Understanding and Improving the Scale Dependence of Trigger Functions for Convective Parameterization Using Cloud-Resolving Model Data

FENGFEI SONG[a] AND GUANG J. ZHANG

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

(Manuscript received 1 October 2017, in final form 1 June 2018)

ABSTRACT

As the resolution of global climate model increases, whether trigger functions in current convective parameterization schemes still work remains unknown. In this study, the scale dependence of undilute and dilute dCAPE, Bechtold, and heated condensation framework (HCF) triggers is evaluated using the cloud-resolving model (CRM) data. It is found that all these trigger functions are scale dependent, especially for dCAPE-type triggers, with skill scores dropping from ~0.6 at the lower resolutions (128, 64, and 32 km) to only ~0.1 at 4 km. The average convection frequency decreases from 14.1% at 128 km to 2.3% at 4 km in the CRM data, but it increases rapidly in the dCAPE-type triggers and is almost unchanged in the Bechtold and HCF triggers across resolutions, all leading to large overpredictions at higher resolutions. In the dCAPE-type triggers, the increased frequency is due to the increased rate of dCAPE greater than the threshold (65 J kg^{-1} h^{-1}) at higher resolutions. The box-and-whisker plots show that the main body of dCAPE in the correct prediction and overprediction can be separated from each other in most resolutions. Moreover, the underprediction is found to be corresponding to the decaying phase of convection. Hence, two modifications are proposed to improve the scale dependence of the undilute dCAPE trigger: 1) increasing the dCAPE threshold and 2) considering convection history, which checks whether there is convection prior to the current time. With these modifications, the skill at 16 km, 8 km, and 4 km can be increased from 0.50, 0.27, and 0.15 to 0.70, 0.61, and 0.53, respectively.

1. Introduction

Because of coarse spatial resolution, convective parameterization schemes are used in global climate models (GCMs) to represent the collective effect of subgrid-scale convection. As an important component of convective scheme, the trigger function mainly determines whether the scheme should be invoked using some criteria for the possibility of convection onset. To a large degree, it determines not only the timing and location of model convection, but also its frequency of occurrence, through which the trigger function exerts substantial influences on the simulation of Madden–Julian oscillation (Zhang and Mu 2005a) and the diurnal cycle of precipitation (Zhang 2003; Bechtold et al. 2004; Xie et al. 2004; Lee et al. 2007, 2008).

Currently, most convective schemes are mass flux based, in which the effect of convection is estimated using a one-dimensional bulk or spectral plume model. It assumes that convection is controlled by large-scale or grid-scale forcing in a statistical sense (Arakawa and Schubert 1974). Based on this assumption, many grid-scale variables, such as convective available potential energy (CAPE), generation of CAPE from large-scale forcing (dCAPE), vertical velocity at the lifting condensation level (LCL; Kain and Fritsch 1990; Bechtold et al. 2001), or near-surface temperature and moisture (Tawfik and Dirmeyer 2014), are used to construct the trigger function. In the past, convective parameterization schemes have been used under the requirement that convection occupies a very small area in a grid cell, that is, the convective cloud fraction σ is much less than 1 ($\sigma \ll 1$; Arakawa et al. 2011; Arakawa and Wu 2013). When the horizontal resolution of GCMs increases to the gray zone (below 10 km), where there is no clear separation between grid scale and subgrid scale, this requirement is no longer satisfied. Arakawa and Wu (2013) suggested that the conventional convective scheme

[a] Current affiliation: Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, Washington.

Corresponding author: Fengfei Song, songfengfei@gmail.com

DOI: 10.1175/JCLI-D-17-0660.1

should include a factor $1-\sigma$ to become scale aware. However, Zhang et al. (2015) pointed out that the assumption of $\sigma \ll 1$ is unnecessary for the conventional convective scheme and the scale awareness factor $1-\sigma$ has already been included implicitly in the formulation of convective eddy transport. Regardless, other components in convective schemes such as trigger function or closure can be scale dependent as well.

In recent years, cloud-resolving models (CRM) have made substantial progress and can simulate convection systems realistically, including precipitation rate (Fridlind et al. 2012). Hence, it provides a great opportunity to use them for convective scheme development and evaluation. In particular, the high-resolution CRM data can be spatially averaged over different subdomains to mimic GCM grid boxes at different resolutions to examine whether the convective scheme is scale aware. Suhas and Zhang (2015) evaluated different closure variables of convective schemes and found that the moisture convergence-based closure should scale down when the resolution increases. Similarly, the CRM datasets are also used to examine the scale dependence of convective transport (Liu et al. 2015; Xiao et al. 2015; Yun et al. 2017).

In this study, we use the CRM output to examine the scale awareness of the four trigger functions evaluated in Song and Zhang (2017). The following are three objectives in this paper: 1) examining whether the trigger functions still work in high resolutions, 2) understanding why the trigger functions are scale dependent physically, and 3) developing a scale-aware trigger function based on the funderstanding. The paper is organized as follows: section 2 describes the trigger functions and CRM dataset, section 3 presents the main results, and conclusions and discussions are shown in section 4.

2. Trigger functions, cloud-resolving model data, and evaluation method

a. Trigger functions

Based on Song and Zhang (2017), the best four trigger functions [undilute and dilute dCAPE, Bechtold, and heated condensation framework (HCF) triggers] are selected to explore whether they are scale dependent. The dCAPE-type triggers (Xie and Zhang 2000; Zhang 2002) use the CAPE generation rate from large-scale forcing in the free troposphere (dCAPE) as the threshold ($\mathrm{dCAPE} > 65\,\mathrm{J\,kg^{-1}\,h^{-1}}$). dCAPE is calculated as the amount of CAPE generated by the large-scale dry static energy (DSE) advection and moisture advection during a time interval:

$$\mathrm{dCAPE} = \frac{\mathrm{CAPE}[T + \mathrm{adv(DSE)}\delta t,\, q + \mathrm{adv}(q)\delta t] - \mathrm{CAPE}(T,\, q)}{\delta t}, \tag{1}$$

where T and q are temperature and moisture, respectively; and $\mathrm{adv(DSE)}\delta t$ and $\mathrm{adv}(q)\delta t$ are the temperature and moisture increments by large-scale advection over a time interval δt, respectively. The dilute dCAPE trigger also considers the dilution process in CAPE calculation, which uses the same entrainment rate as that used in Neale et al. (2008). The Bechtold trigger (Bechtold et al. 2001) uses large-scale vertical velocity at the lifting condensation level to construct a temperature perturbation as follows:

$$\delta T = \mathrm{sign}(w) c_w \left| \frac{w A^{1/2}}{\Delta x} \right|^{1/3}, \tag{2}$$

where $\mathrm{sign}(w)$ is 1 when $w > 0$ and $\mathrm{sign}(w)$ is -1 when $w \le 0$, $c_w = 6\,\mathrm{Km^{-1/3}\,s^{1/3}}$, Δx is the grid spacing in kilometers, w is the grid-scale vertical velocity, and A is the area of the grid cell. Here, we set $A^{1/2}/\Delta x = 1$ in each resolution following Suhas and Zhang (2014). Note this setting is different from the original Bechtold trigger (Bechtold et al. 2001), in which Δx is fixed as 25 km at different resolutions. In the results section, these two settings will be compared. In the Bechtold scheme, convection is triggered when the sum of parcel temperature θ_{parcel} and temperature perturbation δT is greater than the environment temperature θ_{env} at the LCL. The HCF trigger (Tawfik and Dirmeyer 2014) uses the near-surface temperature and moisture to represent the incremental growth of the planetary boundary layer caused by surface heating. In this trigger, it does not require the selection of air parcel. Instead, it gradually increases the near-surface potential temperature and mixes the specific humidity from the surface to the level of neutral buoyancy. When the saturation occurs at the top of the potential mixed level, it returns the total temperature increments referred to as the potential temperature deficit, θ_{def}, which represents the heating needed to achieve the saturation. Hence, convection is triggered in the HCF when $\theta_{\mathrm{def}} \le 0$. More details about these four trigger functions can be found in the above references, Suhas and Zhang (2014) and Song and Zhang (2017).

b. Cloud-resolving model data

In this study, the output from a three-dimensional CRM, the Goddard Cumulus Ensemble (GCE) model

(Tao and Simpson 1993), simulation forced by the Tropical Warm Pool International Cloud Experiment (TWP-ICE) large-scale forcing data is analyzed. The TWP-ICE field experiment was conducted during January and February 2006, which was the northern Australian summer monsoon season. The simulation period is from 2100 UTC 4 February 2006 to 2100 UTC 10 February 2006 and the output interval is 6 min. The model domain is an area of $256 \times 256\,\mathrm{km}^2$ centered on the Atmospheric Radiation Measurement (ARM) Program intensive observation period at the Darwin, Australia, site. The model has a 1-km horizontal resolution and 41 vertical levels. As pointed out by Suhas and Zhang (2015), 1-km CRM simulation is a trade-off between model resolution and domain coverage, and may still have substantial errors in simulating convection. However, we mainly focus on the collective properties of convective clouds within a GCM-sized region, so it may not exert large influence on our results. See Suhas and Zhang (2015) and Zeng et al. (2011) for more details about the simulation and model. As shown in Suhas and Zhang (2015), the model is capable of explicit simulation of tropical convection, with a good agreement between observed and simulated precipitation events. Hence, this CRM simulation can be used as ''pseudo observation'' to gauge the performance of trigger functions.

