and friction velocity) are derived from the hourly averaged surface measurements (de Szoeke et al. 2010). These parameters are linearly time interpolated to the current time. To represent synoptic-scale motion, the Coriolis force at 19°S is applied. Although the simulations are performed in an Eulerian framework, we choose not to apply large-scale advective tendencies for $\theta$ and $q_v$ because they are not measured and their implementation would entail assumptions.

The two-moment microphysics scheme (Feingold et al. 1998) requires representation of the aerosol size distribution. It assumes a lognormal aerosol size distribution, which is a good approximation to the observed aerosol size distribution (Fig. 2). Analysis of the aerosol data shows that for the specific time period under consideration, the ratio of the two lognormal modal
concentrations \( [N_1/(N_1 + N_2)] \) is time independent. Similarly, the geometric-mean diameter \( d_g \) and geometric standard deviation \( \sigma_g \) are approximately constant for each of the aerosol size modes. Therefore, for these simulations, the summation of the aerosol number concentration \( (n_a = N_1 + N_2) \) is a prognostic variable, and each number concentration is partitioned with the constant ratio to compute activation of aerosol for each distribution. Hourly averaged values between 0300 and 0400 UTC are used for all of the aerosol related parameters; for the distribution of the smaller particles, \( N_1 = 59 \text{ mg}^{-1} \), \( d_{g1} = 0.05 \mu \text{m} \), and \( \sigma_{g1} = 1.44 \); for the distribution of the larger particles, \( N_2 = 156 \text{ mg}^{-1} \), \( d_{g2} = 0.18 \mu \text{m} \), and \( \sigma_{g2} = 1.42 \). The geometric standard deviation of both the cloud and rainwater size distributions is 1.2, which is the default value specified in the scheme. No aerosol sources are applied because of the short duration of the simulation, and because aerosol removal by precipitation, although treated by the microphysical scheme, is negligible. Although this is a nonprecipitating stratocumulus case, collision coalescence and drop sedimentation are allowed after the first hour. The domain mean rainwater path for all simulations never exceeds 0.001 \( \text{g m}^{-2} \).

b. Set of simulations

All simulations use a square domain with periodic lateral boundary condition and a domain depth of 2 km. The following set of simulations are performed and listed in Table 1: 6B, our base case, uses \( L = 6.4 \text{ km} \), \( \Delta x = 100 \text{ m} \), and 200 levels \( (\Delta z \approx 10 \text{ m}) \); 12B (25B) uses \( L = 12.8 \text{ (25.6) km} \) and the same grid as 6B; 6H uses \( L = 6.4 \text{ km} \), \( \Delta x = 50 \text{ m} \), and 400 levels \( (\Delta z \approx 5 \text{ m}) \). Our analysis of PSD for \( w \) (e.g., Fig. 10) with 6H shows that the wavelength of \( \lambda = 400 \text{ m} \) (i.e., \( 4\Delta x \) for \( \Delta x = 100 \text{ m} \)) is approximately equal to the largest \( \lambda \) of the inertial subrange. Based on the requirement that \( 6\Delta x \) is much less than inertial subrange (Bryan et al. 2003), \( \Delta x = 100 \text{ m} \) cannot be classified as LES for the current case; thus, 6B, 12B, and 25B are near LES.

Strictly speaking, even the grid spacings of 6H may not be an LES because the major premise of LES is that the grid spacing is sufficiently small to fit well in the inertial subrange, and the simulation is relatively insensitive to the SGS parameterization (e.g., Moeng 1984: Bryan et al. 2003). Convergence of results with finer grids is of interest (e.g., Matheou et al. 2011) but beyond the scope of this study.

Following YF12, the physical time step \( \Delta t \) and the number of substeps for the acoustic mode \( N_{AT} \) are determined with an acoustic Courant number convergence test (appendix B); values are listed in Table 1. This convergence test is performed for 6B and 6H. As in YF12, we emphasize the importance of checking profiles and time series for convergence.

Throughout the study, model analyses are performed based on horizontal domain-averaged statistics recorded every minute and three-dimensional snapshot data saved every 10 min. The variables \( \nabla^2 \phi \), \( w^2 \), and \( \overline{w^3} \) presented in the figures are all resolved-scale quantities, and TKE includes the SGS TKE.

Table 1. List of the named simulations.

<table>
<thead>
<tr>
<th>Name</th>
<th>L (km)</th>
<th>( \Delta x ) (m)</th>
<th>( \Delta z ) (m)</th>
<th>( \Delta t ) (s)</th>
<th>( N_{AT} )</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B</td>
<td>6.4</td>
<td>100</td>
<td>-10</td>
<td>0.2</td>
<td>8</td>
<td>6B-2 modification in sections 6 and 7</td>
</tr>
<tr>
<td>12B</td>
<td>12.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6B-2 modification in section 6</td>
</tr>
<tr>
<td>25B</td>
<td>25.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6B-2 modification in section 6</td>
</tr>
<tr>
<td>6H</td>
<td>6.4</td>
<td>50</td>
<td>-5</td>
<td>0.2</td>
<td>12</td>
<td>6B-2 modification in section 6</td>
</tr>
<tr>
<td>6B-1</td>
<td></td>
<td>As in 6B.</td>
<td></td>
<td></td>
<td></td>
<td>( D = 5 \times 10^{-9} \text{ s}^{-1} ) for free run</td>
</tr>
<tr>
<td>6B-2</td>
<td></td>
<td>As in 6B.</td>
<td></td>
<td></td>
<td></td>
<td>( D = 5 \times 10^{-9} \text{ s}^{-1} ) and interactive surface calculation for free run</td>
</tr>
</tbody>
</table>
5. Framework for a model comparison with observation

a. Initial turbulence in quasi steady state with initial fields

Traditionally, LES is initialized with an observed cloudy boundary layer sounding but the model is in a quiescent state. Here, special effort is made to allow the model to generate initial turbulence that is in a statistically stationary state based on the initial cloudy profile. Establishing the turbulence associated with the initial observed cloud field is essential in order to compare with the time-evolving lidar-observed turbulence.

Our strategy is to apply a nudging technique to maintain the horizontal mean of $u, v, \theta, q_t$, and total number concentration $(n_i = n_a + n_c)$ close to the initial profile while turbulence develops. The nudging tendency is described in appendix C. A similar nudging method has been applied to cloud-system-resolving simulations for deep convection (e.g., Henderson and Pincus 2009). We call this phase the “spinup” run. During the spinup run, we allow turbulence and cloud to adjust and come into balance with one another; subsequently, this statistically stationary state becomes the initial state for further freely evolving simulation (i.e., the “free run,” which, as the name implies, is free of nudging) with the fields evolving according to the unconstrained model equations. For the spinup run, the surface forcing is held constant at 0336 UTC.

