

# The Large-Eddy Simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) Activity for Continental Shallow Convection

William I. Gustafson Jr., Andrew M. Vogelmann, Zhijin Li, Xiaoping Cheng, Kyle K. Dumas, Satoshi Endo, Karen L. Johnson, Bhargavi Krishna, Tami Fairless, and Heng Xiao

**ABSTRACT:** The U.S. Department of Energy’s Atmospheric Radiation Measurement (ARM) user facility recently initiated the Large-Eddy Simulation (LES) ARM Symbiotic Simulation and Observation (LASSO) activity focused on shallow convection at ARM’s Southern Great Plains (SGP) atmospheric observatory in Oklahoma. LASSO is designed to overcome an oft-shared difficulty of bridging the gap from point-based measurements to scales relevant for model parameterization development, and it provides an approach to add value to observations through modeling. LASSO is envisioned to be useful to modelers, theoreticians, and observationalists needing information relevant to cloud processes. LASSO does so by combining a suite of observations, LES inputs and outputs, diagnostics, and skill scores into data bundles that are freely available, and by simplifying user access to the data to speed scientific inquiry. The combination of relevant observations with observationally constrained LES output provides detail that gives context to the observations by showing physically consistent connections between processes based on the simulated state. A unique approach for LASSO is the generation of a library of cases for days with shallow convection combined with an ensemble of LES for each case. The library enables researchers to move beyond the single-case-study approach typical of LES research. The ensemble members are produced using a selection of different large-scale forcing sources and spatial scales. Since large-scale forcing is one of the most uncertain aspects of generating the LES, the ensemble informs users about potential uncertainty for each date and increases the probability of having an accurate forcing for each case.

<https://doi.org/10.1175/BAMS-D-19-0065.1>

Corresponding author: William I. Gustafson Jr., [william.gustafson@pnnl.gov](mailto:william.gustafson@pnnl.gov)

In final form 8 January 2020

©2020 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

**AFFILIATIONS:** Gustafson and Xiao—Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, Washington; Vogelmann, Endo, Johnson, and Fairless—Environmental and Climate Sciences Department, Brookhaven National Laboratory, Upton, New York; Li—Jet Propulsion Laboratory, California Institute of Technology, Pasadena, and Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, Los Angeles, California; Cheng—College of Meteorology and Oceanography, National University of Defense Technology, Changsha, Hunan, China; Dumas and Krishna—Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee

Observation “supersites” are facilities where large numbers of instruments are deployed to provide an integrated measure of a wide array of related variables necessary for researchers to improve process understanding and inform model development. Often the sites aim to measure a wide range of temporal and spatial scales. Supersites focusing on clouds and aerosol have been deployed in many places around the world, and the value of the sites is clear from the many resulting publications. The U.S. Department of Energy’s (DOE) Atmospheric Radiation Measurement (ARM) user facility presently has long-term supersites in Oklahoma, in Alaska, and on Graciosa Island, with previous locations in the west Pacific Ocean and shorter-term, “mobile” deployments worldwide (Turner and Ellingson 2016). Examples of supersites from other organizations include the Cabauw Experimental Site for Atmospheric Research (CESAR) in the Netherlands (Leijnse et al. 2010; Sarna and Russchenberg 2017) and the Barbados Cloud Observatory (Stevens et al. 2016).

However, even the best-instrumented supersite cannot observe the atmosphere sufficiently to infer everything researchers want to know. Even with a growing number of observation supersites around the world providing data, researchers continue to struggle to connect localized observations to scales relevant for model parameterization development. High-resolution models provide a complete, 4D, physically consistent picture of the atmospheric state that, when constrained by observations, presents an opportunity to fill observational gaps and provide additional information for researchers. For example, most observations are point based whereas models provide volumetric representations of the atmosphere—having sufficient measurement samples to identify regional variability can be expensive, and often logistically or technically impossible, whereas models can simulate how processes and the atmospheric state vary over time and space. Thus, the synergy possible by combining observations with a carefully designed modeling framework can open up new research possibilities. A more wholistic view of the atmospheric system is provided with interconnected processes rendered within the model, and the observations provide real-world evaluations of how well the model captures reality. The modeling can also help interpret the observations by providing a dynamical context for the localized measurements to help bridge the scale gap to larger scales for which model parameterizations typically operate.