To assess the performance of trigger functions at different resolutions corresponding to GCMs, the CRM data are averaged over different horizontal subdomains of sizes 128, 64, 32, 16, 8, and 4 km, respectively. When the 1-km horizontal resolution CRM data over the $256 \times 256\,\mathrm{km}^2$ domain is divided into four subdomains of $128 \times 128\,\mathrm{km}^2$ each, we regard the resultant data as equivalent to the output of a GCM of 128-km horizontal resolution. The data for other resolutions are obtained similarly. Then the resultant data are used as the GCM grid-scale input to calculate the convection occurrence in the trigger function. Since the experiment is forced by the ARM large-scale forcing, at the large subdomain (e.g., 128 km), the influence of large-scale forcing is evident, which can also be seen in the following analysis. In addition, at 128 km, there are only four subdomains, the poor sampling may also have some influences on the result. However, as the subdomain becomes smaller, the circulation generated by the model itself becomes more important to the location, frequency, and intensity of convection. When the subdomain falls to about 10 km, it reaches the gray zone, where the grid scale and subgrid scale become hard to distinguish, so how to parameterize convection becomes a problem (Arakawa et al. 2011; Arakawa and Wu 2013).

The temperature, moisture, vertical velocity, and horizontal wind fields are directly from the CRM and averaged into different GCM resolutions. The DSE advection and moisture advection at different resolutions are calculated using the subdomain-averaged fields. Convective precipitation and mass flux are first identified at 1-km resolution and then averaged into different resolutions. Following Suhas and Zhang (2015), convective precipitation is the precipitation occurring at the surface of a convective column, which is defined as an atmospheric column that contains at least one convective grid point. A convective grid point is defined as the grid point where either of the following two conditions is satisfied: 1) the absolute vertical velocity is greater than $2\,\mathrm{m\,s}^{-1}$ or 2) the absolute vertical velocity is greater than $1\,\mathrm{m\,s}^{-1}$ and the sum of the mixing ratios of cloud water, cloud ice, and rainwater is greater than $1 \times 10^{-5}\,\mathrm{kg\,kg}^{-1}$. The mass flux M is calculated by $M = \sigma\rho w$, where σ is convective cloud fraction (the ratio of convective grid points to all grid points in the corresponding subdomains), ρ is air density, and w is the vertical velocity averaged over all convective grid points at a given pressure level.

c. Evaluation method

As in Song and Zhang (2017), in all trigger functions, deep convection requires cloud depth, the thickness between LCL and the neutral buoyancy level, to be greater than 3 km except the HCF trigger. In the HCF, it requires the potential temperature deficit θ_{def} to be nonpositive, that is, $\theta_{\mathrm{def}} \le 0$. In the CRM data, we set a convective precipitation threshold of $0.5\,\mathrm{mm\,h}^{-1}$ at each resolution to exclude small convection events. Based on the predicted and observed convection at each time, the prediction results can be divided into four categories: correct prediction a (observation has convection and trigger predicts it), overprediction b (observation does not have convection but trigger predicts it), underprediction c (observation has convection but trigger misses it), and correct prediction of no convection d (observation does not have convection and trigger does not predict it), forming a 2×2 contingency table. To evaluate the prediction results, the Heidke skill score (HSS; Schaefer 1990; Hogan et al. 2009) is used:

$$\mathrm{HSS} = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}. \tag{3}$$

The highest value of HSS is 1, meaning perfect prediction, 0 means the prediction skill by chance, and negative values mean the prediction is worse than that by chance.

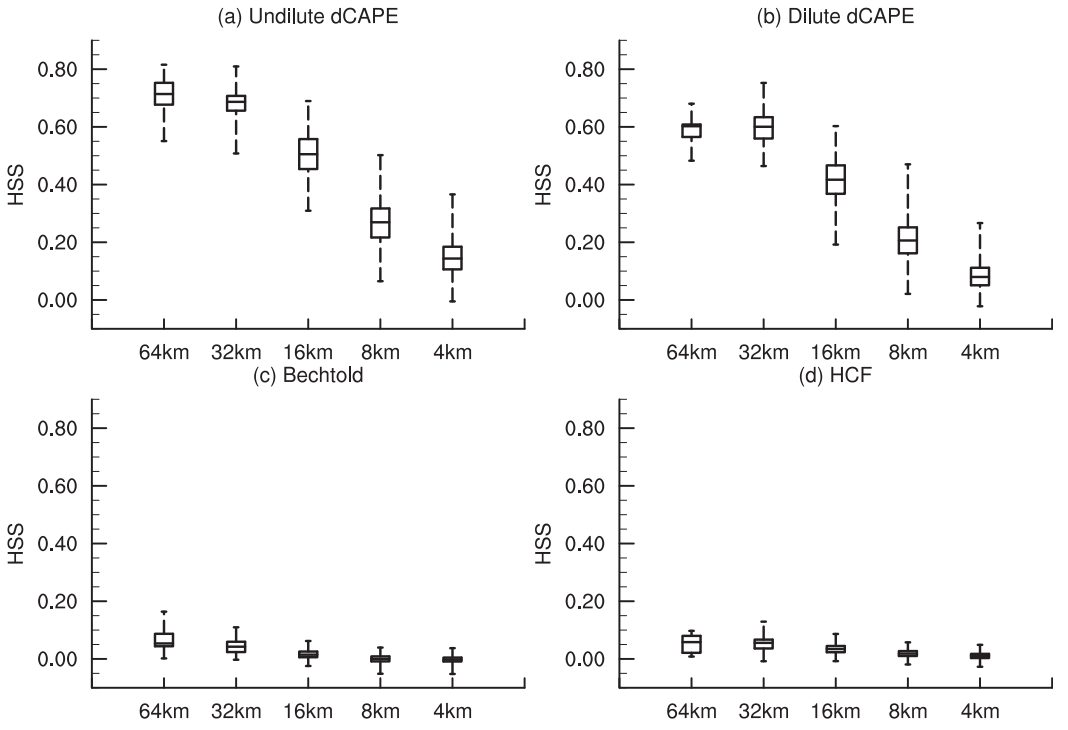

FIG. 1. The box-and-whisker plot of HSS at different resolutions for the (a) undilute dCAPE trigger, (b) dilute dCAPE trigger, (c) Bechtold trigger, and (d) HCF trigger. (from bottom to top) The horizontal lines represent the minimum, 25th percentile, median, 75th percentile, and maximum values, respectively.

3. Results

a. Understanding the scale dependence of the trigger function

At different resolutions, we apply Eq. (3) to all subdomains (equivalent to GCM grid points) and construct the box-and-whisker plots of HSS skill shown in Fig. 1 for dCAPE-type triggers, Bechtold, and HCF triggers. We only show the box-and-whisker plots from 64 to 4 km and omit the 128-km subdomain because of small sample size. For undilute dCAPE trigger, the skill from different subdomains of 128 km varies from 0.54 to 0.67. At 64 and 32 km, the median value of skill is slightly increased (0.71 and 0.69, respectively) and the skill range is also increased (0.55–0.82 and 0.51–0.81, respectively). The relatively lower skill score at 128 km compared to those at 64 and 32 km is likely due to the influence of the prescribed external forcing. From 32 to 4 km, the skill decreases rapidly, with the median value of 0.51, 0.27, and 0.14 for 16, 8, and 4 km, respectively. For dilute dCAPE trigger, it exhibits similar changes. At 128 km, the skill is about 0.32 (at its median value), but it increases to 0.60 at 64 km and 32 km, then decreases quickly from 32 km to 4 km. The Bechtold trigger only has HSS of 0.05–0.13 at 128 km, but it still decreases gradually with the resolution increase. The Bechtold trigger with fixed Δx shows similar results. The low skills of Bechtold at 128 km are consistent with Suhas and Zhang (2014) based on the TWP-ICE field campaign data. For the HCF trigger, the median value of HSS is only 0.02 at 128 km, and remains low across resolutions. Compared to the skill based on the Green Ocean Amazon (GOAmazon) field campaign data (Song and Zhang 2017), the skill of dCAPE-type triggers here at lower resolutions (128 and 64 km) is comparable, but the skill of Bechtold and HCF triggers is much lower. Further examination on the Bechtold trigger shows that the correct prediction and underprediction rates are almost the same in the two cases, but the overprediction rate is about 42% on average here, much larger than 22% in GOAmazon data. Hence, the much lower skill of Bechtold trigger here is due to the higher overprediction rate. For the HCF trigger, the almost zero skill here is also due to the considerably higher overprediction rate (51.5%–56.7%), which is only 4.9% in the GOAmazon. Further examination shows that over the western Pacific, when there is no convection, the frequency of near-surface warm and wet anomaly is much higher than that over the Amazon region. Hence, these two triggers, heavily dependent on the near-surface environment, may easily overpredict convection in these cases.