A suite of experiments based on 6B shows that the existing PBL-top wind shear evaporates cloud, and that the effect is large in the wake of the first large eddy. This point was noted in Wang et al. (2012) for the Coupled Ocean–Atmosphere Mesoscale Prediction System–LES model (Golaz et al. 2005). In our experiments, depending on the input profiles, the effect is so large that nudging cannot prevent the fields from diverging without a very short nudging time scale $\tau$ (i.e., large nudging tendency). Cloud evaporation also occurs for shear-free simulations (e.g., S05) and we observed that the effect is minimum for a shear-free initial condition, as also noted by Wang et al. (2012), who nudge $V_H$ toward the shear profile starting from the shear-free profile during the first hour in order to avoid cloud breakup. We take a different path, and let the wind shear consume cloud water. With the well-mixed assumption in the cloud layer, our initial sounding produces a somewhat larger LWP (95 g m$^{-2}$) than that based on the adiabatic assumption and a ceilometer measurement of cloud base ($\sim$60 g m$^{-2}$). The larger LWP is mitigated by the shear-generated evaporation so that the LWP arrives at approximately the observed value at the end of the spinup run.

Through multiple tests we learned that 1) $\tau$ should be as large as possible; shear-induced cloud evaporation can be prevented with a very short $\tau$—however, the free-run field evolves in a way that the cloud cannot sustain itself without nudging; 2) subsidence allows one to increase $\tau$ by suppressing the PBL top; 3) $q_t$ nudging, as a part of the $q_i$ nudging, is useful since it fills in evaporated cloud where the $q_i$ nudging is very small owing to saturation; and 4) the $q_i$ nudging imposes an upper limit of condensation, whereas nudging only $q_i$ can cause immediate condensation of $q_c$ and results in undesirable growth of $q_c$, especially for small $\tau$.

For the current case, $\tau = 1$ h is short enough to maintain the horizontal mean profile close to the initial soundings and long enough to prevent cloud evaporation after nudging is turned off. We should note that $\tau = 1$ h is optimal for the described configurations (both dynamical core and physics for ARW) and there is no guarantee that the time scale works for other configurations and/or models to produce the observed LWP at the end of the spinup run.

The duration of the spinup run is 4 h. This might not be long enough for 25B ($L = 25.6$ km) to develop meso-scale circulation; the field does, however, reach a steady balanced state within this 4-h period. Savic-Jovcic and Stevens (2008) used a 6-h duration for a square domain with $L = 25.6$ km based on the meso-$\gamma$ scale simulations evident in the domain. Since our subsequent free run has a 6-h duration, the total duration (10 h) is longer than the aforementioned study, and mesoscale circulations are expected to fully develop in the early stages of the free run.

The results of the spinup run for 6B are presented in Fig. 3. The other three spinup simulations (6H, 12B, 25B) show similar results to 6B. There is no notable systematic difference among results for the domain size and resolution tested. All of the ARW profiles are shown with a Tukey box plot for the last hour; the shadings are, however, not distinguishable owing to the narrow distribution of the first-moment quantities, suggesting that the fields are steady as a result of the nudging. The LWP settles down around the observed value ($\sim$60 g m$^{-2}$). The hourly mean model-derived cloud base (defined at the level of $q_c = 0.01$ g kg$^{-1}$) compares well to the lidar estimate (Fig. 3b). Particular differences from the initial conditions are the weaker and more mixed $V_H$ in the cloud layer and the thicker inversion (Fig. 3a). The thickness of the inversion layer grows to approximately 100 m from the initial thickness of 35 m. Aside from the effect of resolution and/or SGS treatment on broadening the simulated inversion layer, PBL-top wind shear is probably responsible for this thickening of the inversion (Wang et al. 2012). A sensitivity test with stronger subsidence results in only marginal lowering of the inversion top and the thickness does not change. Stronger subsidence also results in smaller LWP.
b. Methodology for comparing ARW and lidar

The turbulence in the spinup run is in a quasi steady state due to the steady mean profiles, as expected (Fig. 3b). The lidar measurement over 1 h exhibits large temporal variability while ARW does not. This suggests that the comparison between spatial average (ARW) and temporal average along a line (lidar) requires special consideration.

At a given height, the lidar samples a two-dimensional line as it moves over the ocean surface (or the atmosphere advects over it). The model output is typically an...
average of the field over the entire domain for the sampling time in question. To address this disparity, we sample the model domain as if it were being viewed by the lidar. To do so we employ the LagrangianParcel Tracking Model (LPTM; YR12) within ARW to perform Lagrangian column tracking, which mimics the ship’s sampling of the atmosphere. LPTM places massless parcels (or particles) in the domain, then predicts each parcel trajectory using the spatially interpolated velocity from the host model’s resolved-scale velocity field, and performs the “measurement” with spatial interpolation. To represent the ship measurement, LPTM is modified so that each parcel travels with a fixed $\mathbf{V}_H$ without vertical motion.

The Lagrangian column tracking is performed from 0500 and 0900 UTC during the free run of 6B-2 (a variation on 6B discussed later). Four Lagrangian columns are arbitrarily placed inside the domain, and each column travels with the mean ship velocity [i.e., $\mathbf{V}_H = (-1.0, 0) \text{ m s}^{-1}$]. One column consists of 256 parcels, equally placed in the vertical below 1600 m. Velocity components are saved every second. During the 4-h period, each column travels 14.4 km. The spatial coverage based on the 4-h-mean wind (approximately $-3.5 \text{ m s}^{-1}$) and the mean ship velocity is 36 km in the east–west direction, which is shorter than the spatial coverage based on the observed $u$ owing to the deceleration of $u$ during the free run (discussed later). Additionally, two other cases with different ship velocities [$\mathbf{V}_H = (0, 1)$ and $\mathbf{V}_H = (-0.7, -0.7) \text{ m s}^{-1}$] are performed. There is no significant difference among the three cases for the results discussed below.

The mean and variance profiles are derived in accord with the lidar measurements: 4-h-mean $u$ and $v$ are removed as a background flow. The 4-h-mean $w$ is set to zero. Profiles for mean and variance are then computed for every 10-min segment using the perturbation from the 4-h mean. Since lidar data are available for a period of 10 min, every 20 min, statistics for every other 10-min segment are discarded. Based on the profiles every 20 min, additional statistics are prepared with 40-, 60-, and 80-min means (i.e., averaging two, three, and four profiles).

