The Impact of Variable Land-Atmosphere Coupling on Convective Cloud Populations Observed During the 2016 HI-SCALE Field Campaign

Jerome D. Fast1, Larry K. Berg1, Zhe Feng1, Fan Mei1, Rob Newsom1, Koichi Sakaguchi1, and Heng Xiao1

1Pacific Northwest National Laboratory, Richland, WA, USA

Abstract We use a large-eddy simulation model with a nested domain configuration (297 and 120 km wide) and an interactive land surface parameterization to simulate the complex population of shallow clouds observed on 30 August 2016 during the Holistic Interactions of Shallow Clouds, Aerosols, and Land-Ecosystems campaign conducted in north-central Oklahoma. Shallow convective clouds first formed over southeast Oklahoma and then spread toward the northwest into southern Kansas. By the early afternoon, the relatively uniform shallow cloud field became more complex in which some regions became nearly cloud free and in other regions larger shallow clouds developed with some transitioning to deeper, precipitating convection. We show that the model reproduces the observed heterogeneity in the cloud populations only when realistic variations in soil moisture are used to initialize the model. While more variable soil moisture and to a lesser extent cool lake temperatures drive the initial spatial heterogeneity in the cloud populations, precipitation-driven cold pools become an important factor after 1300 CST. When smoother soil moisture variations are used in the model, more uniform shallow cloud populations are predicted with far fewer clouds that transition to deeper, precipitating convection that produce cold pools. An algorithm that tracks thousands of individual cumulus show that the more realistic soil moisture distributions produces clouds that are larger and have a longer lifetime. The results suggest that shallow and deep convection parameterizations used by mesoscale models need to account for the effects of variable land-atmosphere interactions and cold pools.

Plain Language Summary Models that resolve boundary layer turbulence and clouds have been used extensively to better understand processes controlling the life cycle of shallow convective clouds. Because of their computational expense these models usually employ relatively small domain sizes, usually less than 50 km wide, and constrain the atmospheric forcing so that the predicted clouds are nearly uniformly distributed in space. We employ a more realistic modeling approach to represent the observed complex cloud distributions over Oklahoma on a day during a recent atmospheric sampling campaign. The model results suggest that the observed cloud distributions were caused by two processes that can occur at the same time. During the morning, spatial variations in soil moisture perturbed the heating patterns and altered the timing and intensity of cloud formation. Then during the afternoon cold pools further perturbed the clouds, producing clear skies in some regions and enhancing cloud formation in other regions. While cooler air from evaporating precipitation suppresses cloud formation, cloud formation is enhanced at the edges of the expanding pool of colder air. These processes are missing or treated simply in numerical representations of convective clouds by operational forecasting and climate models that may affect predictions of cloudiness and initiation of precipitation.

1. Introduction

Understanding the evolution of convective cloud populations is critical to accurately simulate the impact of clouds on large-scale dynamical and radiative processes. Gaps in the basic understanding of cloud population dynamics lie at the root of many model biases that span a wide range of spatiotemporal scales, from the diurnal cycle to intraseasonal variations and the global energy budget. At the smallest scales, simulating the location and timing of convective initiation is important, yet difficult for both weather and climate models, particularly when low-level convergence contributes to convection under conditions of weak synoptic forcing. In addition to knowledge gaps on cloud population dynamics, spatial resolution is a

Key Points:
- A relatively large-eddy simulation domain is used to reproduce complex convective cloud populations observed during a recent campaign.
- Cloud distributions are initially driven by soil moisture gradients, followed by cool pools as some cumulus transitions to deeper convection.
- The importance of representing two key processes, soil moisture gradients and cold pools, in convective parameterizations is highlighted.

Supporting Information:
- Supporting Information S1

Correspondence to:
J. D. Fast
jerome.fast@pnnl.gov

Citation:
Fast, J. D., Berg, L. K., Feng, Z., Mei, F., Newsom, R., Sakaguchi, K., & Xiao, H. (2019). The impact of variable land-atmosphere coupling on convective cloud populations observed during the 2016 HI-SCALE field campaign. Journal of Advances in Modeling Earth Systems, 11, 2629–2654. https://doi.org/10.1029/2019MS001727

Received 24 APR 2019
Accepted 26 JUL 2019
Accepted article online 31 JUL 2019
Published online 14 AUG 2019

©2019. The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium provided the original work is properly cited. This manuscript has been co-authored by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830 with the U.S. Department of Energy. The US Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).
contributing factor that contribute to model biases even for so-called “cloud-system resolving models” ($\Delta x$–few kilometers). For example, the horizontal and vertical grid spacing used by models is not likely to be sufficient represent the formation and gradual development of shallow cumulus and the transition to deeper, precipitating convection. Many aspects of mature deep convective systems, such as convective updraft and downdraft intensities, remain poorly resolved as well (Fan et al., 2017; Varble et al., 2014).

Since convective clouds are largely a subgrid-scale phenomenon for global and many regional-scale models, parameterizations are used to represent the vertical redistribution of heat and moisture, precipitation, vertical transport of scalars, and cloud radiative effects associated with convective clouds. Shallow and deep convective clouds are often treated by separate parameterizations. It is common to evaluate shallow convective parameterizations using mean meteorological profiles on “golden days” derived from large-eddy simulation (LES) model predictions (Moeng & Sullivan, 2015; Wyngaard, 1984, 2010) that explicitly resolve both boundary layer turbulence and clouds. Relatively few golden days have been developed for continental cumulus and marine stratocumulus regimes since the observational constraints for the LES models need to be carefully chosen. The Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) shallow cumulus case on 21 June 1997 (ARM97; Brown et al., 2002) is one example that has been widely used over the past two decades as a benchmark for continental shallow cumulus parameterizations (e.g., Bogenschutz & Krueger, 2013; Bogenshultz et al., 2012; Cheng & Xu, 2015; Gentine et al., 2013; Neggers et al., 2009).

Nevertheless, a single case does not robustly evaluate parameterizations that need to represent shallow clouds under a range of atmospheric conditions. Zhang et al. (2017) recently proposed that a composite approach should be used by LES studies when used in conjunction with parameterization evaluation. Their composite shallow cumulus day was developed using SGP measurements on 76 golden days over a 13-year period that were not significantly influenced by synoptic-scale conditions. In contrast, the LES ARM Symbiotic Simulation and Observation workflow (Gustafson et al., 2016, 2017a, 2017b) has recently produced a library of LES model output and concurrent observations, skill scores, diagnostics, and quick-look plots for 20 golden days of shallow cumulus during the summers of 2015 and 2016. Instead of a single LES simulation, each “golden day” has several LES ensemble members based on different large-scale forcing methodologies. This permits parameterizations to be evaluated against LES model output under a wider range of atmospheric conditions that also accounts for uncertainties in large-scale forcing.

Traditional LES studies use small model domains that are typically less than 50 km wide, have no variations in topography, prescribe surface fluxes, and rely on cyclic lateral boundary conditions. Large-scale forcing is prescribed by profiles of mean temperature, humidity, and winds as well as subsidence that can vary in time. Consequently, the studies represent idealized conditions in which shallow cumulus populations are spatially homogeneous. Predictions from LES models are usually evaluated by computing profiles of domain-wide statistics that are compared with measurements at a single site.

Despite the knowledge gained from traditional LES studies, there are many more challenging atmospheric conditions associated with spatially varying cloud distributions. To address this issue, a nested approach can be used (e.g., Moeng et al., 2007; Talbot et al., 2012). This permits forcing to vary in time and space, and thus, more complex cloud populations, including transitions from shallow to deep convection, can be better simulated. In addition, simulations can include topographic variations (e.g., Liu et al., 2011; Rai et al., 2017) and inhomogeneous surface fluxes. While it is desirable to downscale mesoscale forcing to the LES domains, turbulence will be suppressed at the lateral boundaries (e.g., Mirocha et al., 2014). The distance from the inflow boundaries in which turbulent eddies are artificially suppressed depends on the ambient wind speed and the simulation period. To address this issue, some studies (e.g., Munoz-Esparza et al., 2015, 2017; Munoz-Esparza & Kosovic, 2018) have developed approaches to generate inflow turbulence to compensate for the turbulence suppression when transitioning from mesoscale to LES domains.

Despite the computational cost, a few LES studies have been completed that study more realistic cloud populations using very LES domains. For example, Khairoutdinov and Randall (2006) use the System for Atmospheric Modeling to simulate shallow-to-deep convection transitions over the Amazon during the Large-Scale Biosphere-Atmosphere field campaign using a domain that was ~154 km wide, a 100-m grid spacing, and cyclic boundary conditions. Khairoutdinov et al. (2009) used a similar model configuration to simulate maritime tropical deep convection during the GATE phase III field campaign using a domain...
that was ~205 km wide. Other studies have simulated cloud populations over entire countries, including the Netherlands (Schalkwijk et al., 2012, 2015) and Germany (Heinze et al., 2017). The ability to include simulations of soil properties and surface fluxes via a land model coupled with the System for Atmospheric ModelingLES model has also been demonstrated by Lee and Kariroutdinov (2015).

In this study, we use a nested approach to simulate the complex convective cloud populations observed on 30 August 2016 during the Holistic Interactions of Shallow Clouds, Aerosols, and Land-Ecosystems (HI-SCALE) campaign conducted around the ARM SGP site in north-central Oklahoma (Fast et al., 2019).

The early afternoon cloud distribution was complex on this day, with some regions having few shallow clouds, other regions having many shallow clouds, and a few regions having a small number of large clouds that transitions to deeper, precipitating convection. The LES version of the Weather Research and Forecasting (WRF) model is used to investigate the processes responsible for this complex distribution of clouds. Specifically, we hypothesize that spatial variations in land properties and subtle variations in the meteorological profiles may be responsible for the observed inhomogeneous cloud populations. As will be shown later, soil moisture availability and cooler lake temperatures drive the initial spatial variations in the cloud distributions while cold pools become a larger factor in contributing to the variability in cloud distributions after 1300 CST. Section 2 describes aspects of the HI-SCALE campaign relevant to the present modeling study, details of the LES version of the WRF model, and the configuration of the simulations. Section 3 presents the simulated cloud distributions in the context of satellite images, an evaluation of the simulated boundary layer and cloud properties using HI-SCALE and other ARM measurements, the impacts of soil moisture variations and cold pools on cloud distributions, and a sensitivity study that investigates the role of variable soil moisture. A discussion of the results is presented in section 4, and the conclusions are summarized in section 5.

2. Methods

2.1. HI-SCALE Campaign

The Department of Energy’s HI-SCALE campaign was conducted during 2016 near the ARM SGP site in north-central Oklahoma to better understand how a wide range of processes influence the evolving macroscopic and microphysical properties of shallow convective clouds. To supplement the extensive in situ and remote sensing instruments deployed at several ground sites over the SGP site, the Department of Energy Gulfstream-1 (G-1) aircraft collected in situ meteorological, radiation, cloud microphysics, and aerosol measurements within the boundary layer and the lower free troposphere. Two Intensive Observational Periods (IOPs) were conducted by the G-1 aircraft, one in the spring between 24 April and 21 May and the other in the late summer between 28 August and 24 September. The two IOPs can be compared to examine how seasonal changes in the “greenness” of cultivated crops and other vegetation types influences boundary layer and cloud properties. While this study focuses on 30 August during the second IOP, all the flight paths for the second IOP are shown in Figure 1a in the context of the ARM and Oklahoma Mesonet sites in north-central Oklahoma. Sampling was largely conducted within 100 km of the SGP site, but some transects were also performed to sample upwind (usually south) of the SGP.