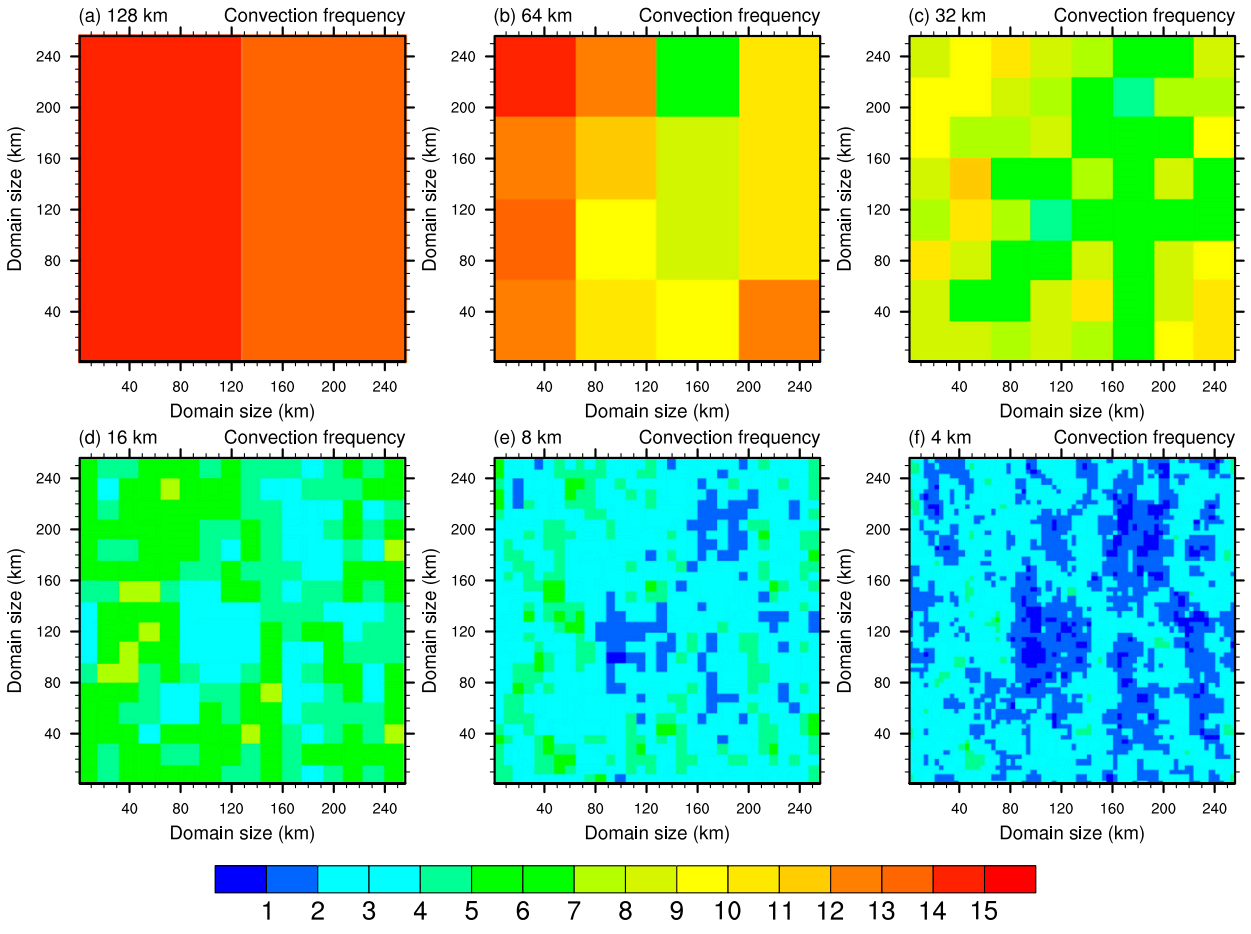

FIG. 2. The spatial distribution of convection frequency (unit: %) in the cloud-resolving model at (a) 128, (b) 64, (c) 32, (d) 16, (e) 8, and (f) 4 km.

As shown in Fig. 1, all the four trigger functions exhibit scale dependence, especially for dCAPE-type triggers. The dCAPE-type triggers perform well at low resolutions, but the performance degrades rapidly as resolution increases. To understand this behavior, we first examine the frequency of convection in the CRM data in each subdomain of different sizes (Fig. 2). Convection frequency in each subdomain is defined as the frequency of convective precipitation greater than 0.5 mm h^{-1} to exclude the very weak convection. Other precipitation thresholds show similar convection frequency change with resolution. At 128 km, convection frequency varies between 13.3% and 14.9%. As the resolution increases, it decreases rapidly. At 4 km, the average convection frequency is 2.3%, only about 1/6 of that at 128 km. This is easy to understand: in the 1-km CRM dataset, at each grid point and a given time convection occurrence is either 1 or 0. When calculating convection frequency in a small subdomain (e.g., 4 km × 4 km), convection occurrence will be 1 if convection occurs in any of the 16 (4 × 4) grid points. In a larger subdomain (e.g., 16 km × 16 km), convection occurrence will be 1 if convection occurs in any of the 256 (16 × 16) grid points. Hence, the larger the area of the subdomain (low resolution), the higher the chance is to observe convection in the CRM output.

Now we turn our attention to convection frequency determined from the four trigger functions (Fig. 3). Similar to Fig. 1, we only show the box-and-whisker plots of convection frequency from 64 to 4 km caused by poor sampling at 128 km. For the dCAPE-type triggers, in general, convection frequency increases with resolutions: it increases from 7.4% at 128 km to 12.2% at 4 km for the undilute dCAPE trigger, and it increases from 3.6% at 128 km to 12.2% at 4 km for the dilute dCAPE trigger. For the Bechtold trigger, convection frequency increases slightly from 128 to 64 km, then remains more or less unchanged when the resolution becomes further higher. The convection frequency in the Bechtold trigger with fixed Δx also keeps almost unchanged with the

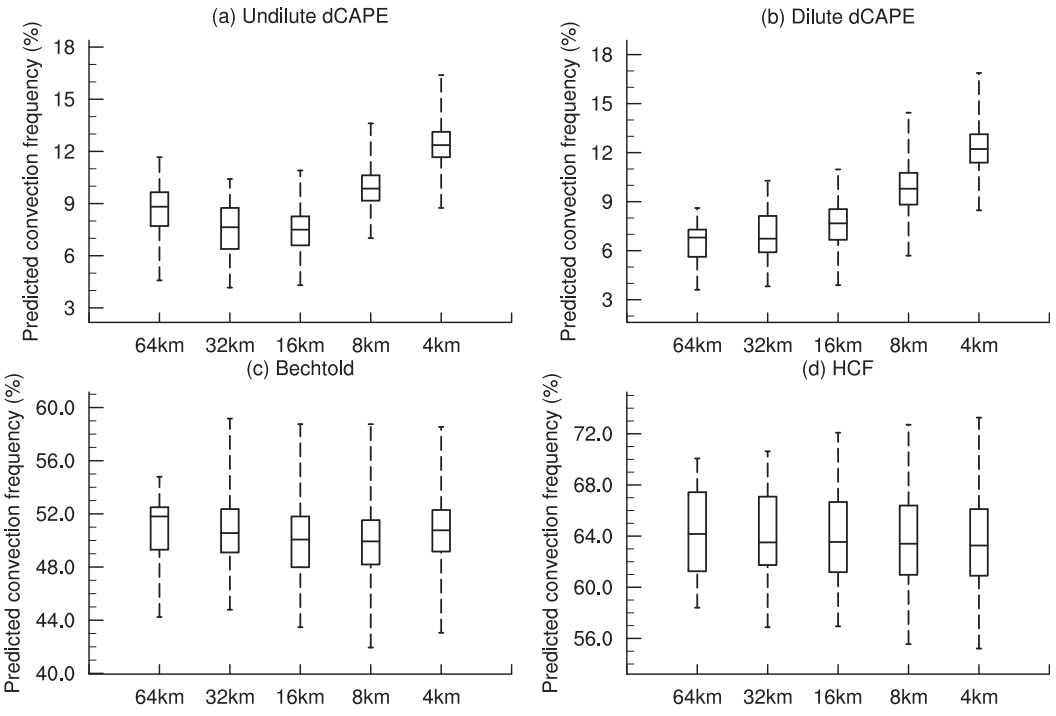

FIG. 3. The box-and-whisker plot of convection frequency at different resolutions for (a) undilute dCAPE, (b) dilute dCAPE, (c) Bechtold, and (d) HCF.

resolution. The convection frequency does not show significant changes with resolution for the HCF trigger.

Considering the decreased convection frequency in the CRM data with resolution, both the increased (Figs. 3a,b) and constant frequency changes (Figs. 3c,d) in the trigger functions lead to an increase of overprediction and decrease of the skill. As an example, Fig. 4 shows the frequency changes of correct prediction, overprediction, and underprediction with resolution in dCAPE-type triggers. At 128 km, the frequency of correct prediction and underprediction is comparable in the undilute dCAPE trigger, and the frequency of underprediction is even higher than that of correct prediction in the dilute dCAPE trigger, leading to the low skill as shown Fig. 1b. This is mainly because the dilution further decreases the dCAPE production and makes it harder to reach the threshold. There is almost no overprediction in both undilute and dilute dCAPE triggers at 128 km. From 64 to 4 km, both the correct prediction and underprediction in dCAPE-type triggers decrease, along with the decreased convection frequency, but the overprediction frequency increases steadily, reaching about 11% at 4 km. Hence, the decrease of the skill with the resolution is mainly due to the increase of overprediction.