The Large-Eddy Simulation (LES) ARM Symbiotic Simulation and Observation (LASSO) activity is a recent undertaking by ARM that blends observation and modeling to better serve researcher needs (ARM 2017; Gustafson et al. 2017, 2019a). The LASSO concept has been motivated by previous work on the Royal Netherlands Meteorological Institute (KNMI) Parameterization Testbed (Neggers et al. 2012), where an LES model was routinely run alongside detailed atmospheric observations taken at CESAR. We have initially implemented LASSO to combine an LES with measurements from ARM’s Southern Great Plains (SGP) atmospheric observatory in Oklahoma for shallow convection conditions. The resulting observation–model combination is run routinely and expands upon ARM’s long history of data gathering to support climate modeling and research into atmospheric processes (Turner and Ellingson

2016). LASSO differs from the KNMI Testbed in that LASSO takes a focused approach by only simulating days meeting the chosen meteorological regime, which is currently shallow convection, and the results are packaged with a suite of data specifically tailored to this regime. The KNMI Testbed takes a more general modeling approach and simulates all days, leaving it to the user to determine if the model is appropriately configured for the conditions on a given day. Overall, the LASSO concept is meant to be adaptable, and future scenarios will be implemented over the coming years for other sites and phenomena.

LASSO combines a suite of information into “data bundles,” which include observations relevant to the particular modeling scenario, the inputs required to run the LES, LES output, quick-look plots, and skill scores and diagnostics that convey how the LES results compare with observed reality. The inclusion of detailed skill scores and diagnostics is unique to LASSO compared to earlier supersite modeling efforts, and these details are possible because of the extensive observations obtained by ARM. The data bundles are produced for many case dates, creating a library that enables statistical analyses beyond what would typically be possible with individual case studies. As described later, considerable effort has gone into facilitating the use of the data bundles that are freely available to the community. The data bundle approach provides significant value to researchers to simplify their use of LASSO.

### **The LASSO audience**

LASSO is designed with three types of user groups in mind: observationalists, modelers, and theoreticians. They each have different priorities and potentially approach LASSO with different needs and expectations, and would thus use the data bundles in different ways. We anticipate observationalists would have little interest in running an LES model themselves, and thus would use the LES output directly, either as “virtual truth” for testing retrieval approaches or, combined with the suite of observations associated with LASSO, to better understand the conditions around the supersite. Thus, an important aspect of the data bundles for observationalists is to include sufficient LES output to meet their needs combined with the most likely set of observations for their applications. In contrast, we anticipate modelers would be interested in the LES output from LASSO as a starting point for further model investigations and the observations would primarily be used to evaluate the simulations. Thus, in addition to the relevant LES output, modelers require the full set of model inputs and other information for reproducing the LES runs. For this reason, the data bundles contain the surface fluxes used for the lower boundary conditions along with the initial conditions and large-scale forcings used to integrate the model (the concept of LES and forcings is described in the “Application of LES to real atmospheres” sidebar). Theoreticians fall somewhat between observationalists and modelers in terms of data needs. They may or may not need to rerun the LES runs, and they are likely to be interested in using LASSO for process study investigations. Thus, theoreticians may be looking for more detailed model outputs, such as profiles of fluxes and budget information.