In situ aerosol sampling was conducted at the SGP Central Facility (C01 in Figure 1) to provide detailed information on aerosol size, composition, and mixing state consistent with the instrumentation on the G-1. The surface measurements provide continuous sampling of aerosols to supplement the periodic sampling made by the aircraft. Additional details of the field campaign objectives, design, and instrumentation are described by Fast et al. (2019).

The HI-SCALE campaign coincided with the deployment of additional instrumentation to reconfigure SGP into a “megasite” (ARM, 2014) during the spring of 2016. To supplement the existing meteorological profiles obtained from a Doppler lidar, Raman lidar, Atmospheric Emitted Radiance Interferometer, and radiosondes at the Central Facility as well as radar wind profilers close to the Central Facility, Doppler lidars and Atmospheric Emitted Radiance Interferometers were installed at four extended facilities surrounding the Central Facility (E32, E37, E39, and E41 in Figure 1). High-time-resolution sampling of wind, temperature, and humidity profiles are now obtained over a region within ~50 km of the Central Facility. In terms of HI-SCALE objectives, another advantage of the ARM megasite is the network soil temperature, soil moisture, and surface heat, moisture, and momentum flux measurements which are needed to better
understand how land-atmosphere coupling affects the variability of boundary layer and cloud properties. The larger Oklahoma Mesonet (McPherson et al., 2007) collects surface meteorology, soil temperature, and soil moisture measurements over the entire state. In this study, we focus on the shallow convective cloud populations observed on 30 August 2016 when some clouds transitioned to cumulus congestus and deep convection. During this morning, a band of clouds and precipitation associated with a slow-moving trough extended from eastern New Mexico into western Kansas; however, clear skies were observed over most of eastern Oklahoma and southeastern Kansas. The G-1 took off at 0835 CST from the Bartlesville airport and made a series of southwest-to-northeast transects at different altitudes over the SGP site to characterize boundary layer meteorology and aerosol properties over the region (Figure 1b). Shallow cumulus formed southeast of the SGP site shortly before 1030 CST as indicated by the image from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra satellite. The aircraft diverted from its planned pattern to sample those shallow cumulus from 1030 CST until the end of the flight at 1126 CST. GOES satellite images showed that by noon, shallow cumulus was present over Oklahoma, southeastern Kansas, and most of the south-central United States (not shown). After refueling, the G-1 performed a short 1-hr flight beginning at 1232 CST (Figure 1b). The G-1 first flew just below cloud base to the Central Facility and then sampled within clouds upon its return to the Bartlesville airport. Video images from the G-1 indicate many instances of short-lived overshooting cumulus penetrating into the free troposphere. As the afternoon progressed, the MODIS Aqua image shows that some shallow cumulus transitioned to deeper convection in some regions, while pockets of clear skies formed in other regions.

2.2. LES

Version 3.8.1 of the WRF model (Skamarock et al., 2008) is used in LES mode to better understand the processes governing the convective cloud populations on 30 August. This version of the code includes a fix to the dynamic core that prevents spurious motions near strong moisture gradients between the boundary layer and free troposphere (Xiao et al., 2015; Yamaguchi & Feingold, 2012). In this study, the simulations use the Noah land surface (version 2.7; Chen & Dudhia, 2001), the RRTMG shortwave and longwave radiation (Iacono et al., 2008), and the Thompson microphysics (Thompson et al., 2008) parameterizations. Since the model run in LES mode, planetary boundary layer and cumulus convective parameterizations are not used because turbulence and clouds are explicitly simulated. The Deardorff (1980) is used to represent turbulence closure in the LES.

A large modeling domain is needed to represent the observed complex variations in the clouds observed during this case study (Figures 1b and 1c) and the processes that drive cloud formation. Therefore, we chose a 297-km-wide outer domain that encompasses much of central Oklahoma and southern Kansas with a 300-m horizontal grid spacing and a 120-km-wide inner nested domain that encompasses the ARM Central and
Extended Facilities using a 100-m horizontal grid spacing. One-way nesting is used in which the boundary conditions for the inner domain are provided by conditions in the outer domain. While a smaller horizontal grid spacing will better resolve small shallow cumulus, the domain size and grid spacing was chosen to limit computational expense. We note that a horizontal grid spacing of 100 m has been used by many other LES studies of shallow clouds, and some of them are described in section 1. A vertical coordinate that uses 304 grid levels extends up to ~16.2 km mean sea level (MSL), with a constant 24-m grid spacing from near the surface up to ~6.2 km MSL that gradually increases above that level. The land surface parameterization uses four soil layers, with default grid spacings of 0.1, 0.3, 0.6, and 1.0 m. As described by Liu and Shao (2013) and Lohou and Patton (2014), the default vertical grid spacing at the ground surface was reduced from 0.1 to 0.01 m so that the soil can better respond to near-surface temperature fluctuations associated with atmospheric eddies for LES applications.

Many LES studies, including those produced by LES ARM Symbiotic Simulation and Observation, also prescribe fixed surface heat and moisture fluxes that do not vary across the domain for real cases. This may be a reasonable assumption for small domains, but it is not a reasonable assumption for the larger domain sizes used in this study. Instead, our LES approach includes spatiotemporal variations in fluxes and consequently land-atmosphere coupling by including a land-surface parameterization.

The simulation period is from 1200 UTC 30 August to 0000 UTC (0600–1800 CST) 31 August. Instead of using periodic boundary conditions for the outer WRF domain as with many LES studies, the initial and boundary conditions for the meteorological variables are based on the National Center for Environmental Prediction Final (FNFL) operational model global tropospheric analyses (National Center for Environmental Prediction, 2000). The FNFL analyses are available at 6-hr intervals on a 1 × 1° grid. In this way, the boundary forcing varies in space and time, similar to mesoscale applications of the WRF model. A similar configure was tested by Rai et al. (2019), using WRF-LES with 40-m horizontal grid spacing driven with the North American Regional Reanalysis. They found that while there was some impact on the simulation of the largest scales, the turbulence structures were well captured. Ideally, we would have used NOAA’s High-Resolution Rapid Refresh operational model (Benjamin et al., 2016) that uses a 3-km grid spacing over the continental United States to provide the initial and boundary conditions for the LES; however, High-Resolution Rapid Refresh was offline between late August and early September of 2016 as it was being updated.

With a LES domain, periodic boundary conditions may adversely affect the evolving cloud populations. For example, a line of deep convection formed just west of the model domain during the early afternoon that propagated into the western side of the LES domain by late afternoon (not shown). Therefore, the meteorological forcing on the western and eastern sides of the model domain differ. However, a disadvantage is that turbulence will be suppressed at the inflow lateral boundaries since the FNFL analyses contain only the large-scale mean conditions. Our simulations showed that for the relatively weak synoptic forcing on 30 August, anomalous suppression of turbulence associated with ambient near-surface inflow along the eastern boundary propagated into the domain by as much as 42 km near the southeast corner of the domain by 1600 CST (2200 UTC). It does not propagate as far as the inner WRF domain which is 88.5 km from the outer domain boundary. In addition, the other three boundaries have either weak inflow or outflow so are not affected significantly by suppressed turbulence.

The initial conditions based on FNFL are compared to radiosonde observations at 0600 CST (1200 UTC) at the ARM Central Facility in Figure S1. As expected, the variability in the radiosonde observations is not represented by the coarse vertical grid spacing of the analyses. There is a qualitative agreement between the analyses and observed radiosonde, but there are some differences. For example, the vertical potential temperature gradient between the surface and 800 m MSL is somewhat higher in the analyses than in the observed profile. The specific humidity in the analyses is 1 to 2 g/kg higher than observed at some altitudes. Finally, the wind speeds are generally lower than observed. While the FNFL analysis indicates the presence of a low-level wind speed maximum of 3 m/s at 700 m MSL, the observed wind speeds at that altitude are closer to 6 m/s.

Instantaneous three-dimensional meteorological variables from the LES model are saved at 1-min intervals. When comparing the predictions with G-1 aircraft observations, variables from the grid cell coinciding with the aircraft location are used and interpolated in time to 1-s sampling time intervals. Since the flight speed of the G-1 is ~100 m/s, the aircraft crosses each 100-m grid cell in the inner WRF domain in about 1 s.
2.2.1. Default Simulation

There are three primary simulations performed for this study to investigate the processes leading to the transition of shallow cumulus to precipitating cumulus congestus and deep convection. The first, called the "default simulation," uses the FNL analyses for the meteorological initial and lateral boundary conditions as described previously. In addition, the initial soil temperature and soil moisture content in WRF is interpolated from the FNL analysis at 12 UTC. The soil moisture content distribution in the upper soil layer is shown in Figure 2a, along with the uppermost in situ measurements (5 cm below ground) from the ARM and Oklahoma Mesonet sites. Estimates from the Global Land Evaporation Amsterdam Model (GLEAM version 3.1; Martens et al., 2017) available on a 0.25 × 0.25° grid, which estimates evaporation and root-zone soil moisture from satellite data and a set of algorithms, are shown in Kansas where there are no in situ measurements. We use GLEAM estimates only in Kansas because comparisons between in situ soil moisture measurements in Oklahoma reveal a wet bias in the GLEAM estimates when in situ measurements become less than 0.3 m³/m³ as shown in Figure S2. We also compared the observations with Soil Moisture Active Passive (SMAP) satellite retrievals (Kumar et al., 2018) and found a bias closer to zero, but differences between SMAP and in situ measurements were often as large as 0.1 m³/m³. The overall spatial pattern in soil moisture over Kansas from SMAP was similar to GLEAM. It is important to note that mixing soil moisture information from more than one source (Koster et al., 2009) introduces uncertainties in the simulations, but as will be shown later does not affect the overall conclusions of this study.

The FNL analysis has a broad variation in soil moisture content over the WRF domain, with highest values between 0.30 and 0.35 m³/m³ in Kansas and lower values between 0.25 and 0.25 m³/m³ in the south-central and east-central sides of the domain. The FNL analysis captures only the broadest north to south variations in observed soil moisture content. There is much more spatial variability and range in the observed soil moisture content, with values reported to be as high as 0.437 and as low as 0.029 m³/m³. Similarly, there are differences between the FNL analysis and the observations in the spatial variability of soil temperature (not shown). The default simulation also uses land use type based on the WRF MODIS 30 s (~1 km) data set. As shown in Figure S3a, most of outer and inner domains are defined by crop or grassland, with a small fraction of the total area defined by water, urban, or deciduous forest.

2.2.2. Revised Simulations

As noted previously, there are potential weaknesses in the configuration of the default simulation associated with two hypotheses:

1. Our first hypothesis is that more realistic variations in soil properties (Figure 2a) and land use (Figure S3a) could significantly influence land-atmosphere coupling and consequently the life cycle of convective clouds.
2. Our second hypothesis is that the small errors in the initial meteorological fields (Figure S1) may contribute to ability of shallow convection to transition into deeper, precipitating convection that occurred in some areas over north-central Oklahoma on 30 August.