In the following analysis, the undilute dCAPE trigger is chosen to investigate the reason for the scale dependence of trigger function and how to develop a scale-aware trigger function. The cumulative probability distribution function (PDF) of dCAPE in the undilute dCAPE trigger at different resolutions is displayed in Fig. 5. The abscissa shows the dCAPE value, and the ordinate shows the corresponding cumulative probability of dCAPE greater than this value. From -300 to $-60\,\mathrm{J\,kg^{-1}\,h^{-1}}$, the cumulative probability of dCAPE is the highest at 128 km, and lowest at 4 km. With the increase of dCAPE, the situation is reversed gradually. When dCAPE is greater than $65\,\mathrm{J\,kg^{-1}\,h^{-1}}$, the corresponding probability is the largest at 4 km and lowest at 128 km. Hence, as resolution increases, the probability of high dCAPE values becomes larger, corresponding to higher predicted convection frequency. The dCAPE is closely related to the vertical velocity (Song and Zhang 2017), and further analysis on cumulative PDF of vertical velocity at 500 hPa at different resolutions shows similar results (figure not shown). This is because for larger averaging subdomains, the large values of smaller-scale upward and downward motions cancel each other to produce a small subdomain-mean vertical motion. The dilute dCAPE trigger shows similar dCAPE distribution changes with resolution (figure not shown). This explains why predicted convection frequency in the dCAPE-type triggers become larger at higher resolutions. For the HCF trigger, since it is based

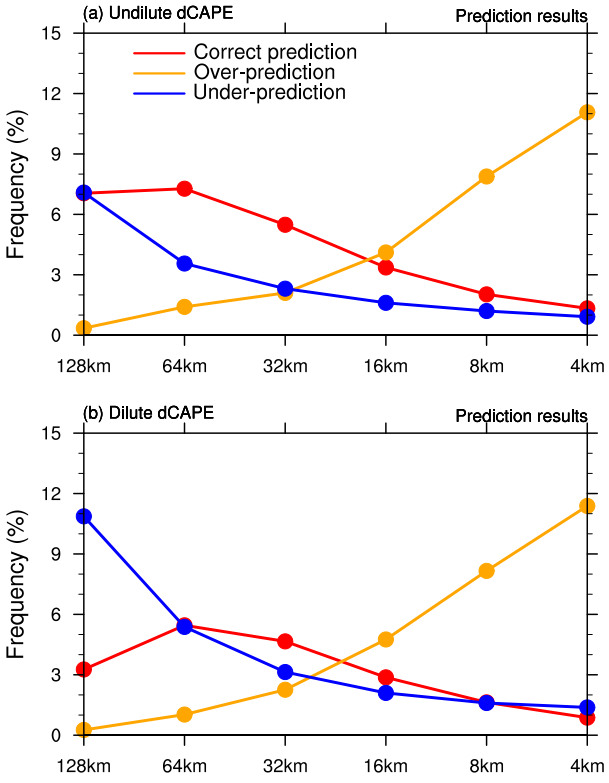

FIG. 4. The frequency of correct prediction (red line, unit: %), overprediction (orange line, unit: %) and underprediction (blue line, unit: %) at different resolutions for the (a) undilute dCAPE and (b) dilute dCAPE triggers.

on near-surface temperature and moisture, which do not show significant changes with resolution, convection frequency does not exhibit substantial variation in Fig. 3d. However, the Bechtold trigger also considers vertical velocity at LCL, so why does its convection frequency remain almost unchanged with resolution? Often, after satisfying the trigger function threshold, the parcel also needs to reach enough height (e.g., 3 km adopted here) to be considered as deep convection, known as the cloud depth constraint. For the dCAPE-type triggers, once the dCAPE threshold is satisfied, the cloud depth is almost always higher than 3 km in each resolution. However, this is not the case for the Bechtold trigger. With the resolution increase, it is more often to be shallower than 3 km after satisfying the temperature requirement. This is because with the increase of resolution, the increased vertical velocity offers higher temperature perturbation δT to the parcel according to Eq. (2). As a result, among the parcels that satisfy the temperature requirement, the fraction of colder and drier parcels is increased, which will be lifted to a lower depth before reaching neutral buoyancy level. Hence, although the frequency of high temperature perturbation δT, which is based on the vertical velocity at the LCL, is indeed increased with resolution (figure not shown), the cloud depth shows opposite effects and these two effects cancelled each other to a large degree. For the Bechtold trigger with fixed Δx, the temperature perturbation δT keeps almost unchanged with the resolution, since the effects of increased vertical velocity and smaller area A in higher resolutions on δT almost cancelled out. In this setting, it keeps constant for the frequency of cloud depth greater than 3 km after satisfying the temperature requirement, so the convection frequency for the Bechtold trigger with fixed Δx also keeps almost unchanged as the resolution increases.

b. *Developing a scale-aware trigger function*

To improve the trigger function, the overprediction and underprediction events need to be reduced to increase the number of correct predictions. Since dCAPE is key to determining the prediction results in the undilute dCAPE trigger function, the investigation of its distributions in the prediction outcome (correct prediction, overprediction, underprediction, and correct prediction of no convection) will give us some insights into how to improve it. According to the four categories of prediction at each resolution, we obtain the corresponding dCAPE values and show them in the box-and-whisker plots in Fig. 6. In the correct prediction, the dCAPE median value and main body (between 25th and 75th percentile) become higher with increasing resolution, with its 25th percentile from 81 J kg^{-1} h^{-1} at 128 km to 407 J kg^{-1} h^{-1} at 4 km. dCAPE in the overprediction also becomes higher in the higher resolutions, but with much slower rate: its 75th percentile only increases from 86 J kg^{-1} h^{-1} at 128 km to 181 J kg^{-1} h^{-1} at 4 km. Hence, dCAPE in the correct prediction and overprediction can be separated from each other to a large degree: the 25th percentile in the correct prediction is much larger than the 75th percentile in the overprediction at all resolutions except 128 km. At 128 km, the overprediction frequency is close to 0 (Fig. 4a), so it does not have a large effect on the skill. The clear separation of dCAPE in these two predictions suggests that the dCAPE threshold can be increased at higher resolutions to reduce the overprediction without much sacrifice of correct prediction. In contrast to correct prediction, dCAPE in the underprediction changes in the opposite direction: the main body shifts from 0–52 to −352–0 J kg^{-1} h^{-1} when varying from 128 to 4 km. In other words, convection events at higher resolutions often have larger dCAPE range. This also means that the relationship between dCAPE and convection becomes worse in higher resolutions (0.61 and 0.47 at 128 km and 4 km, respectively).

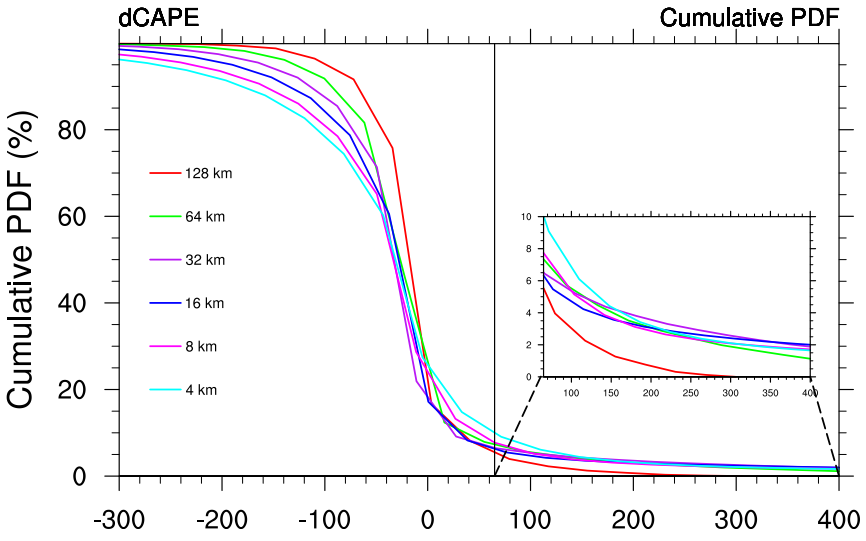

FIG. 5. The cumulative probability distribution of frequency of dCAPE (unit: J kg^{-1} h^{-1}) in the undilute dCAPE trigger at different resolutions. The percentage in the ordinate shows the probability of dCAPE greater than the corresponding abscissa values. The subfigure shows the zoom-in for the dCAPE between 65 and 400 J kg^{-1} h^{-1}.