The temporal variability for the different averaging time scales for one Lagrangian column with ARW’s spatial-mean profiles for 4 h is shown in Fig. 4. For the 10-min time scale, LPTM shows a wide temporal variability, much like the lidar observation, but that variability becomes narrower as the averaging time scale is increased. The variability in the 1-h mean of the LPTM and the variability in the horizontal mean of ARW over the 4-h period are similar in magnitude. Based on this exercise, the comparison for free runs will be performed with the 1-h mean of the lidar profiles (i.e., mean of three 10-min-mean profiles available every 20 min) and the spatial mean of ARW over the last 4 h of the simulation period.

c. Adjustment for the free run

The time evolution of the LWP and inversion-top and -base heights for 6B with the original configuration is shown in Fig. 5a along with two additional cases (6B-1, 6B-2) listed in Table 1 and described below. Case 6B evolves in a way that the inversion-top and -base heights continuously elevate with a steady inversion thickness while the observed cloud-top height is somewhat steady. The LWP of 6B is approximately comparable to the observation. It should be noted that the shading for the LWP represents one standard deviation, computed at each given time; thus, there is significant spatial variability (Fig. 7). The relatively large observed LWP between 0400 and 0600 UTC exists in the model domain (e.g., Fig. 7b). The time series for ARW is calculated as a domain mean and therefore the smaller-scale variability that might be encountered by the ship along its narrow track is smoothed out.

Given the disparity between observed and modeled PBL height, in 6B-1 we applied increased large-scale divergence to alleviate the problem. After experimentation, $D = 5 \times 10^{-6} \text{ s}^{-1}$ was found to produce a steady PBL height. This is approximately 3 times stronger than the originally specified value from the ECMWF operational data ($D = 1.7 \times 10^{-6} \text{ s}^{-1}$). Although the ECMWF-derived $D$ is relatively uncertain, it is quite possible that ARW overestimates the amount influenced by the grid size and $\Delta x / \Delta z$. A rigorous study would be required to examine this point for ARW, and this is not our focus here. Instead the motivation is to apply a set of conditions that mimics a cloud and PBL of reasonable depth while exploring the extent to which this simulation can reproduce the observed wind field (discussed further below). To this end, we use $D$ as a tuning parameter.

Motivated by a desire to let the model compute surface forcing in an internally consistent way, 6B-2 is performed to investigate the effect of the interactive surface calculation, instead of specifying the friction velocity and surface fluxes. The computed surface forcing parameters are smaller than the observations (Fig. 5b). Interestingly, however, 6B-2 produces a smoother LWP time series, a steadier PBL height (Fig. 5a), a slightly stronger $\mathbf{V}_H$, and improved $\overline{w^2}$ and $\overline{w^3}$ (discussed further below). As demonstrated by YR12 with the System for Atmospheric Modeling (Khairoutdinov and Randall 2003), a decrease in the grid aspect ratio would result in higher surface fluxes and, in combination with a smaller $\Delta x$, one might expect further improvements in model performance.

The vertical profiles for 6B, 6B-1, and 6B-2 are shown in Fig. 6. For all cases $u$ is weak compared with the lidar velocity.
measurement, likely because of the absence of treatment of synoptic-scale flow such as the Chilean coastal low-level jet. A comparison between $6H$ (LES) and that with $V_H$ nudging ($\tau = 1$ h) suggests that the difference for $V_H$ is rather weak (not shown). Caveats of this test are (i) a shorter $\tau$ is necessary to exactly match the observed wind with shear in the PBL, which affects TKE generation; (ii) the observed winds are 4 h apart so that the time-interpolated profile is probably not appropriate to use as a mean wind at a given time; (iii) the $V_H$ nudging reduces LWP owing to increased PBL-top shear so that further experimentation is necessary; and (iv) the treatment of the low-level jet should include the large-scale advection of heat and moisture, and these are

![FIG. 4. Temporal variation of the spatial average of ARW and temporal average of the Lagrangian column tracking over the 4-h simulation. The shading covers the minimum and maximum values. The figure uses transparent shadings so that the overlapped area is a mixture of the two colors.](image)
unavailable. Because of these complications and difficulties, we elect not to apply any treatment for the low-level jet.

The simulations all exhibit weaker turbulence than the observations. As expected from Fig. 5b, the turbulence for 6B-2 becomes more top-down oriented (i.e., radiatively driven). Overall, 6B-2 matches the observation better than 6B-1. The negatively skewed profile in $w^3$ for 6B-2 is comparable to the observation, even though $\bar{w}^2$ is underestimated. This indicates that the range of the downdrafts in the model is narrower than the observation, whereas that of the updrafts is much narrower since in order to have similar $w^3$ with smaller $\bar{w}^2$, the skewness has to be larger. Guo et al. (2008), using the Goddard Cumulus Ensemble model (Tao et al. 2003), also showed that a grid of $\Delta z > 10$ m and $\Delta x = 25$ m results in a narrower $w$ distribution compared to aircraft observations of northeastern Pacific stratocumulus. One
can expect a similar smoothing influence of $\Delta x$ on $w$. Since our $\Delta x$ is larger than 25 m, the smearing effect on turbulence is expected to be strong.

On balance, we concluded that the benefits of the interactive surface calculation outweigh the discrepancies between computed and observed values. For this reason, we adopt the modification made for 6B-2 (i.e., increased $D$ and interactive surface calculation) for all free runs discussed below (sections 6 and 7).

6. Model comparison with observation for free runs

a. Results of the free runs

A snapshot of LWP for 25B at the end of the free run (Fig. 7a) shows the appearance of several closed cellular circulations larger than a few kilometers in diameter. These cells are larger than the smaller domains tested. The normalized occurrence of LWP (Fig. 7b), created from the last 6 h of data saved every 10 min, emphasizes the effect of domain size and resolution on the LWP field. With increasing domain size (6B to 12B to 25B), there is a broadening in the distribution and an appearance of larger LWPs. This is also true for the finer resolution (i.e., 6B to 6H) and is at least partially due to an increase in the number of cloudy columns simulated, since the number of horizontal grid points is the same for 6H and 12B. For the larger domain, the broadening is probably related to the appearance of the near-mesoscale closed cellular circulation. The significant overlap between 12B and 25B indicates a tendency to convergence for the larger domains (at fixed resolution). Alternatively, the distribution may achieve convergence at higher resolution with a fixed $L$.