To address the hypotheses, two additional simulations are performed. The first revised simulation, that is, “revised-1,” addresses both hypotheses while the second revised simulation, that is, “revised-2,” addresses only the second hypothesis. By comparing the revised-2 and “default” simulations, differences in the cloud fields will be influenced only by changes to the initial atmospheric state. Differences in cloud fields influenced only by changes in soil properties can be obtained by comparing the revised-1 and revised-2 simulations.

To address the first hypothesis, we account for higher spatial variability in the soil moisture content and soil temperature by replacing the initial conditions from the FNL analyses with values obtained by interpolating the observations using a kriging algorithm with Gaussian covariance weighting. Lake and river temperatures are also replaced with the mean of the G-1 infrared thermometer measurements over the lake and river surfaces (301 K). The initial soil moisture content for the revised-1 simulation is shown in Figure 2b, and as expected the spatial variability and range is in much better agreement with the observations. In addition, we replaced the land use uses values derived from a 30-m National Land Cover Database for 2011 (Yang et al., 2011). The land use type in each WRF grid cell is obtained by taking the dominant land use type among the 30-m National Land Cover Database cells within each 300- or 100-m WRF grid cells. The resulting land use type for the outer WRF domain is shown in Figure 1a, while Figures S3a and S3b compare the land use type between the default and revised-1 simulations for both domains, respectively. The land use has much more heterogeneity in the revised-1 simulation, with a larger fraction of grid cells using deciduous forest, shrubland, and wetland land use types (Figure S3c). Water surfaces are also more realistic (Figure S3b) and now include rivers and small lakes that were not present in the MODIS 30-s data set, particularly for the inner WRF domain.

Zhang and Klein (2010, 2013) used summertime ARM observations over more than a decade to identify the relative importance of several environmental conditions, showing that vertical gradients in temperature, moisture, and winds were associated with the depth of shallow cumulus. Therefore, we nudged the initial potential temperature, specific humidity, and winds to the observed radiosonde profile (Figure S1) to address the second hypothesis. Differences between the radiosonde profile and the default simulation are computed first, and then those differences are applied to the three-dimensional fields by using a horizontal Gaussian function with weights that decrease to zero toward the outer domain boundaries. Therefore, the initial profile in the revised-1 and revised-2 simulations at the Central Facility near the middle of the domain is nearly identical to the radiosonde and additional vertical variability in potential temperature, specific humidity, and winds is introduced into the initial conditions for other parts of the modeling domain. The convective available potential energy computed from both profiles are below 1,000 J/kg (marginally unstable); however, the profile from the FNL analysis at the SGP site is slightly more unstable than the radiosonde. Convective inhibition computed from the profiles are both below 50 J/kg.

3. Results
3.1. Cloud Distributions
For both the default and revised simulations, shallow convective clouds formed initially over the southeastern corner of the outer domain around 0940 CST. Cloud formation gradually expanded over much of the southern and east-central portion of the outer domain during the next 40 min. The MODIS Terra Image (Figure S4a) shows that shallow clouds were present south and east of the SGP site by 1030 CST. While shallow clouds from both simulations initially formed in the right locations, the cloud fraction in those locations was lower than observed at 1030 CST (Figures S4b and S4c), suggesting that the more widespread cloud formation occurred later than observed. Over the inner WRF domain (Figure S5), clouds in both simulations formed between the E37 and E39 sites, consistent with the satellite image at that time. During the following 3 hr, simulated shallow clouds formed over the rest of the outer and inner modeling domains consistent with GOES images (not shown); however, the cloud distributions between the default and revised-1 simulations began to diverge.
By 1350 CST, the observed cloud populations over north-central Oklahoma and southern Kansas were quite complex as indicated by the MODIS Aqua image in Figure 3a. Nearly uniform shallow cloud distributions over Oklahoma earlier in the morning gave way to cellular cloud-free regions with shallow cumulus in between and some clouds transitioning to deep convection. In contrast, the cloud distribution from the default simulation (Figure 3b) is more uniform across the outer domain with somewhat larger clouds along the western boundary and over northeastern corner of the domain. The cloud distribution in the revised-1 simulation (Figure 3c) is far less uniform than the default simulation, including cellular cloud-free regions along with patches of deeper and larger convective clouds. Even though the cloud distributions are different, the cloud fraction of 27.5 and 29.2% is similar for the default and revised-1 simulations, respectively. As will be discussed in more detail later, cold pools were responsible for some of the cloud variability in the revised-1 simulation. For example, the ring of clouds in the southeast corner of the domain at 1350 CST (Figure 3c) is the result of multiple cold pools from precipitating convection around noon. Clouds were suppressed within the cold pools as they expanded and merged, but convergence along the cold pool boundaries favored the formation of additional convection. Clouds did not transition to deeper convection in the default simulation, except along the western boundary.

Figure 4 shows that the inner domain of the revised-1 simulation produces a cloud distribution over the SGP site that is in better agreement with the satellite image than the default simulation. As with the outer domain, the default simulation produced a more uniform cloud distribution than the revised-1 simulation. The revised-1 simulation has a cluster of larger clouds between the E37, E39, and C01 sites as well as larger
clouds along the border of Oklahoma and Kansas, similar to observed. In addition, the regions with fewer and smaller shallow clouds in the revised-1 simulation also agree with the satellite image. While the two simulations have different cloud distributions over the inner domain, the overall cloud fraction of 20.0 and 20.5% is very similar for the default and revised-1 simulations, respectively.

The cloud size, depth, and spatial distribution from the revised-2 simulation are very similar to default simulation. This indicates that changes to the initial vertical profile in temperature, humidity, and winds did not significantly impact the spatial variability in cloud populations or the transitions of shallow cumulus to deeper convection over the inner WRF domain. Since the differences in the simulated cloud distribution shown in Figures 3 and 4 are largely driven by differences in the treatment of surface and soil properties, the analyses from now on will focus on comparing the default and revised-1 simulations.

Another way to examine the differences between the simulations is to use a cloud tracking algorithm (Feng et al., 2018) called the FLEXible object TRaKeR, FLEXTRKR, to determine the position, size, and lifetime of individual clouds during the simulation period within the inner WRF domain. While FLEXTRKR was initially designed to track deep convective clouds, the underlying technique which uses spatial overlap between different times to track any feature is applicable for tracking shallow clouds as well, particularly with the 1-min output frequency of the simulations in this study. Several assumptions are used to define cloud objects for tracking. First, the liquid water path (LWP) field is smoothed with a 10 × 10 window to define a “cloud core” with connected grid cells having LWP > 0.75 kg/m². Then the shallow cloud objects are defined by expanding the region surrounding the cloud core to LWP > 0.2 kg/m². For the remaining clouds that do not have a core, we define a cloud object with LWP > 0.2 kg/m². The minimum area of a cloud object for tracking is 900 m² (nine grid cells). All defined cloud objects are tracked, and those lasting shorter than 4 min are neglected. The algorithm also considers merging and splitting of clouds. The number of cloud tracks meeting these specifications for the default simulation are 5,418, 13,247, and 14,703 between 10–12, 12–14, and 14–18 CST, respectively. There are about 24% fewer clouds tracked in the revised-1 simulation with 4,590, 10,376, and 10,273 between 10–12, 12–14, and 14–18 CST, respectively.

Figure 5a shows the joint frequency of occurrence of cloud lifetime versus maximum equivalent cloud diameter along the cloud track for the default simulation. As expected, the statistics from the cloud tracking algorithm show that small clouds tend to have a short lifetime and larger clouds have a longer lifetime. For example, cloud tracks with a maximum diameter of 500 m have lifetimes up to 11 min but usually less than 8 min. For a maximum equivalent cloud diameter of 1,500 m, most of the clouds have lifetimes between 11 and 22 min. Figure 5b depicts the difference between the revised-1 and default simulations. While the two simulations have similar statistics on the relationship between cloud size and lifetime, the revised-1 simulation has an upward shift in the distribution such that for a given cloud lifetime, the clouds tend to grow larger than the default simulation suggesting the convection in the revised-1 simulation is more vigorous than the default simulation. When these statistics are divided into cloud tracks during the morning, early...
afternoon, and late afternoon periods, the differences between the default and revised-1 simulations are similar to those in Figure 5b. Figure S6 shows that the revised-1 simulation has larger clouds and longer lifetime during much of the day when the results are expressed in terms of percentiles. For the maximum equivalent cloud diameter (Figure S6a), there are large increases in the mean values and more modest increases in the median values, indicating that a relatively few outliers are responsible for the shift to larger clouds in the revised-1 simulation. These outliers are the relatively few shallow clouds that transition to deeper convection during the afternoon. Similarly, the mean cloud lifetime is longer in the revised-1 simulation throughout the day (Figure S6b); however, the median values of the cloud populations from both simulations are nearly the same.

3.2. Cloud and Boundary Layer Properties

To determine the realism of the simulations and quantify how changes in the revised-1 simulation affect cloud and boundary layer properties, we now include some of the HI-SCALE and routine ARM measurements. The simulated total cloud droplet number, cloud droplet number distribution, and liquid water content (LWC) along with the in situ G-1 measurements are shown in Figure 6 in terms of percentiles. The observed values are obtained from the Fast Cloud Droplet Probe (FCDP) and the Multi-Element Water Content System (WCM). Since one cannot expect the observed and simulated clouds to match in space and time along the G-1 flight path, the simulated values are obtained by sampling all simulated cloudy cells within 1 km of the G-1 flight path. The Thompson microphysics parameterization uses a prescribed cloud droplet number concentration of 100 cm$^{-3}$, which is somewhat larger than the observed median of 80 cm$^{-3}$ for the morning G-1 flight (Figure 6a), but significantly larger than 36 cm$^{-3}$ observed for the afternoon G-1 flight (Figure 6b). The observations also exhibit a large range in the total droplet number concentration that is not accounted for in the model.

We then use the assumed cloud droplet number concentration of 100 cm$^{-3}$ and the simulated cloud water mixing ratio to calculate the cloud droplet size distribution following Thompson et al. (2008), see Appendix A). The results in the middle panels of Figure 6 show that the simulated size distribution peak is shifted to larger sizes than observed. The simulated peak diameter is around 15 μm; however, the observed peak diameter is closer to 10 μm. While the simulated median of the droplet concentration is relatively close
to the observed median between 13 and 20 μm, the model underestimates the number of small droplets and overestimates the number of large droplets by one or more orders of magnitude.

As shown in the right panels of Figure 6, the simulated LWC is larger than both the FCDP and WCM observations even when values greater than 24 and 27 μm for the morning and afternoon flights, respectively, are omitted where the observed FCDP concentrations are zero. While the percentiles of LWC from the FCDP and WCM instruments are somewhat different, they are closer to one another than with the simulated values. The mean values between the two instruments are also closer to each other than the median values. While the mean LWC of 0.062 g/m³ from the FCDP is somewhat higher than 0.042 g/m³ from the WCM for the morning flight, the mean from both instruments is 0.123 g/m³ for the afternoon flight (not shown). Differences between the mean and median reflect the skewness of the observed LWC distribution.