To improve the trigger function, the vertical structures of the atmospheric states for different predictions also need to be examined. Figure 7 shows the composites of temperature, moisture, vertical velocity, DSE advection, and moisture advection in four predictions at 8-km model resolution. The composites in other resolutions show similar structures, except that the vertical velocity and advection terms have larger magnitudes at higher resolutions. The low-level cold anomalies are most evident in the correct prediction and underprediction. The moisture anomaly is positive below 600 hPa in the overprediction, but it can extend to much higher levels in both the correct prediction and underprediction (Fig. 7b). The correct prediction has the highest moisture anomalies below 400 hPa, but near the surface, the moisture in these three predictions is comparable. Considering the much colder temperature in the correct prediction and underprediction (Fig. 7a), the

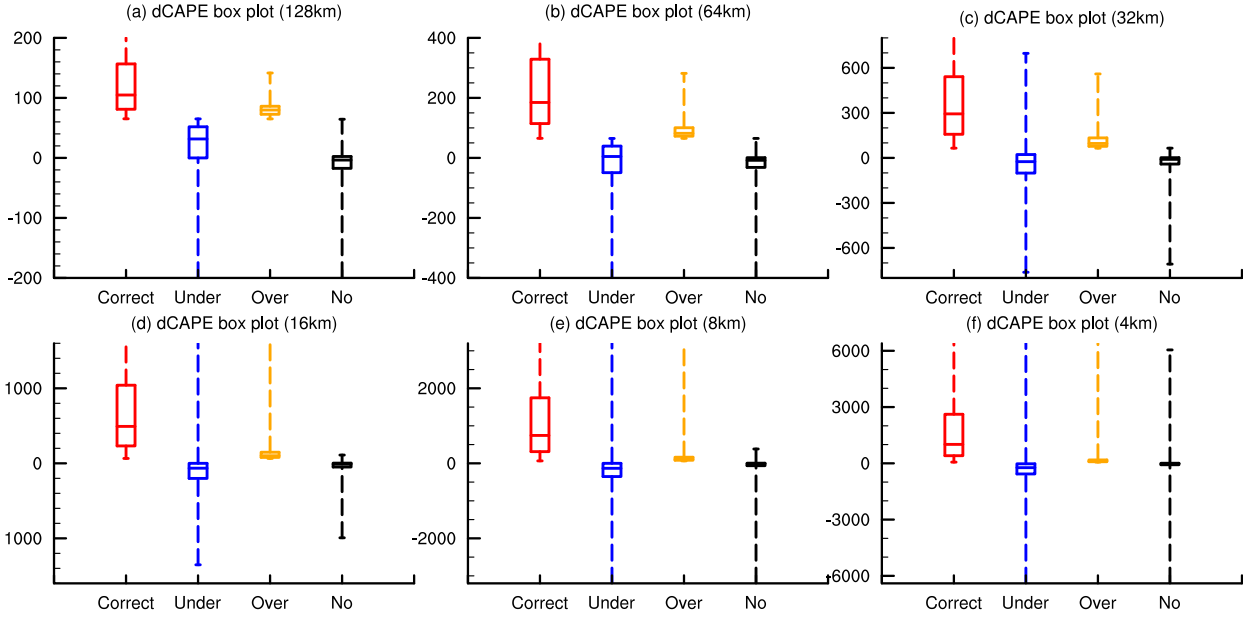

FIG. 6. The box-and-whisker plot of dCAPE (unit: J kg^{-1} h^{-1}) in the undilute dCAPE trigger for correct prediction (red), underprediction (blue), overprediction (orange), and correct prediction of no-convection (black) at (a) 128, (b) 64, (c) 32, (d) 16, (e) 8, and (f) 4 km.

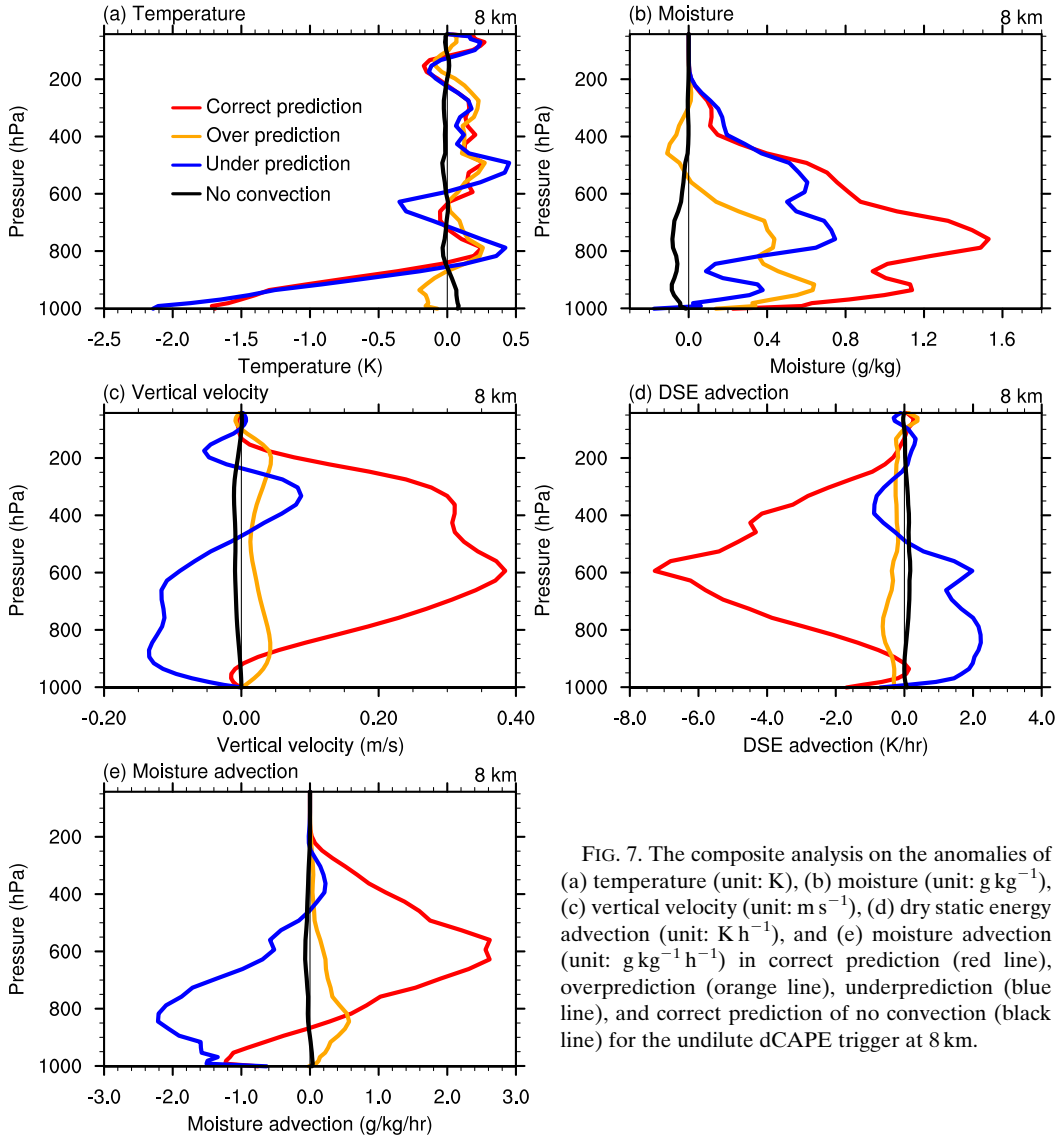

FIG. 7. The composite analysis on the anomalies of (a) temperature (unit: K), (b) moisture (unit: g kg^{-1}), (c) vertical velocity (unit: m s^{-1}), (d) dry static energy advection (unit: K h^{-1}), and (e) moisture advection (unit: g kg^{-1} h^{-1}) in correct prediction (red line), overprediction (orange line), underprediction (blue line), and correct prediction of no convection (black line) for the undilute dCAPE trigger at 8 km.

near-surface relative humidity is much higher when convection occurs. Some previous studies also adopted relative humidity as one of the trigger function thresholds (Sud and Walker 1999; Zhang 2002), but it is unclear whether the inclusion of relative humidity can reduce the scale dependence. Here, when a relative humidity threshold (RH $>$ 80%) is required in the undilute dCAPE trigger at different resolutions, it is found that the HSS can be increased slightly but almost equally at different resolutions (figure not shown). When we turn to the vertical velocity, it is found that the vertical profiles of vertical velocity among the four predictions are considerably different (Fig. 7c). In the correct prediction and overprediction, the upward motion dominates the whole troposphere, but with much weaker magnitude in the overprediction. In the underprediction, the downward motion occurs below 450 and above 250 hPa, and between them is the upward motion. Consistent with Song and Zhang (2017) based on GOAmazon data, the vertical velocity dominates the DSE and moisture advection changes (Figs. 7d,e).

Considering the vital role of vertical velocity in the DSE and moisture advection, which are the two contributors to dCAPE [Eq. (1)], it is important to understand the vertical structure of vertical velocity in different predictions. The subdomain average vertical velocity $\overline{w}$ consists of contributions from three components: the convective updraft part $\sigma_u w_u$, the convective downdraft part $\sigma_d w_d$, and the convection-free environment $(1-\sigma_u-\sigma_d)\tilde{w}$:

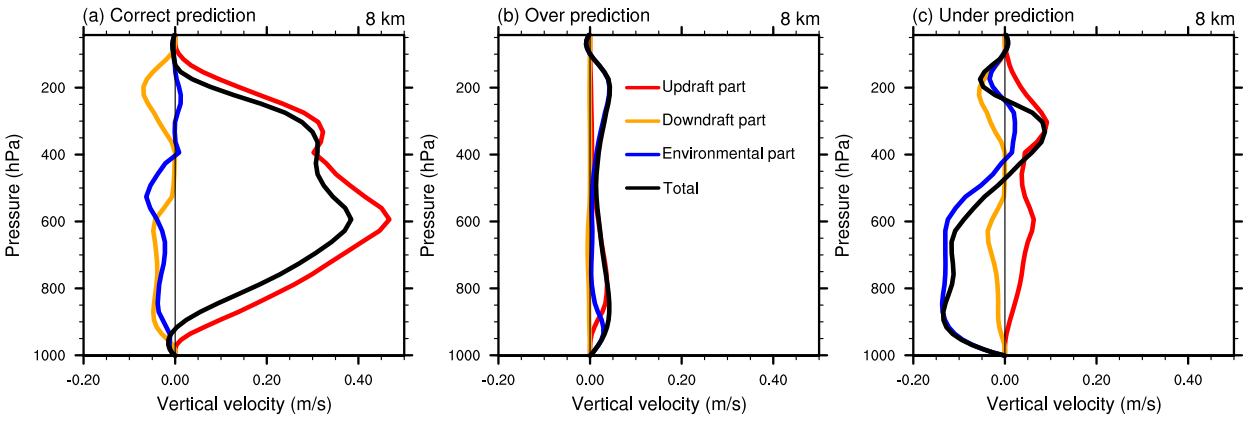

FIG. 8. The updraft part (red line), downdraft part (orange line), environmental part (blue), and total (black line) vertical velocity (unit: m s^{-1}) in (a) correct prediction, (b) overprediction, and (c) underprediction for the undilute dCAPE trigger at 8 km.

$$\overline{w} = \sigma_u w_u + \sigma_d w_d + (1 - \sigma_u - \sigma_d)\tilde{w},$$

where σ_u and σ_d are the updraft and fraction, respectively; w_u, w_d, and $\tilde{w}$ are the vertical velocity averaged over the updraft points, downdraft points, and environmental points, respectively.