The time series of LWP and vertically integrated TKE, $\nabla^2 T$, and $w^2$ are shown in Fig. 8 for all cases. Most of the measured LWP values are covered within the one
standard deviation range for all cases. For the larger domain, there is a large increase in TKE owing to an increase in $V_0^2 H$; $w_0^2$ is similar in magnitude among all the cases (Fig. 8b). This consequence should be expected based on DDJ04. There is also a strong tie between LWP and TKE, which can easily be seen for 12B; for example, when LWP decreases, there are concomitant decreases in TKE and $V_0^2 H$.

The profiles for lidar comparison with the free runs are shown in Fig. 9. The $V_H$ components are not sensitive to the tested resolution and domain sizes (Fig. 9). Interestingly, with increasing $L$, the cloud-base height becomes lower and closer to the lidar estimate. As $L$ increases, $V_H^2$ strengthens inside the mixed layer, especially in the subcloud layer below 800 m. There is good agreement between 25B and the observations below 500 m. A cautionary note: since the spatial coverage for the lidar measurement is 3–4 km while $L = 25.6$ km, this agreement is not necessarily a satisfactory one (see discussion below). The $w$ statistics indicate stronger and more negatively skewed turbulence as the domain becomes larger, but the effect is not strong owing to the smoothing effect of $\Delta x$ on the $w$ distribution (Guo et al. 2008). The temporal variation of the $w$ statistics for 25B is much narrower than for the other cases over the last 4 h (Fig. 8b). This implies that the domain size is perhaps close to large enough so that the mesoscale cellular circulation is nearly free of constraint from surrounding cells.

b. PSD analysis

Following DDJ04, we performed PSD analysis for modeled $u$, $v$, $w$, and $q$, with the procedure described in appendix D. We found that the PSD for $u$, $v$, and $q$, share the same features, thus only the spectra for $w$ and $u$ are presented and discussed here. The PSD of the lidar-observed $w$ is prepared as follows: To estimate the horizontal spatial scale of the variations in $w$ during a single 10-min observation period, 1/2-s-resolution time series at each 30-m height bin were used to calculate the PSD as a function of height. An estimate of $|V_H|$ at each height was used to transform from frequency to wavenumber. The 4-h-mean PSD at selected levels for 6H and 25B are shown in Fig. 10a as a function of $l$. For $w$, the PSD of the lidar data is also shown. The model spectra above the PBL (1500 and 1800 m) are one or two orders of magnitude smaller than those in the PBL. The shapes of the spectra are, however, similar at all levels. Since there is no turbulence in the free atmosphere, the power should have its source in the boundary layer turbulence—for example, via the upward transport of energy by gravity waves (Lilly 1983). For the small $l$, all slopes are steeper than $-5/3$, which is a well-known characteristic of finite difference models (e.g., Bryan et al. 2003). In Fig. 10a, the model and lidar PSDs seem to overlap at some $l$; however, a closer look (Fig. 10c) clearly shows that the magnitude is significantly smaller in the model, which contributes to the underestimation of the model variance.

An interesting feature is that for both $u$ and $w$, 6B, 12B, and 25B ($\Delta x = 100$ m) have generally larger power than 6H ($\Delta x = 50$ m) for $l \approx 600$ m, while the opposite is true for the shorter $l$ (Fig. 10c). Apparently, compared with the observed $w$ spectrum, the low-resolution runs represent turbulence better than 6H for $l \approx 600$ m while the opposite seems true for $l < 600$ m. In addition, the power contributed by the tail at larger $l$ tends to be largest for 25B and smallest for 6B (and 6H). The feature is also observed for the PBL-averaged 4-h-mean PSD (not shown); thus, the difference in the PSDs comes from the difference in domain size and resolution. As
seen in Fig. 9, \( \nu^{2} \) and \( w^{2} \) are similar for 6B and 6H. Case 6H distributes energy at smaller scales and the power is low at \( \lambda \geq 600 \text{ m} \); in contrast, the low-resolution simulations are characterized by power at \( \lambda \geq 600 \text{ m} \) (6Dx). Because of the reduced degrees of freedom in terms of resolvable small scales, 6B accomplishes comparable variance to 6H by generating more energy at \( \lambda \leq 600 \text{ m} \).

Cheng et al. (2010) presented similar results for a set of shallow cumulus simulations and suggested an artificial effect on turbulence that results in modeled clouds becoming more stratiform-like with coarser \( \Delta x \). Eventually, one might expect convergence in \( u \) and \( w \) for an adequately large domain size with the same resolution.

In Fig. 10 there is a notable difference in PSD shape between \( w \) and \( u \) at the larger \( \lambda \); the power of \( w \) becomes weak for \( \lambda \approx 1 \text{ km} \), whereas the power of \( u \) continuously strengthens as \( \lambda \) increases. In the case of \( u \), the slope of the spectra in Fig. 9 for \( \lambda \) is not quite, but close to, \(-\frac{5}{3}\). For the boundary layer, the energy gap between boundary layer turbulence and mesoscale circulation was first observed by Van der Hoven (1957) and Fiedler and Panofsky (1970). Later, enough evidence was reported to suggest that the energy gap does not always exist for horizontal velocity and scalars (Young 1987; Nucciarone and Young 1991; Davis et al. 1996; Duynkerke 1998; Tjernström et al. 2004; Tjernström and Mauritsen 2009). The anticascade of energy has been argued to be responsible for the appearance of the \(-\frac{5}{3}\) slope at the mesoscale (Gage 1979; Lilly 1983). On the other hand, LES tends to show that the spectral peak for scalars appears at a much larger scale than for \( w \) and that the power weakens as \( \lambda \) increases for a sufficiently large domain (Jonker et al. 1999; DDJ04; Moeng et al. 2010). For the free atmosphere, the \(-\frac{5}{3}\) slope has been observed for \( u, v \), and \( \theta \) at \( \lambda \) up to \( O(10^2 \text{ km}) \) (Lilly and Petersen 1983; Nastrom and Gage 1985), and reproduced by the numerical weather forecasting mode of ARW for a domain of \( O(10^3 \text{ km}) \) (Skamarock 2004). We hypothesize that the absence of
larger-scale motions greater than the domain size in the current simulations lacks an energy cascade from the larger scale, and that as a result a spectral peak appears for $V_H$ and scalars.