The ARM site has measurements several instruments that can be used to evaluate the simulated cloud properties. For example, we have also compared the simulated column LWP and precipitable water (PW) with microwave radiometer retrievals at the E32, C01, and E41 sites as shown in Figure S7. While one cannot expect the timing of shallow clouds that pass directly over these sites to be identical, it is desirable that the temporal variability and magnitude of LWP and PW to be similar. Only several observed or simulated clouds pass directly overhead even on a partly cloudy day (Figure S7a). The simulated LWP is smaller than many of the observed clouds at the E32 site. This illustrates the advantage of the G-1 aircraft that can sample far more clouds. The magnitude of the PW (Figure S7b) is similar to observed during the morning, but larger differences are produced during the afternoon that are likely due to errors in moisture transport from the model boundaries. For example, the simulated overall trend in PW at E32 is similar to observed, and the magnitude is too low after 09 CST. This low bias is likely due to the large-scale analyses not representing moisture associated with the trough over western Kansas, western Oklahoma, and the panhandle of Texas that slowly propagates eastward toward the E32 site during the day.

Not surprisingly, the predicted microphysical parameters in Figure 6 are very similar since both simulations use the same microphysics parameterization. As will be discussed later, there are differences in the macrophysical properties of clouds. The simulated errors in droplet number concentration, distribution, and LWC could contribute to errors in the model’s representation of cloud radiative effects. We note that the present simulations omit the effect of aerosols which could shift the simulated cloud droplet distribution to smaller sizes because of the Twomey effect (Twomey, 1977). As shown in Fast et al. (2019) and Figure S8, there are strong gradients in particle concentrations along the G-1 flight legs on this day. Some clouds are embedded in high aerosol concentrations likely originating from refinery and power plant emissions near the SGP site, while other clouds are embedded in lower background aerosol concentrations. The spatial variability in CCN is similar to the variations in particle number concentrations, with peak CCN concentrations within the refinery or power plant plumes (not shown).
The differences in the cloud population size and spatial distribution between the default and revised-1 simulations suggest that clouds on 30 August are strongly influenced by land-atmosphere coupling. Therefore, it is important to assess how well the model represents boundary layer processes and near-surface forcing. Figure 7 shows a comparison of the simulated potential temperature, specific humidity, wind speed, and wind direction from the inner WRF domain with the G-1 measurements. The simulated values are obtained by temporally interpolating the 1-min output to the 1-s measurement times using the 100-m grid cell where the G-1 falls within. The results are then depicted as the mean and range within 100-m altitude bins for measurements west of −96.8 W near the SGP site. Table S1 quantifies the overall difference between the simulated and observed quantities for the entire profile. In general, the profiles and spatiotemporal variability between the observed and simulated values are qualitatively similar. The simulated mean potential temperatures and specific humidity are within 1 K and 1 g/kg, respectively, of the observed means in the boundary layer. The revised-1 simulation is somewhat closer to observed in the lower free troposphere since that simulation constrains the initial conditions by the 12 UTC radiosonde from the SGP site. Wind speeds, however, are less than observed at most altitudes during the morning G-1 flight and the spatiotemporal variability is also smaller than observed. For example, the mean observed wind speed is ~2.5 m/s at 1 km MSL while the simulated mean wind speed from both simulations is closer to 1 m/s. In addition, the model does not produce as strong a wind speed gradient just above cloud level. By the afternoon, the simulated wind speeds and directions are much closer to the observations.

Updrafts produced by the convective eddies support the development of shallow convection. The G-1 also provided measurements of vertical velocity, \( w \), at 1-s intervals that are compared to the simulated \( w \) in Figure 8 in terms of frequency using bins of 0.25 m/s. The simulated \( w \) is obtained along the flight path by interpolating in time between the 1-min output to the 1-s time intervals, but only for altitudes within and below cloud base. While both simulations produce a frequency distribution that is similar to the morning G-1 data, the default simulation distribution is somewhat narrower and closer to observed than the revised-1 simulation, with ~5% more occurrences within 0.25 m/s of zero. The frequency distribution in \( w \) broadens in both the observed and simulated values during the afternoon as convective eddies becomes stronger. For the afternoon G-1 flight, both simulated distributions are close to the G-1 distribution.

While the G-1 collected \( w \) data west of −98.6 W over ~2.2 hr and at selected altitudes, the five Doppler lidars obtained continuous profiles of \( w \) during the day. Figure 9 compares the simulated distribution in \( w \) to the Doppler lidar measurements for the entire daytime period at 40% of the convective boundary layer height, \( z/z_c = 0.4 \). The revised-1 simulation is closer to the observed \( w \) distribution at two (C01, E41) of the five Doppler lidar sites. Both simulations are close to the observed distribution at the E32 and E39 sites. But at the E37 site, both simulations underpredict the frequency of sinking motions between 0 and ~1 m/s and overpredict the frequency of rising motions between 1 and 3 m/s. Figure 9f averages all the frequency distributions among all the sites, showing that the revised-1 simulation is closer to the overall observed distribution. The largest differences between the default and revised-1 simulations are found in the middle of the boundary layer, as shown by statistics at five altitudes in Figure S9. Some of the differences between the simulations are hidden by these statistics computed over the entire day. When the distribution in \( w \) is examined over smaller time intervals, larger differences between the simulations are found at more lidar sites during the morning between 9 and 11 CST as shown in Figure S10. Figure S11 shows that the simulated distributions in \( w \) become more similar between 11 and 13 CST. These results suggest that the model has more difficulty simulating boundary layer turbulence development during early morning hours when land surface properties have more control compared to the late morning.
The convective boundary layer (CBL) height, $z_i$, is determined from the Doppler lidars using a threshold value of vertical velocity variance ($\sigma_w^2 = 0.1$) as described by Berg et al. (2017) and Tucker et al. (2009). The vertical velocity variance and $z_i$ is computed at 10-min intervals using a 30-min running average period. In the model, $z_i$ is determined in the same way using instantaneous vertical velocities saved at 1-min intervals. At times, the model predicts relatively high $\sigma_w^2$ above cloud base so that the threshold value of 0.1 produces an estimate of $z_i$ that is erroneous. For those cases, the height at which the vertical gradient in potential temperature exceeds 0.005 K/m is used instead to better define $z_i$. A qualitative evaluation of the simulated mean temporal variability in $z_i$ as well as its spatial variability is shown in Figure 10a. Prior to 1300 CST, the mean $z_i$ among the five Doppler lidar sites from both simulations are qualitatively similar to the observed mean, although the default simulation is more frequently closer to the observed mean. After 1300 CST, the observed mean $z_i$ levels off while the boundary layer continues to slowly grow in both simulations. Between 1300 and 1600 CST, the simulated mean $z_i$ is as much as 300 m higher than observed. In addition, the range of simulated $z_i$ among the five sites from both simulations is often as large as the observed spatial variability during much of the day. Consistent with the bias in $z_i$, both simulations produce cloud base heights that are 100–500 m higher than observed, as shown in Figure 10b for the C01 site.

The reason that the simulated CBL heights are too high is likely related to the simulated surface sensible heat and latent fluxes as shown in Figure S12. The mean surface sensible heat flux is higher than observed after 08 CST, while the simulated latent heat fluxes are somewhat lower than observed during much of the day. It is important to note that some of

Figure 9. Observed and simulated (inner WRF domain) frequency of vertical velocity at $z/z_i = 0.4$ at the (a) E32, (b) E41, (c) C01, (d) E37, and (e) E41 Doppler lidar sites.

Figure 10. (a) Observed and simulated (inner WRF domain) boundary layer height among the five Doppler Lidar sites and (b) boundary layer and cloud base heights at the Central Facility.
the ARM are sites obtain surface fluxes using the eddy flux correlation system (ECOR) while other sites use the energy balance Bowen ratio system (EBBR). The EBBR system uses the Bowen ratio to close the surface energy balance and the ECOR system uses an eddy covariance approach that has no such conservation restraints that may require corrections (Cheng et al., 2017). These methodologies introduce some uncertainties that affect model-to-observation comparisons, such as Figure S12 that uses both ECOR and EBBR data. In addition, the ECOR and EBBR instruments are located in either crop or grassland sites, respectively, which can contribute to differences in the fluxes (Tang et al., 2019). We also segregated the model evaluation into ECOR and EBBR sites and found that the overall bias and range of simulated values was similar to the results shown in Figure S12.

The bias in sensible heat flux leads to near-surface temperatures that are often positively biased over both the inner and outer WRF domains (Figures S13 and S14). In fact, the largest temperature bias over the inner WRF domain of 1.6 °C occurred at the Central Facility. Simulated latent heat fluxes around that site were between 250 and 350 W/m² during the midafternoon, but observed latent heat fluxes were usually 50 to 100 W/m² higher (not shown) This suggests that either the model's soil moisture content around the Central Facility is too low, or that there are other errors in how the model represents the surface energy balance.

Profiles of $\sigma_w^2$ (not shown) showed that turbulence aloft dissipated during the late afternoon at the C01 and E41 sites after 1630 CST, which is indicative of a collapse in boundary layer intensity and height. In contrast, $\sigma_z$ at the other three sites remained relatively constant between 1630 and 1800 CST. For the E41 site, the decrease in $\sigma_z$ at 1630 CST also corresponds to strong near-surface easterly flow forming over the site (not shown) which may be outflow from a cold pool. While the revised-1 simulation has a dissipating cold pool over the northeastern corner of the inner WRF domain at this time, a stronger cold pool originating from convection in the outer domain did not propagate to the E41 site until close to 1800 CST. While the revised-1 simulation does produce a decrease in $\sigma_z$ in the late afternoon that is closer to observations than the default simulation, the model errors at this time are probably due to cold pools not forming at the right times or not with the same intensity and/or the convective eddies not dissipating fast enough during the late afternoon at some sites.

3.3. Impacts of Soil Moisture Variations and Cold Pools

Since the revised-1 simulation produces cloud distributions that are in reasonable agreement with satellite observations, we now use the model predictions to identify the processes contributing to the spatial...
variability in the cloud populations. Figure 11 shows horizontal and vertical cross sections of clouds and potential temperature variations between 1130 and 1330 CST. Figure 11a shows that most of the clouds form over regions with the driest soil moisture during the morning. In those regions there are large numbers of small clouds and relatively fewer larger clouds. Far fewer clouds formed over the wetter soil regions. The west-to-east vertical cross section shows that the CBL depth over the drier soil moisture regions is also higher than over the wetter soil regions. The cloud depths over the dry soil moisture regions are also higher. The higher latent heat flux over the wet soil regions leads to cooler near-surface temperatures and weaker turbulent eddies (not shown) than over the drier regions. Thus, boundary layer growth and cloud formation over the wetter soil regions lags behind the drier soil regions.

These results are consistent with Lee et al. (2019), in which they used an idealized LES configuration with square patches that have either higher or lower prescribed latent heat fluxes to mimic wet and dry soil conditions, respectively. In that study, clouds preferentially formed over the dry regions when the ambient winds were less than 2 m/s. Interestingly, ambient winds near the top of the CBL over the inner domain of the revised-1 simulation that ranged from 2 to 5 m/s (3.6 m/s average) did not suppress the regional variability in cloud formation. The relatively large wet/dry soil moisture patterns in this study (on the order of tens of kilometers) may be large enough to prevent mixing by the ambient horizontal wind.

One hour later at 1230 CST (Figure 11b), the overall cloud field imposed by the soil moisture variations is still present; however, the easterly to southeasterly ambient winds advect clouds formed over the dry soil regions toward the wet soil regions. Animation of the cloud fields reveals that these clouds dissipate as the convective eddies become weaker over the wet soil region. By 1330 CST (Figure 11c), a larger fraction of the clouds advected over the wet soil regions do not dissipate as convective eddies become stronger and the CBL height differences between the dry and wet soil region diminish.