The broad target audience for LASSO presents a challenge to clearly communicating the available information and how to effectively use it. Experience over the last several years has revealed assumptions and misconceptions from different constituencies that sometimes require subtle education. This involves both the observations and the simulations and requires a balance between protecting new users from misapplying data versus allowing expert users to judge for themselves the appropriateness of data for specific purposes. Within LASSO, data are quality controlled but that does not minimize the need to understand how to properly interpret observations. New users are encouraged to review the LASSO documentation (Gustafson et al. 2019a), relevant ARM technical reports, and related references as well as to contact experts to help the users with new applications of the data. For example, some retrievals only work in certain situations and observational uncertainty can be difficult to convey and often requires

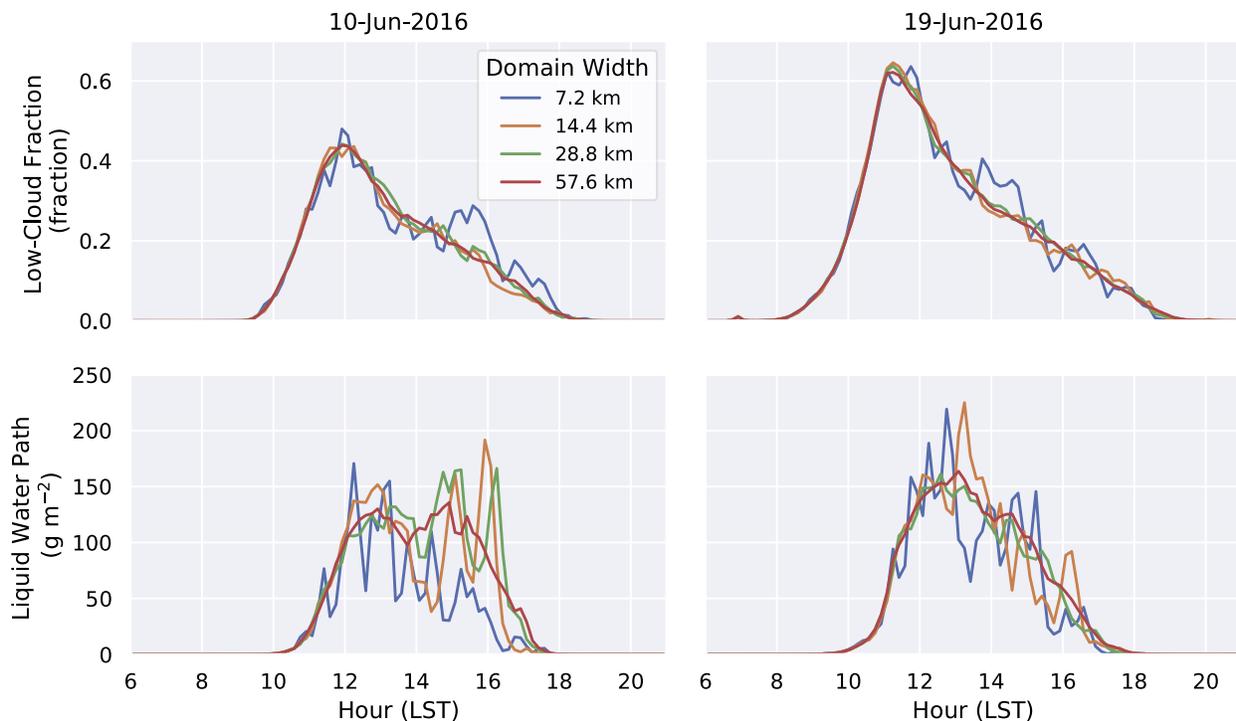
## Application of LES to real atmospheres

A typical LES domain is smaller than the spatial scales of synoptic and most mesoscale variability, so one traditionally assumes that the time variability of these large scales can be represented homogeneously throughout the LES domain. The implication is that the LES only provides submesoscale information, which is essentially small-scale detail on top of the mean coarser-scale information in which the LES resides. This permits use of a simpler domain configuration with “doubly periodic” boundaries, where the air mass that leaves one side of the domain reenters on the opposite side. For this configuration, spatially detailed boundary conditions are not used; instead, the mesoscale and synoptic influences are imposed as a “large-scale forcing” consisting of a single profile applied uniformly throughout the domain for each domain-mean tendency,  $\partial X/\partial t$ , of moisture and temperature. Optionally, this approach could be combined with “nudging,” where the mean model state is relaxed toward observations using a response time scale to offset the accumulation of error when large-scale conditions are not well known (e.g., Randall and Cripe 1999). We note that “nesting” can also be used for LES modeling such that a larger-scale model provides temporally and geographically varying flow along the lateral boundaries of the LES, much like how a regional NWP model is embedded within a global forecast model. The doubly periodic approach without nudging is used for LASSO.