Considering the above discussion and looking back at Fig. 10, as DDJ04 suggested, our largest domain (i.e., $L = 25.6$ km) is not large enough. It is easy to understand why $V_H^2$ increases as the domain widens, following DDJ04. Since larger power appears at larger $\lambda$ up to the spectral peak, the total power (i.e., area under the curve) effectively increases. For the current simulation, increasing the domain size strengthens $w^2$ only marginally through mesoscale fluctuation, since the characteristic $\lambda$ (approximately the boundary layer depth) has already appeared within the domain.

How large should the domain be for our case? Wood and Hartmann (2006) found that the characteristic scale of closed cells increases significantly when the PBL height is higher than 1 km; the characteristic scale jumps from 10 to 40 km as the PBL height changes from 760 to
1100 m. This implies that $L \approx 50$ km might be required for the current case since the PBL height is approximately 1300 m. If multiple cells are required for an adequate simulation, $L = 200$ km would be necessary. Is the duration of our simulations long enough? The required duration of the simulation should scale with the size of the domain. The vertically averaged PSDs plotted every hour from 0500 UTC are shown in Fig. 10b.
power across all $\lambda$ does not change over the course of the free run for 6H, which suggests that the turbulence has fully developed by 0500 UTC. This translates to a 6-h duration being long enough for $L = 6.4$ km. On the other hand, the power for the larger $\lambda$ increases with time and becomes almost steady after 0900 UTC for 25B. A simulation period of 10 h is required for the turbulence and closed cellular circulations to develop fully for $L = 25.6$ km. Applying the same estimate to $L = 12.8$ km, one finds that the necessary duration is 8 h (not shown).

For the lidar observations, the behavior of $\overline{V^2}$ with increasing spatial scale is expected; that is, $\overline{V^2_{ij}}$ should increase as the spatial scale characterizing variation becomes larger. As noted above, the good agreement between the observation and 25B is misleading. That result simply suggests that the modeled $\overline{V^2_{ij}}$ becomes comparable to the observed value representing a 3–4-km extent with use of $L = 25.6$ km. It reinforces that the model in its current configuration generally underestimates the variance at the same spatial scale.

### 7. Discussion

Generally, GASS LES intercomparison studies (Moeng et al. 1996; Bretherton et al. 1999b,a; Stevens et al. 2001; Brown et al. 2002; Siebesma et al. 2003; Duynkerke et al. 2004; S05; Ackerman et al. 2009; vanZanten et al. 2011; Blossey et al. 2013) show a wide spread of variability among participating models. Sensitivity of model results to various SGS and advection schemes are often discussed. Here, we perform three additional free runs based on 6B. One simulation is run without physical diffusion but with surface forcing. The other two simulations are run with second- and fourth-order momentum advection schemes.

The results for the run without physical diffusion and the run with the second-order momentum advection scheme are shown in Fig. 11. The shading of Fig. 11b covers the last 6 h for both model and observation, which is different to the time scale used in the previous section (4 h). The run with the fourth-order momentum advection scheme produces similar results to the run with the second-order scheme. The results are considerably different from those with the original configuration (6B). This is not a surprise since like the abovementioned GASS LES intercomparisons, these additional simulations are produced by models that differ from the originally configured ARW.

Turning off the physical diffusion produces larger LWP and stronger turbulence (S05) toward the end of the simulation, and provides a better match to the lidar measurement. Since the spatial scale for lidar (3–4 km) and the model domain size ($L = 6.4$ km) is somewhat comparable, this result is promising. The increase in turbulence is expected since the TKE does not dissipate through the SGS scheme but dissipates only through the model’s inherent numerical diffusion. Physically, it can be understood as a radiation–turbulence feedback: all else equal, larger TKE results in larger LWP unless the turbulence intensity overcomes the control of the radiative cooling, which leads to a negative feedback (Moeng and Schumann 1991; Randall et al. 1992; Moeng et al. 1995). This result suggests that the model’s inherent numerical diffusion may be large or the SGS scheme may not be adequate for the original model configuration used in this study.

The second-order (and fourth order) momentum advection results in larger horizontal TKE than that of the original free run in the early stages when the LWP is similar. The even-order momentum advection preserves the shape of the turbulence profiles from the spinup run. The LWP decreases toward the end of simulation, as does the turbulence. This might be related to the buildup of energy at the shortest $\lambda$, which produces aliasing error to the longer $\lambda$, resulting in a less reliable solution for the physical phenomena represented by these waves (Skamarock 2004). There is a possibility that the spinup run with the even-order momentum advection is necessary for successful simulation of this case. This will require an iteration of the test involving modification of the sounding and $\tau$ and is not pursued here.

### 8. Summary

Using a set of input measurements from the RHB during VOCALS-REx, we have examined the ability of the ARW to reproduce the turbulence profiles measured by NOAA’s high-resolution Doppler lidar. The influences of domain size ($L = 6.4$–25.6 km) and resolution are also examined.

A two-stage simulation strategy was developed to generate a realistic simulation in order to compare with the time-evolving observations. In the first “spinup” stage, the mean profiles are nudged to the observed profiles allowing the turbulence to develop and become steady. After some iteration and calibration of the initial soundings and $\tau$, the spinup run produces an LWP comparable to the observation. This was deemed a prerequisite for comparison, given the close link between LWP and turbulence.

The simulated turbulence and observed turbulence are derived from two different averaging methods: the former uses a spatiotemporal mean, while the latter samples limited parts of the domain. A methodology is proposed to identify the appropriate spatiotemporal averaging scales for comparing these two different statistics. We
utilize Lagrangian column tracking, which imitates the ship measurement inside the simulation domain. The analysis shows that depending on the averaging time scale, the temporal variability of the column statistics changes substantially. The temporal variation over 4 h between the spatial-mean ARW saved every minute and the 1-h-averaged Lagrangian column processed like the lidar measurements is comparable. Thus, with this averaging time scale, reasonable agreement would be expected if ARW were to reproduce the observed turbulence.

In the second, “free running” stage, we first modified our configuration so that a steady PBL top was maintained and the turbulence became more radiatively driven. These modifications result in turbulence closer to that observed, and better than the turbulence generated by the original free run, but the modeled turbulence is still underestimated. As $L$ increases, the $V_H$ turbulence intensifies, while the $w$ turbulence shows weak sensitivity to these changes. A PSD analysis shows that the $V_H$ turbulence improves significantly in response to increasing $L$ until the spectral peak appears, while the $w$
turbulence improves only slightly in response to the increased \( L \). Once the spectral peak appears, the simulated turbulence improves very gradually when increasing resolution and \( L \). Convergence is expected for both variables at sufficiently large \( L \), provided that the resolution is fixed. Compared to the effect of \( L \), the overall effect of resolution is minor for the resolutions tested. However, we found that the shape of the spectra differs with resolution, which may impose an artificial bias on turbulence and cloud water. Analysis of the effect of \( L \) and resolution on the variability of LWP suggests that it is expected to converge with sufficiently large domain and sufficiently high resolution.