In addition to the impact of soil moisture variations, the lakes also have a noticeable effect on cloud formation. Similar to the wet soil regions, the cooler lake temperatures suppress turbulent convective eddies so that updrafts are not strong enough to initiate cloud formation. At 1130 CST, the impact of the lakes is largely confined to the lakes and the surrounding land downwind (west) of the lakes. As surface temperatures surrounding the lakes increase, clouds are still suppressed as the cooler temperatures and weaker turbulent eddies are transported from the lake over the land. As the day progresses, the regions of suppressed turbulence and clouds expand westward at 1230 (Figure 11b) and 1330 CST (Figure 11c). Some smaller clouds eventually form in these regions as additional surface heating compensates for the suppressed turbulence transported from the lakes. By 1330 CST, the area impacted by the lakes is actually an order of magnitude or larger than the surface area of the lake itself. This effect also occurs over more lakes located on the outer domain (not shown). These results suggest that the impact of man-made reservoirs on cloud formation contributes to the distribution of cloud populations.

The results shown in Figure 11 are consistent with Zhang and Klein (2010, 2013) in their analysis of long-term measurements made at the SGP. They found that greater inhomogeneity in the morning boundary layer among the ARM and Oklahoma Mesonet sites were more likely to be associated with days in which shallow cumulus transitioned to deep convection. They also showed that humidity between 2 and 4 km was the environmental factor with the highest statistical significance associated with the transition from shallow cumulus to deep convection. However, we find that the shallow-cumulus revised-1 simulation transitioned to deep convection even though the free tropospheric humidity was less than the default simulation (Figure S1). By comparing the default and revised-2 simulation, we also found that changing the initial temperature and moisture profiles to be closer to the observed radiosonde observations had little effect on the simulated shallow cumulus. The shallow cumulus in the revised-2 simulation did not transition to deep convection more frequently than default simulation. For the 30 August case, the driving factor in the morning and early afternoon appears to be the variability in the boundary layer properties driven by land-atmosphere coupling processes.

While soil moisture and lakes contribute to the spatial variability of surface temperature and shallow cloud distributions during the morning and around noon, the model results show that cool pools also play a role after 1300 CST as some of the shallow clouds transition to deeper precipitating clouds. To quantify the impact of cold pools, we computed surface virtual potential temperature anomalies less than −1 °C associated with cold pools at each 1-min interval. Then, those anomalies at 1-min intervals were summed up between 1300
and 1500 CST to produce a “footprint” of cold pools over the inner model domain as shown in Figure 12a. We note that cloud shading also produces localized and transient cooling at the surface; however, the magnitude of the temperature perturbations is lower than the cutoff used for the virtual potential temperature anomalies and therefore does not contribute to the pattern shown in Figure 12a. In the south-central portion of the inner WRF domain, many small cold pools formed that usually lasted only for several minutes. Surface potential temperatures decreased a few degrees, but quickly increased to back to ambient temperatures as the stronger downdrafts dissipated. Interestingly, cold pools often formed in the same regions. The footprint in the south-central portion of the domain is made up of many short-lived cold pools during the 2‐hr period, rather than one continuous cold pool event. Animation of the cloud fields suggests that these short-lived cold pools affect cloud formation nearby, but not over distances greater than a few kilometers. We note that there were many smaller cold pools with potential temperature anomalies between 0 and −1 °C not included in Figure 12a. The cold pool in the northeast corner of the domain results from a larger and more persistent cold pool that will be described in more detail later.

To demonstrate the impact of cold pools on cloud macroscopic properties, we computed cloud base, top, and depth over the south-central portion of the inner model domain denoted by the dashed box in Figure 12a. The area in the box was then divided further into regions with soils moisture >0.25 m³/m³ (west side of dashed box) and <0.25 m³/m³ (east side of dashed box) in the revised-1 simulation. In the default simulation, there are much smaller gradients in soil moisture in this region (Figure 2a), but the same areas are used in the analysis. The cloud base, top, and depth are expressed in terms of percentiles in Figures 12b and 12c for the wetter and drier soil regions, respectively. For the west side of the box, the cloud base from the two simulations shown in Figure 12b are very similar for most of the day. After 1400 CST, the cloud tops and thus the cloud depth are somewhat higher in the revised-1 simulation. Cloud fraction in the revised-1 simulation was smaller than the default simulation during the entire day for the wetter soil region as seen in the top panel. Much larger differences in cloud base, top, and depth between the two simulations are produced over the east side of the dashed box as seen in Figure 12c. The revised-1 simulation has somewhat higher cloud bases than the default simulation as a result of the drier soil moisture conditions that lead to warmer near-surface temperatures. The cloud tops and therefore cloud depths are also higher throughout the day so that the differences between the two simulations are much larger than the wetter soil moisture in the west side of the dashed box. In general, the results show that the revised-1 simulation produced deeper convective clouds over the dry soil regions, and some of those clouds produced precipitation and short-lived cold pools. The deeper convective clouds are then advected over the wet soil regions where they eventually dissipate. While the wet soil region also has deeper clouds, the lower 75th percentiles indicate that are fewer deep clouds than in the dry soil region.

Figure 12. (a) Footprint of all cold pools (color shading) between 13 and 15 CST over the inner WRF domain with temperature anomalies <−1 °C superimposed on soil moisture distribution (black contours). Percentiles of cloud base, top, and depth within the dashed box in (a) for soil moisture, (b) greater than 0.25, and (c) less than 0.25 m³/m³.
While short-lived cool pools formed preferentially in one region of the inner model domain, larger more persistent cold pools formed in the northwest corner of the domain and in other regions of the outer model domain (not shown). Figure 13 shows the evolution of the persistent cold pool at 3 times, similar to Figure 11, but between 1415 and 1545 CST in the northeast corner of the domain, and we now include vectors to illustrate near-surface cold pool outflow. At 1415 CST, the cold pool as defined by the stronger winds and cooler temperature is ~12 km wide. The vertical cross section shows strong upward motions supporting the convection along the leading (western) edge of the clouds and strong downdrafts along the eastern edge of the clouds where precipitation reaches the surface. At this time, near-surface potential temperatures are 2 to 2.5 K cooler than the surrounding air. Long after convection dissipates, the cold pool continues to expand as seen by the outflow wind pattern and suppressed turbulence. By 1500 CST the cold pool is ~25 km wide (Figure 13b). While near-surface potential temperatures have rebounded, they are still cooler than the area west of the cold pool. The vertical cross section shows that the boundary layer is not yet well mixed in terms of potential temperature and turbulence is still suppressed (not shown) as evident from the lack of clouds forming at the boundary layer top. In addition, another cold pool formed to the east. By 1545 CST (Figure 13c), the two cold pools have merged and the front becomes less distinct as near-surface potential temperatures and boundary layer turbulence continues to increase and become more like the ambient regional conditions. In addition, a strong cold pool forms on the outer domain that begins to propagate into the northeast corner of the inner model domain at 1500 CST. The cold pool front is more pronounced at 1545 CST, bringing in cooler air and stronger winds that become superimposed on the dissipating cold pool.

Therefore, the cloud distribution after 1300 CST becomes increasingly influenced by cold pool outflow. Some cold pools are small and transient and only influence the local cloud field, while other clouds transition to deeper convection with higher precipitation rates that lead to cold pools that persist longer and expand farther. Therefore, the distribution of clouds shown in Figures 1c and 4a result from two primary mechanisms. First, spatial variability in heat and moisture fluxes resulting from soil moisture and lake temperatures control the initial distribution of clouds up to about 1300 CST as the ambient winds advect shallow clouds from east to west. Second, cold pools resulting from some clouds that transition to precipitating convection after 1300 CST increasingly affect the cloud distributions over a larger fraction of the model domain. While the outflow suppresses cloud formation within the cold pools, convergence along the cold pool boundaries enhances convection at the cold pool boundaries.

3.4. Sensitivity to Soil Moisture Variations

Previous idealized and real-world modeling studies have demonstrated the importance of soil moisture variations in creating secondary circulations that influence cloud development (e.g., Chen & Avissar, 1994; Lee...
et al., 2019; Segal & Arritt, 1992; Xiao et al., 2018). For idealized soil moisture variations and no synoptic forcing during the day, warmer (cooler) air over dry (wet) soil regions will set up convergence (divergence) zones near the surface and rising (sinking) motions in the boundary layer. Aloft, the wind patterns are reversed. Clouds will therefore preferentially form over the dry soil moisture regions demonstrated by Lee et al. (2019) and other studies. However, the secondary circulations are more complex and difficult to identify in real-world conditions. Consistent with theory and idealized simulations, both the default and the revised-1 simulations have weaker near-surface wind speeds over the wetter soil regions at 12 CST as shown in Figure S15. The revised-1 simulation produces higher wind speeds of 4 to 6 m/s along the soil moisture gradient between the E39 and Central Facility sites near (~1,700 m MSL) and above (~2,185 m MSL) the boundary layer. This acceleration is consistent with secondary circulations imposed on the larger-scale easterly winds. In addition, the simulated mean vertical motions in the boundary layer within 20 × 20-km boxes upwind and downwind of wet soil region (see Figure S15) were positive and negative, respectively. Figure S16 shows the frequency of vertical velocity at ~1,700 m MSL in which the revised-1 simulation had mean vertical motions of 3.6 and −2.4 cm/s in the upwind and downwind boxes, respectively. These mean rising and sinking motions are consistent with the concept of secondary circulations induced by variable soil moisture. In contrast, the default simulation had less of a difference in vertical motions in those two boxes since the soil moisture gradients were weak.

Both the default and revised-1 simulations produce wind speed and direction profiles similar to the radiosonde observations at 12 CST (Figure S15); however, the mean wind speed profile within 2 km of the Central Facility is somewhat closer to the observed profile. The revised-1 simulation has lower wind speeds in the boundary layer and higher wind speeds near and above the top of the boundary layer, suggesting that the more realistic soil moisture distributions contributed to improvements in the simulated winds. It is important to note that the LES model produces substantial variability in wind speed and direction near the Central Facility, as evident in the percentiles and peak values within 2 km of the site. While the model exhibits variability in the winds, the radiosonde encounters only a limited number of eddies as it rises through the boundary layer. Therefore, the single realization of the radiosonde must be put into the proper context of the simulated wind profile variability. In addition, the ARM wind profile network (E32, E37, E39, E41, and Central Facility sites) may capture some aspects of secondary circulations, but that depends on the local soil moisture gradients near these sites.

To further examine the role of soil moisture variability on regional cloud distribution for the atmospheric conditions observed on 30 August, we performed four additional sensitivity simulations for only the outer model domain. The initial specified soil moisture distributions used by these simulations are depicted in

![Figure 14](image-url)
the top panels of Figure 14. In one simulation, a constant soil moisture of 0.23 m³/m³ is used for the entire domain. Simulations “variable-1” and “variable-2” have the same soil moisture distribution as in the revised-1 simulation, except that a smoother has been applied to moderately (variable-2) or strongly (variable-1) reduce the spatial gradients in soil moisture. In contrast, the “variable-4” simulation increases the soil moisture gradients used by the revised-1 simulation (also called “variable-3”) by making the wet soil regions wetter and the dry soil regions drier.