By decomposing the vertical velocity into different contributions (updraft, downdraft, and the environment), we can learn the physical nature of different predictions. We select 8-km model resolution as an example, and results from other resolutions are similar. As shown in Fig. 8, convective updraft contributes the most to vertical velocity in the correct prediction (Fig. 8a), representing the mature phase of convection. In the overprediction, the updraft part contributes most in the lower level (between 850 and 500 hPa) and the environmental part dominates in other levels (Fig. 8b). In the underprediction, the domain-averaged downward motion in the lower half of the troposphere (from surface to 500 hPa) mainly comes from the convection-free environment, whereas the upper-level upward motion (from 450 to 250 hPa) is mainly from the convective updraft part. Meanwhile, the downdraft in the underprediction is similar to that in the correct prediction (Figs. 8a and 8c). Through this decomposition, we can obtain some insights into the physical nature of overprediction and underprediction: the overprediction probably occurs at the developing phase of convection, in which the updraft begins to appear in the lower level and there is almost no downdraft, whereas the underprediction occurs at the decaying phase of convection, in which the downdraft is still active, but the updraft is only evident in the upper level.

To confirm this, lead/lag composites of vertical velocity and mass flux in three predictions are shown in Fig. 9. Lag zero corresponds to the time of a given prediction by the trigger function. For correct prediction, it corresponds to the mature phase of convection, with strong upward motion and mass flux in the whole troposphere, indicating that the convective cloud is dominating. The convection event begins at the low level about 1 h before the correct prediction, and gradually develops from the low level to the upper level, with the gradual development of convective clouds. At the time of correct prediction, the upward motion reaches the maximum at 600 hPa, and downward motion begins to appear at the low level. Later, the downward motion extends to the midtroposphere, while the upward motion remains for about 1.5 h in the upper troposphere. Convective mass flux (sum of both updraft and downdraft) is similar to the domain-averaged vertical velocity, with the maximum developing at lower levels and gradually moving up as time progresses. The difference between the vertical velocity and mass flux in the low level after the time of correct prediction shows that the downward motion is mainly convection free, along with the stratiform clouds.

At the time of overprediction, there is upward motion throughout the troposphere (Fig. 9c). However, since the upward motion in the upper level is mostly from convection-free areas (Fig. 8b), there is little resemblance between the domain-averaged upward motion and convective mass flux (Figs. 9c,d). About a half hour prior to the time of overprediction, upward motion develops in the lower levels with downward motion aloft. After the time of overprediction, upward motion dominates the entire troposphere, with the center of maximum motion gradually moving up to the upper troposphere. Both of these patterns are found in the convective mass flux (Fig. 9d), suggesting the development of convective clouds during this period. The composite time evolution of convective mass flux from the CRM simulation

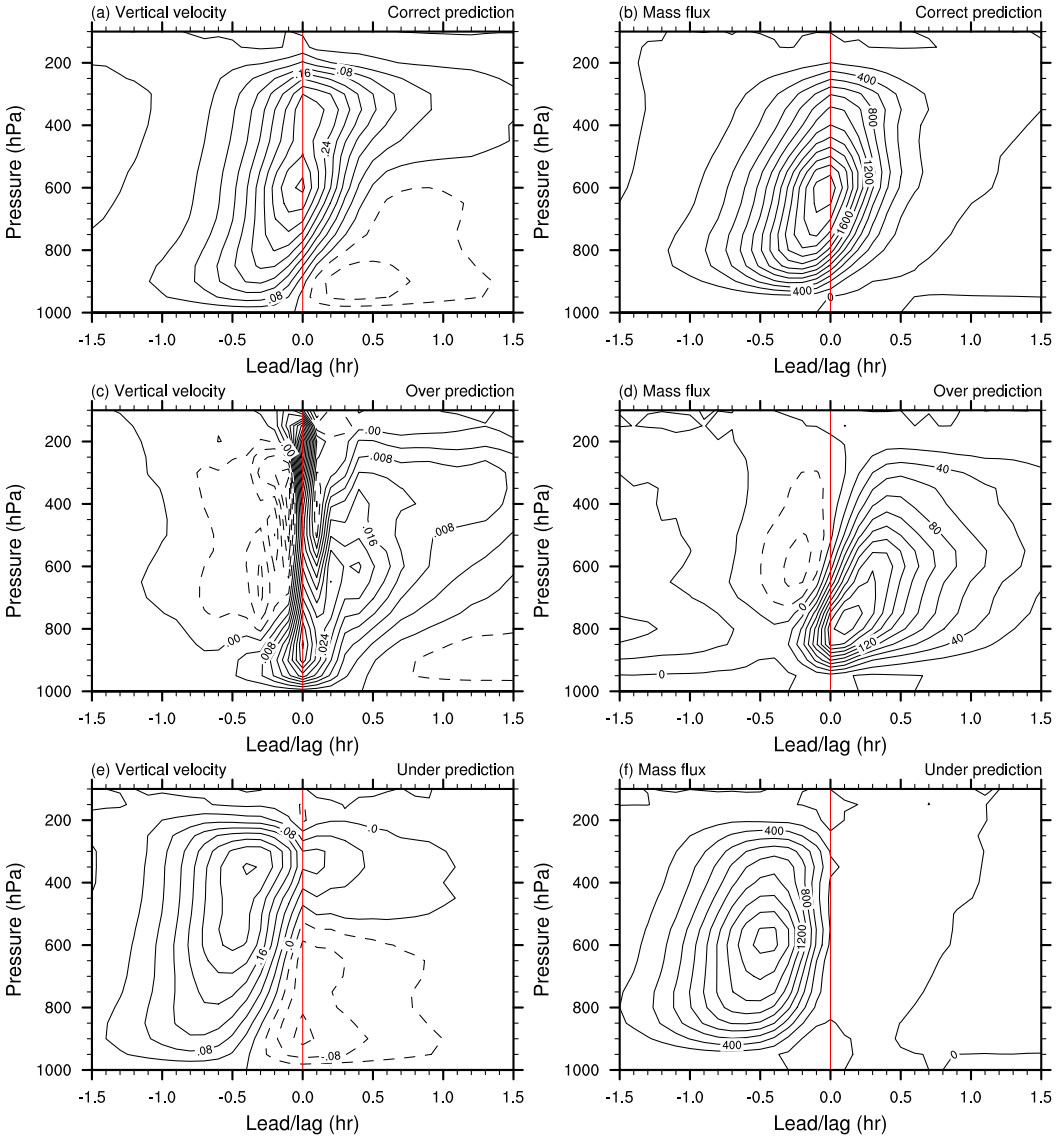

FIG. 9. The lead–lag composite analysis for underprediction for the undilute dCAPE trigger at 8 km of (left) vertical velocity (unit: $m\,s^{-1}$) and (right) mass flux (unit: $kg\,m^{-2}\,s^{-1}$) for (top) correct prediction, (middle) overprediction, and (bottom) underprediction.

signifies the early development stage of convection, likely in shallow convection phase. This corroborates the earlier studies showing that convection schemes tend to predict convection too early and fail to capture the shallow convection stage (Guichard et al. 2004; Xu et al. 2005). At the time of underprediction, there is strong downward motion below 500 hPa and upward motion above. Within the previous 1.5 h, there is strong domain-averaged upward motion and convective mass flux, signifying the strong convection. The composite time evolution suggests that it corresponds to the decay stage of convection. Since the convection-free environmental motion contributes the most to the vertical velocity (Fig. 8), the convective mass flux is weak, which suggests the dominant clouds have changed from the convective clouds to the stratiform clouds at the time of underprediction.