Use of the traditional, doubly periodic LES approach has resulted in confusion over how to interpret and use LASSO LES

output. Unlike an NWP simulation, one cannot assume a point-to-point comparison between columns within the LES and a specific point on the ground. Instead, each column in a doubly periodic domain is statistically identical, such that nothing exists to differentiate one column from another. The best way to view the LES is as a statistical representation of the region experiencing the same large-scale forcing. Because all columns are identical from the large-scale perspective, another implication is that transitions between synoptic states manifest differently in an LES domain compared to an NWP domain. For example, instead of a frontal passage propagating across the domain, the entire LES domain sees the front at the same time as the average of the frontal influence over the forcing region for the particular moment.

Because every column is statistically identical, the size of the domain is somewhat irrelevant as long as it is large enough to statistically hold a cloud population consistent with the large-scale forcing. Once that is achieved, the primary advantage of a larger domain is the ability to have better statistical sampling. Figure SB1 shows that the results for domain-averaged, in-cloud liquid water path and cloud fraction are more consistent from time to time for larger domains. We found the cloud statistics to be too noisy with a 14.4-km domain, so we chose a 25-km domain for operational LASSO simulations to balance robustness of the results with computational cost (Gustafson et al. 2017).



**Fig. SB1.** Comparison of different LES domain sizes for (top) low-cloud fraction and (bottom) in-cloud LWP ( $\text{g m}^{-2}$ ) for (left) 10 and (right) 19 Jun 2016. Note the fluctuations of the domain-averaged cloud fraction and LWP for the 7.2- and 14.4-km domains tend to diminish for larger domains. These simulations are drawn from the Alpha 2 set of LASSO data using the SAM. The figure is an adaptation of plots in Gustafson et al. (2017).

context. Within LASSO, when possible, we use complementary observations to help make potential issues apparent. For example, cloud fraction measurements each have different biases and are notoriously difficult to measure in a way comparable to a model, so we include cloud fraction from the total-sky imager (TSI) (Morris 2005), which is based on a hemispheric image, alongside shallow-cloud fraction from ARM's Active Remotely-Sensed Cloud Locations (ARSCL) Value-Added Product (VAP) (Clothiaux et al. 2000), which is based on time averages from vertically pointing instruments. The chances of upper-level clouds contaminating the shallow-cloud fraction from the TSI is reduced when these two cloud fraction estimates correlate. Modelers should not blindly use the observations without first understanding how the measurements are made and any inherent assumptions. Likewise, expectations sometimes need to be tempered regarding what the LES generates and how to use the model output (see sidebar "Application of LES to real atmospheres" for examples). While we refer to the LES as virtual truth, it is still a model with built-in limitations and uncertainties. For example, the choice of microphysics or use of specified surface fluxes instead of an interactive land model can alter results, and the resolution is never fine enough to explicitly resolve all the relevant processes, especially surrounding microphysics, mixing, and entrainment that strongly impact shallow convection (e.g., Endo et al. 2019). Whenever the LES is used as a basis for understanding processes, or to serve as a proxy for the real-world meteorological state, one needs to determine whether the configuration is appropriate for the given use.