The resulting cloud fields at 1350 CST for the inner portion of the modeling domain indicated by the dashed box are shown by the bottom row of panels in Figure 14. The simulation with constant soil moisture produces a shallow cloud field that is nearly uniform over most of the domain. The variable-1 simulation with a relatively smooth soil moisture distribution produces a cloud field that is very similar to the constant soil moisture simulation. This is consistent with what was produced in the default simulation. When slightly larger soil moisture gradients are imposed, as in variable-2, a somewhat more complex cloud field is produced with noticeably more larger clouds in the center of the domain. As shown previously, the variable-3 simulation which is the same as the revised-1 simulation now produces regions with larger clouds and regions that are nearly cloud-free near the center of the domain. Finally, when the soil moisture gradients are exaggerated in variable-4, the cloud field nearly becomes bimodal with regions that are cloud free and other regions containing large clusters of deep convection. These results suggest that for the ambient meteorological conditions on 30 August, small gradients in soil moisture over the domain do not significantly impact on the variability in shallow clouds. There are also few instances of shallow clouds transitioning to deeper convection at this time. As soil moisture gradients reach some critical value, the variations in surface heat and moisture fluxes begin to affect boundary layer growth in a way that also influences and cloud formation. As has been shown previously, using an initial soil moisture distribution based on the observations leads to conditions that permit the transition of some of the shallow clouds to deeper, precipitating convection that is consistent with satellite images. When the soil moisture distribution extremes are exaggerated, too much deep convection is produced at this time.

We now quantify the variations in the cloud fields from the simulations shown in Figure 14 in terms of temporally varying cloud fraction. Cloud fraction is determined within 22 × 22 subregions that encompass the outer model domain, using a cutoff of LWP >0.01 kg/m² to define cloudy grid cells. Each subregion has an area of 182.25 km² and consists of 45 × 45 model grid cells that are 300 m wide. We then express cloud fraction in terms of percentiles that are computed over an hour-long period for 81 of those subregions that fall within the inner model domain to examine the cloud fraction variability near the SGP site. As shown in Figure 15, all of the simulations produce peak cloud fractions around 1300 CST. Cloud fraction generally decreases during the rest of the afternoon until deep convection propagates into the region during the late afternoon around 1800 CST. There are differences among the simulations. For example, at 1300 CST the variable-3 simulation has a lower median cloud fraction than the simulations with lower or no spatial soil moisture gradients. However, the maximum value of 0.72 is much larger than the maximum value from the other three simulations that are between 0.50 and 0.55. This suggests that the cloud fraction over the inner model domain is transitioning to a bimodal state. In other words there are more subregions with fewer clouds, but also a few subregions with more clouds that lead to a large cloud fraction. This trend is amplified in the variable-4 simulation in which the median drops even further and the maximum value continues to increase. This tendency among the simulations persists until 15 CST. Then, the variable-3 and variable-4 simulations begin to have higher median cloud fractions than the other three simulations. As the afternoon progresses, more and more clouds transition to deeper convection in those two simulations that slowly increase the number of subregions to have higher cloud fractions.

4. Discussion

Because we show that soil moisture is an important factor in controlling convective cloud populations on this day when the synoptic circulations are relatively weak, it is important to consider methods that could improve the initial soil conditions used by the LES model. Running an off-line land-surface model that is driven by observed atmospheric quantities for a sufficiently long period to generate realistic soil moisture and temperature distributions would have more spatial variability than distributions derived strictly from observations as shown in Figure 2b. For example, Lin et al. (2017) used the WRF-Hydro framework in which the
Noah Multi-Parameterization is driven by the North American Land Data Assimilation System-2 (Xia et al., 2012) analyses to evaluate streamflow and evapotranspiration over Texas during a three-year period. A similar model configuration was used by Somos-Valenzuela and Palmer (2018) to examine the water budget in the northeastern United States over a 36-year period. Nevertheless, the predicted soil quantities could still contain uncertainties that are attributed to assumptions used by the land-surface parameterizations and the available spatial and temporal resolution of the atmospheric analyses. The grid spacing of North American Land Data Assimilation System-2 is 1/8th degree which may still be too coarse to account for local variations in precipitation that are crucial in obtaining realistic variations in soil moisture. The specific spatial resolution of soil moisture needed by models is unclear since the idealized LES study of Lee et al. (2019) suggests that the impact of soil moisture variations depends on the strength of the background winds. As described by Santanello et al. (2016), data assimilation techniques could also be combined with these models to better constrain the soil conditions. Therefore, it would be useful to explore some of these techniques using the 30 August 2016 case to assess the sensitivity of predicted convective cloud populations to smaller-scale variations in soil moisture.

The impact of soil moisture on precipitation has been the subject of several studies (e.g., Findell & Eltahir, 2003; Guillod et al., 2015; Taylor et al., 2012). The strength and sign of the soil moisture-rainfall coupling over the SGP region are not consistent among those studies, ranging from strong positive to weak or no coupling (e.g., Ferguson & Wood, 2011; Findell et al., 2011; Koster et al., 2004; Zeng et al., 2010). Our results clearly depict a negative spatial coupling as discussed by Guillod et al. (2015) and lend support to the importance of the soil moisture gradient on convective initiation (Taylor et al., 2018; Weaver, 2004). It is interesting to note that the morning atmospheric soundings over most of our simulation domain are identified as wet-soil advantage for convection according to the framework by Findell and Eltahir (2003), which contrasts with our finding. We are currently investigating the implication of the key processes identified in our simulations (soil moisture variability and cold pools) for the land-atmosphere coupling metrics developed by Findell and Eltahir (2003) and others (see Santanello et al., 2017).

It is important to remember that moisture fluxes in the Noah land surface parameterization are determined from three pathways: (1) evaporation directly from the soil, (2) evaporation from the canopy, and (3) plant transpiration. The leaf area index (LAI) in the Noah parameterization depends on the land use type and a greenness fraction that is obtained by interpolating in space and time from MODIS climatology. Previous studies using observations from the SGP site suggest that the evaporative fraction (EF), defined as the ratio of the latent flux to the sum of the latent and sensible heat fluxes, is more strongly correlated with LAI and with surface soil moisture (e.g., Bagley et al., 2017; Phillips et al., 2017; Tang et al., 2018). In the revised-1 simulation, the correlation of EF with LAI and surface soil moisture was 0.47 and 0.29, respectively, which is consistent with many of the observed correlations in the previous studies. This result also points out that treatment of the LAI in the model and its effect on the relative partitioning of latent and sensible heat fluxes between the soil and vegetation that warrants further investigation in the future. In addition, we use the Noah land surface parameterization in this study, but we did not explore the sensitivity of the partitioning of latent and sensible heat fluxes to different treatments of land surface parameterizations and their

Figure 15. Cloud fraction over the outer model domain as a function of time among the five soil moisture sensitivity simulations shown in terms of percentiles at 30-min intervals. See text for a description of how cloud fraction is computed.
impact on cloud formation. In addition to vegetation controls on transpiration, the treatment of other factors such as surface roughness and aerodynamic resistance in land surface parameterizations will affect the surface fluxes.

While including higher variability in the initial temperature, moisture, and wind profile did not have a significant impact on the simulated cloud fields, the simulated cloud populations may be affected by uncertainties introduced by the National Center for Environmental Prediction FNL at the lateral boundaries. As mentioned previously, High-Resolution Rapid Refresh with its data assimilation and 3-km grid spacing would likely produce better boundary conditions than the coarse global FNL analyses. For example, the current initial conditions contain no information on cloud liquid water or ice, but satellite observations at 06 CST (12 UTC) show that stratiform and cirrus clouds, associated with the trough extending from eastern New Mexico to western Kansas, were present over the northwest corner of the outer WRF domain. Since the LES model did not form these clouds during the early morning, the simulated surface temperatures over the northwest corner of the domain were biased too high (Figure S14a) and the simulated humidity was too low (Figures S14b and S14c) which may have affected boundary layer properties and the regional wind field later in the afternoon. Nevertheless, the LES model did form areas of deep convection near the western boundary that slowly propagated eastward late in the afternoon, consistent with satellite images.

It is also possible that the choice of microphysics parameterization may affect the simulated shallow cloud populations and transitions from shallow to deep convection. Therefore, we performed another sensitivity simulation identical to the default configuration, except that the Morrison microphysics parameterization (Morrison et al., 2005, 2009) was used. The macroscopic properties of the shallow clouds (i.e., size, depth, spatial distribution) were very similar to the cloud fields when using the Thompson parameterization, as shown in Figures 3b and 4b, during the entire afternoon. In addition, neither simulation produced transitions to deeper convection during the middle of the afternoon as in the revised-1 simulation. There were differences in the droplet number concentrations, with the Morrison parameterization producing a much higher droplet number concentration for droplet diameters around 5 μm than the Thompson parameterization and the observations (Figure 6). Li et al. (2015) performed WRF simulations of precipitating shallow marine cumuli during the RICO experiment using both the Thompson and Morrison parameterizations found that the Morrison parameterization produced more and stronger cold pools. Therefore, it is possible that cloud populations predicted by the revised-1 simulation might be sensitive to the choice of microphysics treatment once cold pools start to form after 13 CST.

One factor not considered in this study that affects the observed droplet number distribution is the variability in aerosol concentrations. Fast et al. (2019) and Figure S8 show a large spatial gradient in aerosol number concentrations along the afternoon G-1 flight path on 30 August (Figure 1c). The highest concentrations greater than 8,000 cm$^{-3}$ just east of the Central Facility are likely due to emissions from nearby refineries and a coal-fired power plant. Aerosol concentrations further to the east are an order of magnitude lower. A comparison of two Condensation Particle Counter instruments with a 3- and 10-nm-diameter cutoff, respectively, reveals that about half of the number concentration is from aerosols smaller than 10 nm. Additional analyses of the measured aerosol number concentration, size distribution, and composition in relation to the CCN concentrations will reveal how aerosols affect the cloud droplet number concentrations. Such analysis will also provide a benchmark for future LES modeling studies that represent cloud-aerosol interaction processes.

The complex convective cloud populations on 30 August, as shown in Figure 1c, as well as the extensive measurements collected over north-central Oklahoma at the SGP site make this an ideal case to evaluate the performance of parameterizations of shallow and deep convection. These parameterizations are designed to change the vertical stability, generate and redistribute heat, remove and redistribute moisture, predict convective precipitation, and affect surface heating and atmospheric radiation associated with subgrid-scale convective clouds. Various assumptions and approaches are used by the parameterizations to account for multiple processes operating within one vertical column of the model grid.

As shown in this study, local-to-regional-scale variations in soil moisture and cold pools affect the convective cloud populations on 30 August in addition to larger-scale forcing mechanisms. To decide when and where to activate convection, parameterizations use a trigger function often based on convective available potential energy that depends on vertical profiles of temperature and humidity or a trigger function based on multiple
factors (e.g., Bombardi et al., 2016; Tawfik et al., 2015). As described by Zhang (2002), much of the spatial variability in convective available potential energy in the midlatitudes over land is due to thermodynamic changes in the boundary layer air. While convective parameterizations are inherently one-dimensional in the vertical, horizontal variations in soil moisture will indirectly contribute to spatial variations in the triggering of convection. As shown in Figure 11, wetter soil moisture cools the boundary layer that inhibits the formation of cumulus and conversely drier soil moisture leads to a relatively warmer boundary layer with more active turbulent eddies that favors cumulus formation. If a regional or global model grid adequately resolves the spatial variability of soil moisture and the boundary layer parameterization adequately represents the differences in turbulent mixing between the regions, then the convective parameterization should be able to represent the initial distribution of shallow cumulus as seen in our LES study. Nevertheless, it is likely that some subgrid-scale variability in soil moisture will be missed and not captured by weather and climate models. In addition, parameterized convective clouds are not advected across the model grid cells; therefore, convection will be fixed to the surface forcing contrary to the advection of cumulus from dry soil regions to wet soil regions as shown in Figure 11.