As discussed above, the correct prediction corresponds to the mature phase of convection, while overprediction and underprediction are corresponding to the developing and decaying phases of convection, respectively. This is also the case at other resolutions. It means, regardless of resolutions, the current undilute dCAPE trigger overemphasizes the developing phase, but does not consider the decaying phase well. To reduce the overprediction, we can increase the dCAPE threshold as 25th percentile of dCAPE in the correct

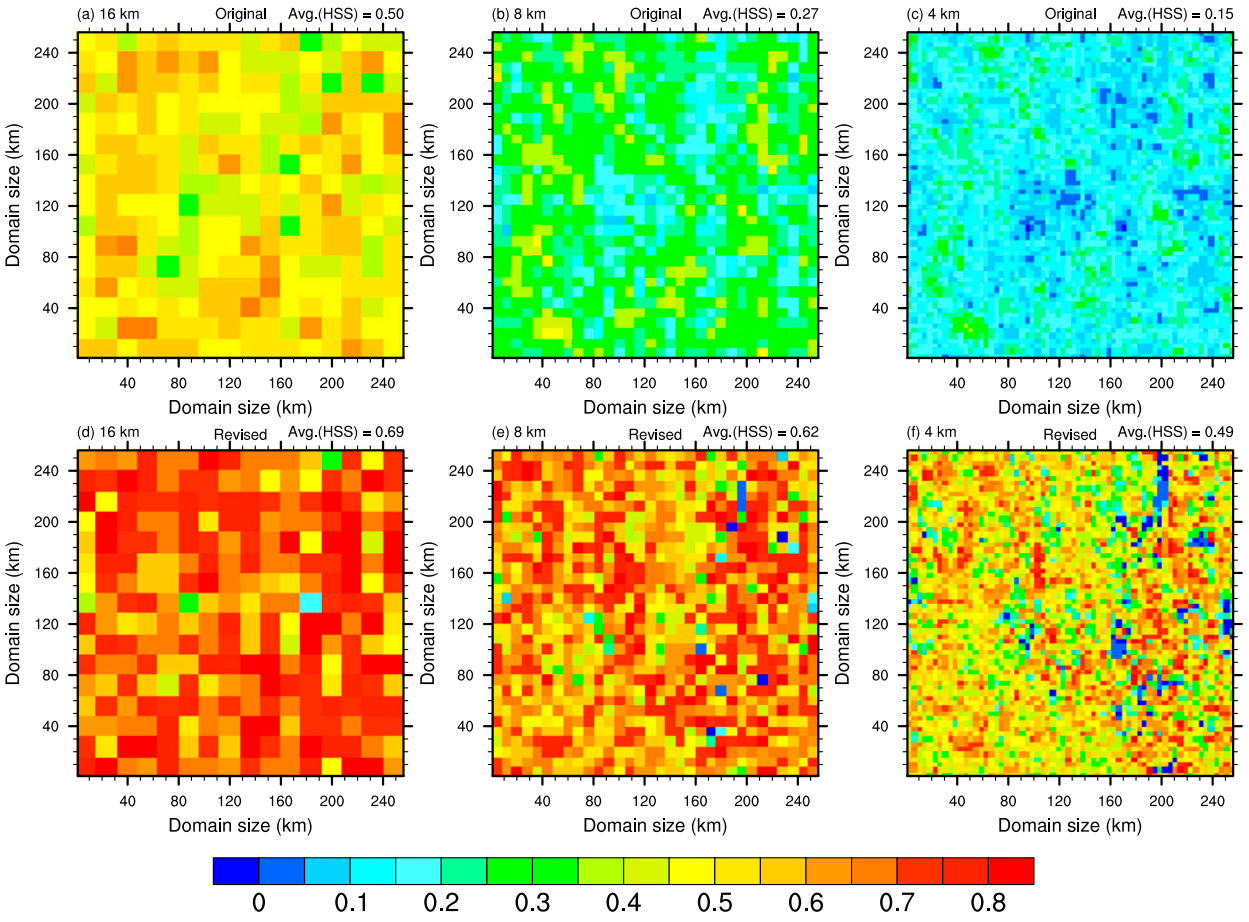

FIG. 10. The spatial distribution of HSS for the (top) original and (bottom) revised undilute dCAPE trigger at (a),(d) 16; (b),(e) 8; and (c),(f) 4 km. The number at the right corner represents the average of HSS.

prediction at different resolutions. To reduce the underprediction, we need to consider the previous state or history of convection. If convection was predicted at the preceding time step (6 min here), then the dCAPE threshold is reduced to 0 at the current time step for triggering convection. In effect, this modification increases the chance of convection in the decaying phase being predicted. We should add a note of caution here that setting dCAPE threshold to zero after convection onset could potentially lead to perpetual convection (gridpoint storm) in the full GCM because of the feedback between convection and vertical motion, although this did not happen in past simulations (Zhang and Mu 2005b). These two revisions are applied to the three high resolutions (16, 8, and 4 km) to improve their performances. The horizontal distribution of HSS for the original and revised undilute dCAPE trigger at three different resolutions is shown in Fig. 10. For the original trigger, the HSS decreases rapidly from 0.50 at 16 km to 0.15 at 4 km. After the revision, the skill is increased from 0.50 to 0.70 at 16 km, from 0.27 to 0.61 at 8 km, and from 0.15 to 0.53 at 4 km. To examine the possible influence of data independence on the results, we also divide the dataset in two parts: the first half is used to obtain the 25th percentile of dCAPE in the correct prediction as the dCAPE threshold and the second half is used to evaluate the skill after the revision. It is found that the skill of the revised trigger is 0.69, 0.62, and 0.49 for 16 km, 8 km, and 4 km, respectively, comparable to the skill shown in Fig. 10. The revised undilute dCAPE trigger also works better than the original one at the lower resolutions. For the domain average, the skill is improved from the original 0.62, 0.71, and 0.69 to 0.70, 0.75, and 0.75 for 128 km, 64 km, and 32 km, respectively. Hence, this revision can be used across the resolutions. We also looked into the relative contributions of two revisions to the improvement. It is found the HSS can be increased to 0.62, 0.55, and 0.48 for 16 km, 8 km, and 4 km, respectively, by only increasing the dCAPE threshold. Hence, the dCAPE threshold has the most important impact on the improvement,

and the convection indicator plays a secondary role. Similar revisions also apply to the dilute dCAPE trigger well (figure not shown).

4. Conclusions and discussion

As the resolution of global climate models (GCMs) increases, whether current trigger functions of the convection scheme, which are often designed for coarse resolution, still work remains unknown. In this study, we investigate this issue based on undilute dCAPE, dilute dCAPE, Bechtold, and HCF trigger functions by using cloud-resolving model data (CRM). The 1-km CRM data are averaged over 128-, 64-, 32-, 16-, 8-, and 4-km subdomains to represent the GCM resolutions. It is found that all these four trigger functions are scale dependent. For dCAPE-type triggers, they perform well at the lower resolutions, including 128, 64, and 32 km, with the HSS skill of ~0.60, except for the dilute dCAPE trigger at 128 km. However, the skill decreases rapidly to only 0.14 and 0.08 at 4 km for undiluted and dilute dCAPE triggers, respectively. For Bechtold and HCF triggers, they do not perform well across the resolutions, due to excessive overpredictions. The scale dependence of these trigger functions is attributed to the different convection frequency changes with the resolution from the CRM. The actual convection frequency decreases gradually from 14.1% at 128 km to 2.3% at 4 km as the domain becomes smaller at higher resolutions. However, the convection frequency increases with the resolution in the dCAPE-type trigger, and remains almost unchanged in the Bechtold and HCF triggers. In both situations, the overprediction becomes more with the resolution and deteriorates the skills.

As the undilute dCAPE trigger has the most scale dependence, it is selected to understand why the predicted convection frequency becomes higher and how to develop a scale-aware trigger function. The undilute dCAPE trigger is mainly based on dCAPE, which represents the large-scale free tropospheric forcing. It is found that as resolution increases, the high dCAPE values occur more often. The vertical velocity dominates the dCAPE change, and as the resolution increases, the high upward motion also becomes more often. The frequency of dCAPE greater than the threshold ($65\,J\,kg^{-1}\,h^{-1}$) is increased from 7.4% at 64 km to 12.4% at 4 km. The box-and-whisker plot of dCAPE shows that the main body (between 25th percentile and 75th percentile) of correct prediction and overprediction can be separated from each other across different resolutions, so increasing the dCAPE threshold in the trigger can reduce large overprediction without much loss of correct prediction.

The composite analysis suggests that the vertical velocity exhibits different structures among four predictions, and this difference holds across the resolutions. Both the correct prediction and overprediction show the upward motion throughout the troposphere, but with much larger magnitude in the correct prediction. The upward motion only occurs between 450 and 250 hPa in the underprediction and the rest are occupied by the downward motion. By decomposing the vertical velocity into updraft, downdraft, and environmental components, it is found that the convective updraft contributes most in the correct prediction. The convective updraft also dominates the low-level upward motion in the overprediction and upper-level upward motion in the underprediction. The upper-level upward motion in the overprediction and downward motion in the underprediction are mostly convection free. Meanwhile, the downdraft components in the correct prediction and underprediction are very similar. The lead–lag composite analysis on the vertical velocity and mass flux further confirms that the overprediction and underprediction correspond to the developing and decaying phases of convection, respectively. Based on these results, a revision on the undilute dCAPE trigger is proposed by increasing the dCAPE threshold and considering the convection history. If the previous time step (6 min here) is convective, then the dCAPE threshold is tuned down to zero at the current time step. The increase of dCAPE threshold reduces overpredictions and considering convection history reduces some underpredictions. After the revision, the undilute dCAPE trigger works better than the original one not only at the higher resolutions, but also at the lower resolutions. For the domain average, the HSS skill can reach 0.70, 0.62, and 0.53 at 16 km, 8 km, 4 km, respectively, comparable to those at the lower resolutions.