To date, we have seen users aligned with all three of the anticipated usage categories. For example, the first published, observation-based application of LASSO uses LASSO LES output as proxies for cloud fields to improve cloud-radar scan strategies (Oue et al. 2016). The first published use of LASSO for model development is by Angevine et al. (2018), who use LASSO to evaluate planetary boundary layer parameterizations in a single-column model framework. A use in line with the theoretician mindset is an investigation of aerosol-cloud interactions in shallow cumuli, in which the large number of available LASSO forcing data are used in conjunction with observed aerosol to improve the ability to statistically identify aerosol-cloud interactions in simulations while accounting for covariability in meteorology and other conditions (Glenn et al. 2020). Often the line between modeling and theory blurs with the modeling leading to new theory, such as with the examination of cloud spatial organization and size distributions by Neggers et al. (2019). In their study, they draw from five LASSO cases and rerun them with an altered domain configuration to enable larger cloud structures and refined resolution. Users have also considered how to use LASSO to inform issues involving 3D radiation and vertical velocity in relation to clouds (Endo et al. 2019; Gristey et al. 2020).

In addition to individual researchers, other projects are beginning to incorporate LASSO data into their workflows. For example, the Developmental Testbed Center's Global Model Test Bed (GMTB), which is part of the Common Community Physics Package (CCPP; <https://dtcenter.org/community-code/common-community-physics-package-ccpp>) being developed for use with the next-generation forecast models, now includes LASSO forcing data within its distribution (Firl et al. 2019). This permits modelers to run the GMTB single-column model for LASSO cases as physics parameterizations are developed and improved, which helps address the issue of having an end-to-end parameterization development workflow. This synergistic coupling of activities will extend the reach of LASSO, bringing the ARM data into the NOAA and NCAR communities where considerable model parameterization development occurs.

### **Core LASSO concepts**

The choice of how LASSO has been implemented for shallow convection is founded on five core concepts. The first is that individual case studies are insufficient to gain robust understanding. We instead approach LASSO as a library of cases from which to draw statistically

robust conclusions. This improves upon the typical LES approach of using single, finely tuned case studies with unknown representativeness, such as the now-classic continental shallow cumulus case based on SGP data that is part of the Global Energy and Water Cycle Experiment Cloud System Study (GCCS) Intercomparison (Brown et al. 2002). The value of the library approach to running LES has been demonstrated in different ways. Schalkwijk et al. (2015) take the approach of running every day of 2012 over Cabauw, the Netherlands, whereas van Laar et al. (2019) selectively run 146 shallow convection cases for the Jülich Observatory for Cloud Evolution (JOYCE) supersite in western Germany (Löhnert et al. 2015). LASSO follows the latter approach of focusing on a particular phenomenon instead of attempting to model all weather regimes.

The second core concept is that model inputs should not be fine-tuned on a case-by-case basis to make the model output match observations. If the LES differs from observations, there are likely multiple causes, and adjusting one or more details empirically or by trial and error can mask underlying modeling issues, resulting in misinterpretations. Instead of hand tuning initial conditions, surface fluxes, or large-scale forcings to make the model match observations, we instead use quality-controlled observations to assess the model results in the context of the observations and their uncertainties. Where possible, we use synergistic information from different instruments, such as multiple estimates of cloud fraction, noted above, as well as a combination of surface temperature and humidity from meteorological stations alongside mid–boundary boundary layer temperature and humidity from the Raman lidar (Newsom et al. 2013; Turner et al. 2002).

We address the issue of uncertainty in the large-scale forcing, which arguably is the largest contributor to model success on a day-to-day basis, in a unique way for our LES modeling. LASSO employs an ensemble of large-scale forcing datasets, as outlined in the “Large-scale forcings used by LASSO” sidebar. Thus, for each case date with shallow convection, we

## Large-scale forcings used by LASSO

The large-scale forcings used by LASSO come from three sources. Two are NWP based, while the third uses a forecast product as a background field that is optimally adjusted to be consistent with observations.

The first forcing derives from the ECMWF Integrated Forecast System (IFS) and uses the Diagnostics in the Horizontal Domains (DDH) system to extract closed budget terms to construct the large-scale forcing for three scales. Operationally available observations constrain the results via 4D variational data assimilation. The smallest extracted forcing scale is the single column closest to the SGP Central Facility, which has a scale of 9 km for the 2017 cases. Also extracted are the average forcing over 114- and 413-km scales.