Representing cold pool effects in convective parameterizations is more complicated since many processes need to be accounted for that evolve over time. Since cold pools are driven by evaporating precipitation within downdrafts, there is microphysical component that is not normally included in many convective parameterizations. A range of approaches, from simple to comprehensive, have been developed to represent the impact of cold pools on convective clouds. Simple approaches may include only the effects of near-surface gustiness associated with convective precipitation (e.g., Zeng et al., 2002) or modify parameters in existing convective parameterizations (e.g., Hohenegger & Bretherton, 2011). Other approaches that represent subgrid-scale variability using probability density functions (e.g., Berg et al., 2013; Larson & Schanen, 2013) to indirectly account for some cold pool effects. Comprehensive approaches explicitly represent many cold pool mechanisms that evolve in time (e.g., Park, 2014; Schlenmer & Hohenegger, 2014). Many of these studies have shown improved performance in simulating convection, but a limited number of metrics have been used. For example, Rio et al. (2009) developed a density current parameterization to represent a population of expanding circular cold pools in a grid cell and show that the inclusion of this effect improved the timing of convective precipitation at the ARM site for a case during June 1997. Del Genio et al. (2015) describe modifications to the cumulus parameterization used in a climate model that regulate the occurrence of weakly entraining convection by cold pools. They show that its inclusion led to an improved simulation of column water versus cloud top height when compared to radar observations during the ARM MJO Investigation Experiment in the Indian Ocean during 2011. However, as far as we know, there has been no systematic evaluation of convective parameterizations that include cold pool effects for cloud populations such as those on 30 August 2016 (Figure 1c) in terms of spatial variability in cloud fraction and precipitation.

5. Summary and Conclusions

In this study, we use LES predictions in relation with ARM and satellite measurements to identify the primary processes responsible for the complex convective cloud population observed on one day during the HI-SCALE field campaign. We use a nested LES domain configuration large enough to encompass the variations in cloud size and amount with which the outer and inner domain is 297 (Δx = 300 m) and 120 km (Δx = 100 km) wide, respectively. Three primary simulations are performed, the first (default) with meteorology and soil property variations downscaled from global model analyses, a second (revised-1) that includes finer-scale vertical variations in meteorology and larger spatial variability in soil properties consistent with the network of soil measurements, and a third (revised-2) that includes only finer-scale vertical variations in initial atmospheric state. These simulations are evaluated using ground-based ARM measurements, HI-SCALE aircraft measurements, and MODIS satellite images of cloud distributions.

Overall, the three simulations qualitatively reproduce many of the observed boundary layer temperature, water vapor, and horizontal winds. The simulated distributions of updrafts and downdrafts in the boundary layer driven by convective eddies are qualitatively similar to those observed by aircraft and Doppler lidars, although the distribution from the revised-1 simulation that incorporates more spatial variability in soil moisture is in better agreement with the lidar observations at the E39 and E41 sites during the entire day and better at other lidar sites during the morning. The largest difference between the default and revised-
1 simulations, however, is in the predicted cloud macroscopic properties and spatial cloud distribution. The default simulation with soil moisture downscaled from a global model had relatively uniform cumulus fields over the inner WRF domain. In contrast, the revised-1 simulation produced more variability in cloud size and distribution more consistent with MODIS satellite images during the afternoon.

We show that the spatial heterogeneity in the cloud populations results primarily from two factors that become superimposed during the afternoon.

1. The first factor is soil moisture variability, and cooler lake temperatures to a lesser degree, that control cumulus formation early in the day. Regions with drier soil moisture had warmer boundary layers, higher boundary layer depths, and had more vigorous shallow cumulus forming earlier in the day than regions with wetter soil moisture. As cumulus formed over drier soil moisture regions are transported by easterly winds over wetter soil regions, they tend to dissipate during the morning and early afternoon. The boundaries of drier and wetter soil regions favor the transition of shallow cumulus to deeper, precipitating convection.

2. The second factor is cold pools that form after 13 CST. There are many precipitating convective cells that last only a few minutes and only perturb the local cloud fields. The fewer, longer-lasting, and stronger convective cells produce cold pools that are more persistent and expand over large areas that are tens of kilometers wide. The cold pools suppress turbulence and cloud formation so that clear skies are present over relatively large areas. The cloud distributions forced by soil moisture variations are perturbed by cold pools that expand and overlap each other. These processes are likely responsible for the complex observed population of convective clouds.

A cloud-tracking algorithm is used to quantify the cloud lifetime and size predicted by the two simulations. We show that the revised-1 simulation has larger clouds and longer lifetime during much of the day over the inner WRF domain. Additional sensitivity simulations were also performed using the outer WRF domain to examine the effect of smoother or more variable soil moisture than the revised-1 simulation. A comparison of these simulations also shows that smoother soil moisture variations led to more uniform cumulus distributions. While the median cloud fraction was similar among the simulations, larger soil moisture variability induced more and more regions with higher cloud fraction. This effect into a bimodal effect for soil moisture variations larger than the revised-1 simulation, with some regions having low cloud fractions and other regions having cloud fractions approaching 100%.

The results of this study suggest that the 30 August 2016 case during HI-SCALE would be useful to evaluate convective parameterizations used by larger-scale weather and climate models. Since the cloud distributions are likely governed by soil moisture distributions and cold pools, these are factors that will need to be assessed by such simulations when representing the cloud fraction, precipitation, and cloud radiative effects simulated by convective parameterizations.

Acknowledgments
The HI-SCALE field campaign was supported by the Atmospheric Radiation Measurement (ARM) Climate Research Facility and the Environmental Molecular Science Laboratory (EMSL), both are U.S. Department of Energy (DOE) Office of Science User Facilities sponsored by the Office of Biological and Environmental Research. The modeling studies were performed in EMSL’s Cascade computer. The WRF community model was made available from the National Center for Atmospheric Research (NCAR) at http://www2.mmm.ucar.edu/wrf/users/. This research was supported by the Atmospheric Science User Facilities sponsored by the DOE by the Battelle Memorial Institute. DOE’s National Laboratory is operated by DOE by the Battelle Memorial Institute, both are U.S. Department of Energy laboratories.

References
Bagley, J. E., Kueppers, L. M., Billesbach, D. P., Williams, I. N., Biraud, S. C., & Torn, M. S. (2017). The influence of land cover on surface energy partitioning and evaporative fraction regimes in the U.S. Southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 122, 5793–5807. https://doi.org/10.1002/2017JD026740
Benjamin, S., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Monthly Weather Review*, 144(4), 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1
Berg, L. K., Gustafson, W. J. Jr., Kassinanov, E. I., & Deng, L. (2013). Evaluation of a modified scheme for shallow convection: Implementation of CuP and case studies. *Monthly Weather Review*, 141(1), 134–147. https://doi.org/10.1175/MWR-D-12-00136.1
Berg, L. K., Newsom, R. K., & Turner, D. D. (2017). Year-long vertical velocity statistics derived from Doppler lidar in the continental convective boundary layer. *Journal of Applied Meteorology and Climatology*, 56, 2441–2454. https://doi.org/10.1175/JAMC-D-16-0359.1
Bogenschutz, P. A., & Krueger, S. K. (2013). A simplified pdf parameterization of subgrid-scale clouds and turbulence for cloud-resolving models. *Journal of Advances in Modeling Earth Systems*, 5, 195–211. https://doi.org/10.1002/jame.20018
Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Schanen, D. P., Meyer, N. R., & Craig, C. (2012). Unified parameterization of the planetary boundary layer and shallow convection with a higher-order turbulence closure in the community atmosphere model: Single-column experiments. *Geoscientific Model Development*, 5(6), 1407–1423. https://doi.org/10.5194/gmd-5-1407-2012
Bombardi, R. J., Tawfik, A. B., Manganello, J. V., Marx, L., Shin, C.-S., Halder, S., et al. (2016). The heated condensation framework as a convective trigger function in the NCEP Climate Forecast System version 2. *Journal of Advances in Modeling Earth Systems*, 8, 1310–1329. https://doi.org/10.1002/2016MS000668
Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., et al. (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, 128(582), 1075–1093. https://doi.org/10.1256/003590002320732110

https://doi.org/10.1029/2017JD026740
https://doi.org/10.1175/JAMC-D-16-0359.1
https://doi.org/10.1002/jame.20018
https://doi.org/10.1002/2016MS000668
https://doi.org/10.1256/003590002320732110
https://doi.org/10.1175/2017JD026740
https://doi.org/10.1175/MWR-D-15-0242.1
https://doi.org/10.1175/MWR-D-12-00136.1
https://doi.org/10.1002/jame.20018
https://doi.org/10.5194/gmd-5-1407-2012
Liu, S., & Shao, Y. (2013). Soil-layer configuration requirement for large eddy atmosphere and land surface coupled modeling. *Atmospheric Science Letters*, 14(2), 112–117. https://doi.org/10.1002/asl.426

Liu, Y., Warner, T., Liu, Y., Vincent, C., Wu, W., Mahoney, B., et al. (2011). Simultaneous nested modeling from the synoptic scale to the LES scale for wind energy applications. *Journal of Wind Engineering and Industrial Aerodynamics*, 99(4), 308–319. https://doi.org/10.1016/j.jweia.2011.01.013

Lohou, F., & Patton, E. (2014). Surface energy balance and buoyancy response to shallow cumulus shading. *Journal of the Atmospheric Sciences*, 71(2), 655–682. https://doi.org/10.1175/JAS-D-13-0145.1

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

McPherson, R. A., Fiebrich, C., Crawford, K. C., Elliott, R. L., Kilby, J. R., Grimley, D. L., et al. (2007). Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmospheric Oceanic Technology*, 24(3), 301–321. https://doi.org/10.1175/JTECH1976.1

Mirocha, J., Kosovic, B., & Kirkil, G. (2014). Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the Weather Research and Forecasting model. *Monthly Weather Review*, 142(2), 806–831. https://doi.org/10.1175/MWR-D-13-00064.1

Moeng, C.-H., Dudhia, J., Klemp, J., & Sullivan, P. (2007). Examining two-way grid nesting for large eddy simulation of the PBL using the WRF Model. *Monthly Weather Review*, 135(6), 2295–2311. https://doi.org/10.1175/MWR3406.1

Moeng, C.-H., & Sullivan, P. P. (2015). Large eddy simulation. In G. North, J. Pyle, & F. Zhang (Eds.), *Encyclopedia of Atmospheric Sciences*, (2nd ed., Vol. 4, pp. 232–240). Elsevier.