Sensitivities associated with the model dynamics and physics are more significant than those associated with grid spacing and \( L \). This fact sheds light on the ARW behavior for a prescribed configuration of the model dynamics and physics; for a given choice of core parameterizations (e.g., advection, diffusion), the model results are relatively insensitive to resolution and domain size. Much work remains to investigate the influence on ARW simulations of \( \Delta x/\Delta z \) and associated influence on entrainment, SGS scheme, inclusion of the Chilean coastal low-level jet, advection, and diffusion, especially for the near-LES approach.

Acknowledgments. This study is supported by the NOAA Climate Program Office through the Climate Process Team, Cloud Macrophysical Parameterization and its Application to Aerosol Indirect Effects (PI V. Larson). Useful discussion with the Climate Process Team is acknowledged. The authors would like to express their gratitude to the team on board the RHB for their efforts in acquiring such a valuable dataset. In particular, thanks are due to D. Covert for the aerosol size distribution data and P. Zuidema for the LWP data. The ECMWF operational data are courtesy of I. Sandu and M. Koehler.

#### APPENDIX A

Input Soundings and Data

The VOCALS-REx datasets are publicly available. Table A1 lists the data source and specific data used to construct input profiles as well as surface forcing.

The initial soundings for \( u, v, \theta_i \), and \( q_t \) are constructed with the following method based on the sonde and lidar data. From the sonde data, \( u, v, \theta, q_t \), and relative humidity are used. From the lidar data, \( u, v \), and the number of samples at each level are used. Curve fitting and idealization are applied to obtain smooth profiles.

First, the sonde and lidar data are vertically interpolated with the monotonic cubic interpolation of Steffen (1990) to every 1-m height. Then, six levels are set:

\[
\begin{align*}
\text{Just above surface} & \quad z_S = 100 \text{ m} \\
\text{Cloud base} & \quad z_C = \text{highest level of 95% relative humidity} \\
\text{Mixed-layer top} & \quad z_B = 1240 \text{ m} \\
\text{Inversion-layer top} & \quad z_I = 1272 \text{ m} \\
\text{for} & \quad \mathbf{V}_H \quad z_{BV} = 1130 \text{ m} \\
\text{for} & \quad \mathbf{V}_H \quad z_{IV} = 1272 \text{ m}
\end{align*}
\]

The mixed-layer and inversion-layer tops for \( \mathbf{V}_H \) are set differently from those for the thermodynamic variables since the jump in \( \mathbf{V}_H \) is not as sharp as that for thermodynamic variables. These values are determined iteratively by comparing the resulting soundings with the raw data. For the background \( \mathbf{V}_H \) at 0814 (1217) UTC, \( z_{BV} = 1200 \) (1250) m and \( z_{IV} = 1350 \) (1400) m.

The \( \mathbf{V}_H \) profile is constructed based on the lidar and sonde data. The lidar profile is used from the surface to the highest level with a sample count of at least 300; then, the sonde profile is connected above. The free-atmospheric value is set constant and the value at \( z_{IV} \) is used. The mixed-layer profile is then obtained with a quadratic polynomial fit based on the data between \( z_S \) and \( z_{BV} \). The mixed-layer and free-atmospheric profiles are connected via monotonic cubic interpolation.

Observations of \( \theta_i, q_t, \) and \( q_v \) are not available but \( \theta \) and \( q_v \) are. For \( \theta_i \) and \( q_t \), the profile for the subcloud layer is made using a linear fit based on the \( \theta \) and \( q_v \) profiles between \( z_S \) and \( z_C \). Using the well-mixed assumption in

---

**Table A1. List of data sources and data names.**

<table>
<thead>
<tr>
<th>Data Source</th>
<th>URL/data name</th>
</tr>
</thead>
<tbody>
<tr>
<td>General data source</td>
<td><a href="http://data.eol.ucar.edu/master_list/?project=VOCALS">http://data.eol.ucar.edu/master_list/?project=VOCALS</a></td>
</tr>
<tr>
<td>Radiosonde</td>
<td>NOAA R/V RHB radiosonde 10-m gridded soundings</td>
</tr>
<tr>
<td>Lidar</td>
<td>NOAA R/V RHB lidar data</td>
</tr>
<tr>
<td>Surface</td>
<td>NOAA R/V RHB meteorological, radiation, GPS, and flux data</td>
</tr>
</tbody>
</table>
the cloud layer (i.e., constant profile with the value at \(z_C\)), the profile is extended up to \(z_B\). For \(\theta_i\), a cubic polynomial fit is used to construct the free-atmospheric profile (above \(z_I\)). For \(q_t\), a value of 1 g kg\(^{-1}\) is specified for the free atmosphere. The mixed-layer and free-atmospheric profiles are connected with a linear fit. With these profiles, the cloud base is evaluated with a saturation adjustment. The value at the new cloud base is obtained by a linear fit with the same coefficient previously used for the subcloud-layer profile. The cloud-layer profile is reconstructed with the well-mixed assumption, and then the inversion profile (between \(z_B\) and \(z_I\)) is recomputed.

APPENDIX B

Acoustic Courant Number Convergence

YF12 evaluated the LES mode of ARW and found sensitivity of results to the acoustic Courant number \([c_t \Delta t/(N_{AT} \Delta x)]\), where \(c_t\) is the speed of sound. The strength of the sensitivity depends on case configuration, and the source of the sensitivity is currently unknown. It is advisable to perform a convergence test so that the result is independent of this sensitivity.

For \(\Delta x = 100\) m and \(\Delta z \approx 10\) m, the combination of \(\Delta t = 1\) s and \(N_{AT} = 8\) gives a stable solution based on the Courant–Friedrichs–Lewy stability criteria for both physical and acoustic modes. Using the 6B configuration without nudging, 10 simulations with different combinations of \(\Delta t\) and \(N_{AT}\) are performed for a 2-h duration. The results are presented in Fig. B1. If only one-dimensional time series are relied on for the convergence test, one would mistakenly judge the convergence at \(\Delta t = 1\) s and \(N_{AT} = 12\). The true convergence, however, occurs at \(\Delta t = 0.2\) s and \(N_{AT} = 8\) for the one-dimensional time series as well as for the horizontal mean profiles. The result illustrates that the one-dimensional time series is not a rigorous-enough requirement for convergence.