It is found that the vertical structures of vertical velocity in the overprediction and underprediction are different from those obtained from the Amazon region (Song and Zhang 2017). Over the Amazon region, the underprediction exhibits bottom-heavy upward motion in all the levels, while the overprediction exhibits low-level downward motion and upper-level upward motion. Based on this structural contrast, Song and Zhang (2017) proposed to include the low-level upward motion to improve the undilute dCAPE over the Amazon region. Obviously, this improvement does not apply to the western Pacific warm pool region here, underlining the difficulty in convection parameterization in global models. Meanwhile, we only used the CRM data forced by large-scale forcing from one location here. This may limit our conclusions to some extent. For example, the extent to which the trigger functions are scale

dependent may be location dependent. Whether the increase of dCAPE threshold in the dCAPE-type trigger to improve its scale awareness applies to other locations also needs to be examined. However, the fact that these four trigger functions are scale dependent is not location dependent, because the mechanism responsible for the scale dependence is not dependent on location qualitatively. More importantly, we have proposed a framework to understand the scale awareness of trigger functions used in the global climate models. This is an important and unanswered question, especially for global climate models whose resolution is greatly increased but not high enough to resolve convection explicitly.

Acknowledgments. This material is based upon work supported by the National Science Foundation Grant AGS-1549259 and the Department of Energy, Office of Science, Biological and Environmental Research Program (BER), under Award DE-SC0016504, and by the Scientific Discovery through Advanced Computing (SciDAC) program of the U.S. Department of Energy (DOE) Office of Advanced Scientific Computing Research and Office of Biological and Environmental Research through PNNL Subcontract 190110. We thank Prof. Courtney Schumacher and two anonymous reviewers for their helpful comments and suggestions. We also thank Drs. Wei-Kuo Tao and Xi-Ping Zeng for providing us with their CRM simulation data. The data used for this analysis are available from the authors upon request.

REFERENCES

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part 1. *J. Atmos. Sci.*, **31**, 674–701, https://doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2.

——, and C.-M. Wu, 2013: A unified representation of deep moist convection in numerical modeling of the atmosphere. Part I. *J. Atmos. Sci.*, **70**, 1977–1992, https://doi.org/10.1175/JAS-D-12-0330.1.

——, J.-H. Jung, and C.-M. Wu, 2011: Toward unification of the multiscale modeling of the atmosphere. *Atmos. Chem. Phys.*, **11**, 3731–3742, https://doi.org/10.5194/acp-11-3731-2011.

Bechtold, P., E. Bazile, F. Guichard, P. Mascart, and E. Richard, 2001: A mass flux convection scheme for regional and global models. *Quart. J. Roy. Meteor. Soc.*, **127**, 869–886, https://doi.org/10.1002/qj.49712757309.

——, J.-P. Chaboureau, A. Beljaars, A. K. Betts, M. Kohler, M. Miller, and J.-L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model. *Quart. J. Roy. Meteor. Soc.*, **130**, 3119–3137, https://doi.org/10.1256/qj.03.103.

Fridlind, A. M., and Coauthors, 2012: A comparison of TWP-ICE observational data with cloud-resolving model results. *J. Geophys. Res.*, **117**, D05204, https://doi.org/10.1029/2011JD016595.

Guichard, F., and Coauthors, 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quart. J. Roy. Meteor. Soc.*, **130**, 3139–3172, https://doi.org/10.1256/qj.03.145.

Hogan, R. J., E. J. O'Connor, and A. J. Illingworth, 2009: Verification of cloud fraction forecasts. *Quart. J. Roy. Meteor. Soc.*, **135**, 1494–1511, https://doi.org/10.1002/qj.481.

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

Lee, M.-I., and Coauthors, 2007: An analysis of the warm-season diurnal cycle over the continental United States and northern Mexico in general circulation models. *J. Hydrometeor.*, **8**, 344–366, https://doi.org/10.1175/JHM581.1.

——, S. D. Schubert, M. J. Suarez, J.-K. E. Schemm, H.-L. Pan, J. Han, and S.-H. Yoo, 2008: Role of convection triggers in the simulation of the diurnal cycle of precipitation over the United States Great Plains in a general circulation model. *J. Geophys. Res.*, **113**, D02111, https://doi.org/10.1029/2007JD008984.

Liu, Y.-C., J. Fan, G. J. Zhang, K.-M. Xu, and S. J. Ghan, 2015: Improving representation of convective transport for scale-aware parameterization: 2. Analysis of cloud-resolving model simulations. *J. Geophys. Res. Atmos.*, **120**, 3510–3532, https://doi.org/10.1002/2014JD022145.

Neale, R. B., J. H. Richter, and M. Jochum, 2008: The impact of convection on ENSO: From a delayed oscillator to a series of events. *J. Climate*, **21**, 5904–5924, https://doi.org/10.1175/2008JCLI2244.1.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Song, F., and G. J. Zhang, 2017: Improving trigger functions for convective parameterization schemes using GOAmazon observations. *J. Climate*, **30**, 8711–8726, https://doi.org/10.1175/JCLI-D-17-0042.1.

Sud, Y. C., and G. K. Walker, 1999: Microphysics of clouds with the Relaxed Arakawa–Schubert Scheme (McRAS). Part I: Design and evaluation with GATE Phase III data. *J. Atmos. Sci.*, **56**, 3196–3220, https://doi.org/10.1175/1520-0469(1999)056<3196:MOCWTR>2.0.CO;2.

Suhas, E., and G. J. Zhang, 2014: Evaluation of trigger functions for convective parameterization schemes using observations. *J. Climate*, **27**, 7647–7666, https://doi.org/10.1175/JCLI-D-13-00718.1.

——, and ——, 2015: Evaluating convective parameterization closures using cloud-resolving model simulation of tropical deep convection. *J. Geophys. Res. Atmos.*, **120**, 1260–1277, https://doi.org/10.1002/2014JD022246.

Tao, W.-K., and J. Simpson, 1993: The Goddard Cumulus Ensemble Model. Part I: Model description. *Terr. Atmos. Oceanic Sci.*, **4**, 19–54, https://doi.org/10.3319/TAO.1993.4.1.35(A).

Tawfik, A. B., and P. A. Dirmeyer, 2014: A process-based framework for quantifying the atmospheric preconditioning of surface-triggered convection. *Geophys. Res. Lett.*, **41**, 173–178, https://doi.org/10.1002/2013GL057984.

Xiao, H., W. I. Gustafson, S. M. Hagos, C.-M. Wu, and H. Wan, 2015: Resolution-dependent behavior of subgrid-scale vertical transport in the Zhang–McFarlane convection parameterization. *J. Adv. Model. Earth Syst.*, **7**, 537–550, https://doi.org/10.1002/2014MS000356.

Xie, S., and M. Zhang, 2000: Impact of the convection triggering function on single-column model simulations. *J. Geophys. Res.*, **105**, 14 983–14 996, https://doi.org/10.1029/2000JD900170.

——, J. S. Boyle, R. T. Cederwall, G. L. Potter, and W. Lin, 2004: Impact of a revised convective triggering mechanism on

Community Atmosphere Model, version 2 simulations: Results from short-range weather forecasts. *J. Geophys. Res.*, **109**, D14102, https://doi.org/10.1029/2004JD004692.

Xu, K.-M., and Coauthors, 2005: Modeling springtime shallow frontal clouds with cloud-resolving and single-column models. *J. Geophys. Res.*, **110**, D15S04, https://doi.org/10.1029/2004JD005153.

Yun, Y., J. Fan, H. Xiao, G. J. Zhang, S. J. Ghan, K.-M. Xu, P.-L. Ma, and W. I. Gustafson Jr., 2017: Assessing the resolution adaptability of the Zhang–McFarlane cumulus parameterization with spatial and temporal averaging. *J. Adv. Model. Earth Syst.*, **9**, 2753–2770, https://doi.org/10.1002/2017MS001035.

Zeng, X., W.-K. Tao, T. Matsui, S. Xie, S. Lang, M. Zhang, D. Starr, and X. Li, 2011: Estimating the ice crystal enhancement factor in the tropics. *J. Atmos. Sci.*, **68**, 1424–1434, https://doi.org/10.1175/2011JAS3550.1.

Zhang, G. J., 2002: Convective quasi-equilibrium in midlatitude continental environment and its effect on convective parameterization. *J. Geophys. Res.*, **107**, 4220, https://doi.org/10.1029/2001JD001005.

——, 2003: Roles of tropospheric and boundary layer forcing in the diurnal cycle of convection in the U.S. southern great plains. *Geophys. Res. Lett.*, **30**, 2281, https://doi.org/10.1029/2003GL018554.

——, and M. Mu, 2005a: Simulation of the Madden–Julian oscillation in the NCAR CCM3 using a revised Zhang–McFarlane convection parameterization scheme. *J. Climate*, **18**, 4046–4064, https://doi.org/10.1175/JCLI3508.1.

——, and ——, 2005b: Effects of modifications to the Zhang–McFarlane convection parameterization on the simulation of the tropical precipitation in the National Center for Atmospheric Research Community Climate Model, version 3. *J. Geophys. Res.*, **110**, D09109, https://doi.org/10.1029/2004JD005617.

——, J. Fan, and K.-M. Xu, 2015: Comments on ''A unified representation of deep moist convection in numerical modeling of the atmosphere. Part I.'' *J. Atmos. Sci.*, **72**, 2562–2565, https://doi.org/10.1175/JAS-D-14-0246.1.