Morrison, H., Curry, J. A., & Khvorostyanov, V. I. (2005). A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *Journal of the Atmospheric Sciences*, 62(6), 1665–1677. https://doi.org/10.1175/JAS3446.1

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*, 137(3), 991–1007. https://doi.org/10.1175/2008MWR2556.1

Munoz-Esparza, D., & Kosovic, B. (2018). Generation of inflow turbulence in large-eddy simulations of nonneutral atmospheric boundary layers with the cell perturbation method. *Monthly Weather Review*, 146(6), 1889–1909. https://doi.org/10.1175/MWR-D-18-0077.1

Munoz-Esparza, D., Kosovic, B., van Beek, J., & Mirocha, J. (2015). A stochastic perturbation method to generate inflow turbulence in large-eddy simulation models: Application to neutrally stratified atmospheric boundary layers. *Physics of Fluids*, 27(3), 035102. https://doi.org/10.1063/1.4913572

Munoz-Esparza, D., Lundquist, J. K., Sauer, J. A., Kosovic, B., & Linn, R. R. (2017). Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign: From weather to boundary-layer eddies. *Journal of Advances in Modeling Earth Systems*, 9, 1572–1594. https://doi.org/10.1002/2017MS000960

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. (2000). Updated daily. NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6M043C6.

Neggers, R. A. J., Köhler, M., & Beljaars, A. C. M. (2009). A dual mass flux framework for boundary layer convection. Part I: Transport. *Journal of the Atmospheric Sciences*, 66(6), 1465–1487. https://doi.org/10.1175/2009JAS3635.1

Park, S. (2014). A unified convection scheme (UNICON). Part I: Formulation. *Journal of the Atmospheric Sciences*, 71(11), 3902–3930. https://doi.org/10.1175/JAS-D-13-0323.1

Phillips, T. J., Klein, S. A., Ma, H.-Y., Tang, Q., Xie, S., Williams, I. N., et al. (2017). Using ARM observations to evaluate climate model simulations of land-atmosphere coupling on the U.S. southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 122, 11,524–11,548. https://doi.org/10.1002/2017JD027141

Rai, R. K., Berg, L. K., Kosovic, B., Haupt, S. E., Mirocha, J. D., Ennis, B. L., & Draxl, C. (2019). Evaluation of the impact of horizontal grid spacing in Terra Incognita on coupled mesoscale-microscale simulations using the WRF Framework. *Monthly Weather Review*, 147(3), 1007–1027. https://doi.org/10.1175/MWR-D-18-0282.1

Rai, R. K., Berg, L. K., Kosovic, B., Mirocha, J. D., Pekour, M. S., & Shaw, W. J. (2017). Comparison of measured and numerically simulated turbulence statistics in a convective boundary layer over complex terrain. *Boundary Layer Meteorology*, 163(1), 69–89. https://doi.org/10.1007/s10546-016-0217-y

Rio, C., Hourdin, F., Grandpeix, J. Y., & Lafore, J. P. (2009). Shifting the diurnal cycle of parameterized deep convection over land. *Geophysical Research Letters*, 36, L07809. https://doi.org/10.1029/2008GL036779

Santanello, J. A., Dirmeyer, P. A., Ferguson, C. R., Finkell, K. L., Tawfik, A. B., Berg, A., et al. (2017). Land- atmosphere interactions: The LoCo perspective. *Bulletin of the American Meteorological Society*, 98(6), 1253–1272. https://doi.org/10.1175/BAMS-D-17-0001.1

Santanello, J. A. Jr., Kumar, S. V., Peters-Lidard, C. D., & Lawston, P. M. (2016). Impact of soil moisture assimilation on land surface model spin-up and coupled land–atmosphere prediction. *Journal of Hydrometeorology*, 17(2), 517–540. https://doi.org/10.1175/JHM-D-15-0072.1

Schalchlik, J., Griffith, E., Post, F. H., & Jonker, H. J. J. (2012). High-performance simulations of turbulent clouds on a desktop PC: Exploiting the GPU. *Bulletin of the American Meteorological Society*, 93(3), 307–314. https://doi.org/10.1175/BAMS-D-11-00059.1

Schalchlik, J., Jonker, H. J. J., Siebesma, A. P., & Van Meijgaard, E. (2015). Weather forecasting using GPU-based large-eddy simulations. *Bulletin of the American Meteorological Society*, 96(5), 715–723. https://doi.org/10.1175/BAMS-D-14-00114.1

Schlemmer, L., & Hohenegger, C. (2014). The formation of wider and deeper clouds as a result of cold-pool dynamics. *Journal of the Atmospheric Sciences*, 71(8), 2842–2858. https://doi.org/10.1175/JAS-D-13-0170.1

Segal, J., & Arritt, R. W. (1992). Nonclassical mesoscale circulations caused by surface sensible heat flux gradients. *Bulletin of the American Meteorological Society*, 73(10), 1593–1604. https://doi.org/10.1175/1520-0477(1992)073<1593:NNCMCS>2.0.CO;2

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). A description of the Advanced Research WRF version 3. *NCAR Tech. Note*. NCAR/TN-475+STR, 113 pp. https://doi.org/10.5065/D68S4MVH.

Somos-Valenzuela, M., & Palmer, R. N. (2018). Use of WRF-Hydro over the northeast of the U.S. to estimate water budget tendencies in small watersheds. *Water*, 10(12), 1799. https://doi.org/10.3390/w10121799

Talbot, C., Bou-Zeid, E., & Smith, J. (2012). Nested mesoscale large-eddy simulations with WRF. Performance in real test cases. *Journal of Hydrometeorology*, 13(5), 1421–1441. https://doi.org/10.1175/JHM-D-11-048.1
Tang, Q., Xie, S., Zhang, Y., Phillips, T. J., Santanello, J. A., Cook, D. R., et al. (2018). Heterogeneity in warm-season land-atmosphere coupling over the U.S. Southern Great Plains. Journal Geophysical Research: Atmospheres, 123, 7867–7882. https://doi.org/10.1029/2018JD028463

Tang, S., Xie, S., Zhang, M., Tang, Q., Zhang, Y., Klein, S. A., et al. (2019). Differences in eddy-correlation and energy-balance surface turbulent heat flux measurements and their impacts on the large-scale forcing fields at the ARM SGP site. Journal Geophysical Research: Atmospheres, 124, 3301–3318. https://doi.org/10.1029/2018JD029689

Tawfik, A. B., Dirmeyer, P. A., & Santanello, J. A. (2015). The heated condensation framework. Part I: Description and Southern Great Plains case study. Journal of Hydrometeorology, 16, 1929–1945. https://doi.org/10.1175/JHM-D-14-0117

Taylor, C. M., Jeu, D., Richard, A. M., Guichard, F., Harris, P. P., & Dorigo, W. (2012). Afternoon rain more likely over drier soils. Nature, 489(7416), 423–426. https://doi.org/10.1038/nature11377

Taylor, C. M., Prigent, C., & Dadson, S. J. (2018). Mesoscale rainfall patterns observed around wetlands in sub-Saharan Africa. Quarterly Journal of the Royal Meteorological Society, 144(716), 2118–2132. https://doi.org/10.1002/qj.3311

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. Monthly Weather Review, 136(12), 3095–3115. https://doi.org/10.1175/2008MWR2387.1

Tucker, S. C., Serff, C. J., Weickmann, A. M., Brewer, W. A., Banta, R. M., Sandberg, S. P., et al. (2009). Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles. Journal of Atmospheric and Oceanic Technology, 26, 673–688. https://doi.org/10.1175/2008JTECHA1157.1

Twomey, S. (1977). The influence of pollution on the shortwave albedo of clouds. Journal of the Atmospheric Sciences, 34(7), 1149–1152. https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOA>2.0.CO;2

Varble, A., Zipser, E. J., Fridlind, A. M., Zhu, P., Ackerman, A. S., Chaboureau, J.-P., et al. (2014). Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft properties. Journal of Geophysical Research: Atmospheres, 119, 13,891–13,918. https://doi.org/10.1002/2013JD021371

Weaver, C. P. (2004). Coupling between large-scale atmospheric processes and mesoscale land-atmosphere interactions in the U.S. Southern Great Plains during summer. Part I: Case studies. Journal of Hydrometeorology, 5(6), 1247–1258. https://doi.org/10.1175/Jhm-397.1

Wyngaard, J. C. (1984). Large-eddy simulation: Guidelines for its application to planetary boundary layer research. U.S. Army Research Office Contract 0804, pp. 122.

Wyngaard, J. C. (2010). Turbulence in the atmosphere. (p. 393). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511840524

Xia, Y., Mitcheh, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): I. Intercomparison and application of model products. Journal of Geophysical Research, 117, D03110. https://doi.org/10.1029/2011JD016048

Xiao, H., Berg, L. K., & Huang, M. (2018). The impact of surface heterogeneities and land-atmosphere interactions on shallow clouds over ARM SGP site. Journal of Advances in Modeling Earth Systems, 10(6), 1220–1244. https://doi.org/10.1175/2018MS001286

Xiao, H., Endo, S., Wong, M., Skamarock, W., Klem, J. B., Fast, J. D., et al. (2015). Modifications to WRF’s dynamical core to improve the treatment of moisture for large-eddy simulations. Journal of Advances in Modeling Earth Systems, 7, 1627–1642. https://doi.org/10.1002/2015MS000532

Yamaguchi, T., & Feingold, G. (2012). Technical note: Large-eddy simulation of cloudy boundary layer with the Advanced Research WRF model. Journal of Advances in Modeling Earth Systems, 4, M09003. https://doi.org/10.1029/2012MS000164

Yang, L., Jin, S., Danielson, P., Homer, C., Gasse, L., Bender, S. M., et al. (2011). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 108–123. https://doi.org/10.1016/j.isprsjprs.2018.09.006

Zeng, X., Barlage, M., Castro, C., & Fling, K. (2010). Comparison of land–precipitation coupling strength using observations and models. Journal of Hydrometeorology, 11(4), 979–994. https://doi.org/10.1175/2010JHM1226.1

Zeng, X., Zhang, Q., Johnson, D., & Tao, W. K. (2002). Parameterization of wind gustiness for the computation of ocean surface fluxes at different spatial scales. Monthly Weather Review, 130(8), 2125–2133. https://doi.org/10.1175/1520-0493(2002)130<2125:POWGFT>2.0.CO;2

Zhang, G. (2002). Convective quasi-equilibrium in midlatitude continental environment and its effect on convective parameterization. Journal of Geophysical Research, 107(D14), 4220. https://doi.org/10.1029/2001JD001005

Zhang, Y., & Klein, S. A. (2010). Mechanisms affecting the transition from shallow to deep convection over land: Inferences from observations of the diurnal cycle collected at the ARM Southern Great Plains site. Journal of the Atmospheric Sciences, 67(9), 2943–2959. https://doi.org/10.1175/2010JAS3661.1

Zhang, Y., & Klein, S. A. (2013). Factors controlling the vertical extent of fair-weather shallow cumulus clouds over land: Investigation of diurnal-cycle observations collected at the ARM Southern Great Plains site. Journal of the Atmospheric Sciences, 70(4), 1297–1315. https://doi.org/10.1175/JAS-D-12-0131.1

Zhang, Y., Klein, S. A., Fan, J., Chandra, A. S., Kollias, P., Xie, S., & Tang, S. (2017). Large-eddy simulation of shallow cumulus over land: A composite case based on ARM long-term observations at its Southern Great Plains site. Journal of the Atmospheric Sciences, 74(10), 3229–3251. https://doi.org/10.1175/JAS-D-16-0317.1