APPENDIX C

Nudging

For a variable \(\phi\), the nudging tendency is computed with the observed and horizontal-mean profiles as

\[
N_{ijk} = \frac{\partial \bar{\phi}_k - \bar{\phi}_k}{\tau},
\]
where the tilde represents the observation, the superscript \( t \) is the time level, and the subscripts \( i \) and \( j \) (\( k \)) represent the horizontal (vertical) components. We retain the horizontal indices for the following discussion, even though the above nudging term depends only on the vertical direction. As described in YF12, the horizontal mean profile is constructed with a vertical interpolation at a high-resolution height coordinate, which has 10 times the model vertical levels, owing to the fact that ARW is formulated with a mass vertical coordinate. The nudging tendency is, then, computed on each high-resolution level, and interpolated back to the local level.

In ARW, \( q \), \( n_t \), and \( n_t \), are not prognostic variables but \( q_v \), \( q_v \), and \( n_v \) are. For this reason, \( N_{i,j,k} \) for \( q_t \) and \( n_t \) is partitioned by a mass or number weighting. For example, \( N_{i,j,k} \) for \( q_v \) is given as

\[
\left( \frac{\partial q_v}{\partial t} \right)_{\text{nudging}} = \left( \frac{q_v}{\partial t} \right)_{i,j,k} \left( q_v \right)_{i,j,k} \left( \frac{\partial q_v}{\partial t} \right)_{\text{nudging}}. \tag{C2}
\]

The application of \( N_{i,j,k} \) given above has an undesirable effect on the monotonicity of the variable; for instance, an undesirable kink of water vapor at the inversion top forms because of nudging. A simple discrete form of the advection equation for a variable \( \phi \) including a nudging term can be written as

\[
\phi_{i,j,k}^t + \Delta t = \phi_{i,j,k}^t + A_{i,j,k} \Delta t + N_{i,j,k} \Delta t, \tag{C3}
\]

where \( A \) represents the advective tendency. By denoting

\[
\phi_{i,j,k}^{\text{min}} = \min(\phi_{i,j,k}^t, \phi_{i,j,k}^{t-1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}),
\]

\[
\phi_{i,j,k}^{\text{max}} = \max(\phi_{i,j,k}^t, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}, \phi_{i,j,k+1}).
\]

The local minimum and maximum are an estimation of the lower and upper limit for \( \phi_{i,j,k}^{\text{adv}} \) without knowing the actual \( \phi_{i,j,k}^{\text{adv}} \). Since \( \phi_{i,j,k}^{\text{min}} \) involves a comparison with the upper- and lower-level values, \( \phi_{i,j,k}^{\text{min}} \) can be smaller than \( \phi_{i,j,k}^{\text{max}} \) while \( \phi_{i,j,k}^{\text{max}} \) can be larger than \( \phi_{i,j,k}^{\text{adv}} \).

The nudging tendency is limited as follows: When \( N_{i,j,k} > 0 \), it is limited by

\[
N_{i,j,k} = \begin{cases} 
0 & \text{if } \phi_{i,j,k}^{\text{min}} \geq \phi_{i,j,k}^{\text{adv}} \\
\max \left( N_{i,j,k}, \frac{\phi_{i,j,k}^{\text{min}} - \phi_{i,j,k}^{\text{adv}}}{\Delta t} \right) & \text{if } \phi_{i,j,k}^{\text{min}} < \phi_{i,j,k}^{\text{adv}}.
\end{cases}
\]

(C8)

Quick inspection of (C5) tells us that when \( N_{i,j,k} \) becomes a sink (i.e., \( N_{i,j,k} < 0 \)), then nudging creates a new minimum if \( \phi_{i,j,k}^{\text{adv}} \) is minimum. Likewise, when nudging acts as a source, it creates a new maximum if \( \phi_{i,j,k}^{\text{adv}} \) is maximum. The new minimum becomes a new lower bound for a monotonic advection limiter and it is advected. The effect increases with smaller \( \tau \) and accumulates over time.

To achieve monotonicity after nudging for advection equation, \( N_{i,j,k} \) has to be limited so that the creation of a new minimum and maximum is prohibited. For this purpose, we developed a monotonic nudging limiter in the spirit of the flux-corrected transport (Zalesak 1979) as described below.

We construct the profiles of the layer minimum and maximum by

\[
\begin{cases}
\phi_{i,j,k}^{\text{min}} = \min(\phi_{i,j,k}^t) & \text{for all } i \text{ and } j \\
\phi_{i,j,k}^{\text{max}} = \max(\phi_{i,j,k}^t) & \text{for all } i \text{ and } j.
\end{cases}
\]

(C6)

Like the layer-mean profile, the high-resolution height coordinate is used to construct the layer-minimum and -maximum profile before they are interpolated back to the local level. The limiter also requires the local minimum and maximum by

\[
\begin{cases}
\phi_{i,j,k}^{\text{min}} = \min(\phi_{i,j,k}^t) & \text{for all } i \text{ and } j \\
\phi_{i,j,k}^{\text{max}} = \max(\phi_{i,j,k}^t) & \text{for all } i \text{ and } j.
\end{cases}
\]

(C7)

When \( N_{i,j,k} > 0 \), it is limited by

\[
\begin{cases}
N_{i,j,k} = 0 & \text{if } \phi_{i,j,k}^{\text{min}} \leq \phi_{i,j,k}^{\text{adv}} \\
N_{i,j,k} = \min \left( N_{i,j,k}, \frac{\phi_{i,j,k}^{\text{max}} - \phi_{i,j,k}^{\text{min}}}{\Delta t} \right) & \text{if } \phi_{i,j,k}^{\text{max}} > \phi_{i,j,k}^{\text{adv}}.
\end{cases}
\]

(C9)

For example, when nudging acts as sink, the limiter does not apply nudging if \( \phi_{i,j,k}^{\text{min}} \) is smaller than \( \phi_{i,j,k}^{\text{adv}} \) to prevent possible creation of a new minimum. The nudging tendency is limited with the local nudging tendency, which requires \( \phi_{i,j,k}^{\text{adv}} \) to be \( \phi_{i,j,k}^{\text{min}} \) over \( \Delta t \); that is, \( (\phi_{i,j,k}^{\text{adv}} - \phi_{i,j,k}^{\text{min}})/\Delta t \).
if \( \phi_{i,j,k}^{\text{min}} \) is larger than \( \phi_{i,j,k}^{\text{max}} \). Last of all, \( N_{i,j,k} \) with the limiter is written as one line of code:

\[
N_{i,j,k} = \max \left[ \min(N_{i,j,k}, 0), \min \left( \frac{\phi_{i,j,k}^{\text{min}} - \phi_{i,j,k}^{\text{min}}}{\Delta t}, 0 \right) \right] + \max(N_{i,j,k}, 0), \max \left( \frac{\phi_{i,j,k}^{\max} - \phi_{i,j,k}^{\max}}{\Delta t}, 0 \right).
\]

(C10)

APPENDIX D

Computing Power Spectral Density

Our discrete Fourier transform, which converts a function in physical space (or time) to wavenumber (or frequency), denoted by \( k \), spectra, has the form

\[
s(n\Delta k) = \frac{1}{N} \sum_{m} \left[ f(m\Delta x) \exp \left( -2\pi i \frac{mn}{N} \right) \right], \tag{D1}
\]

where \( k = n\Delta k, \Delta k = 1/L \) is an interval of \( k, \Delta x = L/N \) is a space (or time) interval, \( L \) is a length of space (or period), \( N \) is a number of data points, and \( m \) and \( n \) are integers ranging \( 0 \leq m \leq N-1 \) and \( -0.5N < n \leq 0.5N \). Hence, \( n \) is a discrete form of normalized \( k \). It should be noted that the unit of a spectrum is the unit of \( f \).

The same argument made for the one-dimensional case between \( p \) and \( P \) can be applied, so that

\[
P(n\Delta k, n\Delta k) = L_{x}L_{y}p(n_{x}\Delta k_{x}, n_{y}\Delta k_{y}) . \tag{D7}
\]

Because of two dimensionality, we cannot directly compare it to the one-dimensional PSD. Following DDJ04, the two-dimensional \( k \) are transformed to cylindrical coordinates \((k, \theta)\) so that

\[
\begin{align*}
    n_{x}\Delta k_{x} &= n\Delta k \cos \theta \\
n_{y}\Delta k_{y} &= n\Delta k \sin \theta,
\end{align*}
\]

(D8)

where \( n\Delta k = [(n_{x}\Delta k_{x})^{2} + (n_{y}\Delta k_{y})^{2}]^{1/2} \) is the horizontal plane \( k \), and \( \theta = \tan^{-1}[(n_{y}\Delta k_{y})/(n_{x}\Delta k_{x})] \). Then, \( V \) can be expressed in continuous form as

The power spectrum is the square of the absolute value of \( s \):

\[
p(n\Delta k) = |s(n\Delta k)|^{2} . \tag{D2}
\]

The advantage of using (D1) is that simple summation over all \( k \) is equivalent to the total variance \( V \):

\[
V = \frac{1}{N} \sum_{m} f^{2}(m\Delta x) = \sum_{n} p(n\Delta k) . \tag{D3}
\]

Frequently, only positive \( k \) values are used for the lateral axis. For this situation, in order to conserve \( V \), one adds \( p(-|n\Delta k|) \) to the corresponding \( p(|n\Delta k|) \) and discards all \( p(-|n\Delta k|) \). The range of \( n \) becomes \( 0 \) and \( 0.5N \).

Instead of using power spectra, the power spectral density (PSD) \( P \) is useful, especially comparing data of different \( L \), because the area under the curve of PSD is \( V \) so that

\[
\sum_{n} p(n\Delta k) = \sum_{n} [P(n\Delta k)\Delta k] . \tag{D4}
\]

It is straightforward to deduce that the PSD is given as

\[
P(n\Delta k) = Lp(n\Delta k) . \tag{D5}
\]

The unit of the PSD is the unit of \( f^{2} k^{-1} \). One can show that (D5) is valid by starting with the continuous form of the spectral density obtained from the Fourier transform.

For two-dimensional data (e.g., a horizontal plane at each level), our discrete Fourier transform has the form

\[
s(n_{x}\Delta k_{x}, n_{y}\Delta k_{y}) = \frac{1}{N_{x}N_{y}} \sum_{m_{x}} \sum_{m_{y}} \left[ f(m_{x}\Delta x, m_{y}\Delta y) \exp \left[ -2\pi i \left( \frac{m_{x}n_{x}}{N_{x}} + \frac{m_{y}n_{y}}{N_{y}} \right) \right] \right] . \tag{D6}
\]

\[
V = \left\{ \begin{array}{cc}
\int L_{x}L_{y}p(k_{x}, k_{y})dk_{x}dk_{y} & \text{Cartesian coordinate} \\
\int L_{x}L_{y}k_{p}(k, \theta)dkd\theta & \text{cylindrical coordinate}
\end{array} \right. . \tag{D9}
\]

In discrete form, \( V \) in the cylindrical coordinate is expressed as

\[
V = \sum_{i} \left[ L_{x}L_{y} \sum_{\Delta\theta} n\Delta k p(n\Delta k, \theta_{i})\Delta\theta \right] \Delta k , \tag{D10}
\]

so that PSD is given by

\[
P(n\Delta k) = L_{x}L_{y}n\Delta k \sum_{i} [p(n\Delta k, \theta_{i})\Delta\theta] : \tag{D11}
\]

\( i = \{1, \ldots, I\} \) is the index for angles.
The two-dimensional power spectrum has two symmetry components. Before integrating $p$ over the angle in (D11), $p(-n\Delta k, \theta)$ is added to the corresponding $p(n\Delta k, \theta)$. This procedure constrains the range of $n$ between 0 and 0.5N and the range of angles between $-\pi/2$ and $\pi/2$. Adding one symmetry component ($-\pi/2 \leq \theta \leq 0$) to the other symmetry component further reduces the range of the angle to $0 \leq \theta \leq \pi/2$.

One can use various ways to integrate $p$ over angles. Here we use $\Delta \theta = \pi/(2n_{\text{arc}}N)$ and $n_{\text{arc}}$ is constant for all $k$. In this way, the length of arc for all $k$ is the same (i.e., $n\Delta k\Delta \theta = \Delta k\pi/2n_{\text{arc}}$). Bilinear interpolation is used to diagnose $p$ at every angle. The value of $n_{\text{arc}}$ is decided by iteratively incrementing $n_{\text{arc}}$ from 1 until the maximum improvement of the angle integration of $p$ for all $k$ becomes less than $10^{-3}$.

For zonal velocity, $p(0)$ is much larger than that of $n = 1$, and the use of interpolation significantly overestimates $p(\Delta k)$ after the angle integration, so the perturbation of $f$ is used to compute $p$ instead of the raw $f$.

REFERENCES