The second NWP-derived forcing uses the MSDA methodology to directly incorporate a selection of ARM observations into the data assimilation process (Li et al. 2015a,b, 2016). The LASSO MSDA configuration uses the Gridpoint Statistical Interpolation (GSI) software associated with WRF in a 3D variational setup with 2-km grid spacing. A unique aspect of MSDA is a scale-separation algorithm that optimizes the observation error covariances to the grid spacing of each nested domain to produce high-resolution gridded estimates of the atmosphere. The Global Forecast System analyses are used as the initial background field. Ingested data includes conventional and satellite observations used by NOAA for operational data assimilation plus ARM radiosonde profiles

and horizontal winds retrieved from the radar wind profilers. Large-scale average forcings from MSDA are extracted at scales of 75, 150, and 300 km.

The third forcing source is the ARM constrained variational analysis (VARANAL) product (Xie et al. 2004; Zhang and Lin 1997; Zhang et al. 2001). VARANAL uses the NOAA Rapid Refresh (RAP) analyses as a background field that is combined with observations using a variational analysis methodology to obtain an estimate of the meteorological state for a given region. The methodology directly incorporates ARM data, such as surface fluxes, to constrain the energetic balance of the atmosphere. VARANAL is generated for a 300-km spatial scale.

The forcings are derived from fundamentally different methodologies, and thus do not always match. Of note, MSDA and VARANAL both directly incorporate ARM observations, but these methodologies optimally use different observation types. The MSDA is good at ingesting profiles and point-based information about the meteorological state. In contrast, VARANAL is influenced more by flux information to constrain the gridded background field. Since LASSO shallow convection days have no precipitation, the rain rate does not constrain VARANAL for LASSO. While rain rate is a very strong constraint in VARANAL on rainy days, surface fluxes play a much larger role on the shallow convection days without rain.

produce an ensemble of LES using a range of forcing scales and sources. We interpret the ensemble differently than for ensembles in weather forecasting. When forecasting, one does not know a priori which ensemble member is closest to reality, and the ensemble average is considered the best predictor with each member being an equally plausible forecast within the ensemble spread. With LASSO, the ensemble is produced after reality has happened, and we have observations to indicate the success and failure of each member. Thus, even though each large-scale forcing used to generate the ensemble is considered equally plausible, the model output is clearly differentiable as good or bad, and thus can be used accordingly. Additionally, the ensemble spread provides one measure of the relative uncertainty of the large-scale forcing for the particular day. Practically, the ensemble approach enables the operational production of the quality LES by eliminating the time needed to manually tune the forcings for each day, which would be untenable when producing many simulations per year.

The third core concept is that the data need to be easily usable. This drives the choice of data bundles for packaging the different types of datasets and makes it easy for users to obtain what they need. LASSO has a tiered approach with the bundles separated into two pieces. Metadata, model inputs, observations, summarized model output, skill scores, quick-look plots, and diagnostics are tarred into a relatively small file, which is about 30 MB per simulation and easily downloadable. This meets the needs of users who do not need full 4D model output, for example, if they are doing initial analyses to identify which specific cases they want. Output of the full model volume every 10 min and the associated LES statistics are tarred into a second file, which is three orders of magnitude larger at around 30 GB, and this file contains information for more detailed studies.

The fourth core concept is that users should be able to easily find and retrieve the specific simulations they need for their research, or more succinctly, “discovery and deliverability” should be quick and easy. Toward this end, a substantial amount of time during the pilot project was spent determining the appropriate set of observations to include in the data bundles, combined with methods for users to see how well each simulation behaves. A set of skill scores is generated for each simulation, which is explained in detail in the LASSO technical documentation (Gustafson et al. 2019a). The skill scores and associated quick-look plots form the foundation for information served via the LASSO Bundle Browser interface (<http://archive.arm.gov/lassobrowser>; Fig. 1). Users can query the Browser for specific metadata values, such as dates, forcing details, and skill score values. Graphs at the top of the page dynamically respond to search criteria, while matching data bundles are listed in a table at the bottom of the page. Users can select data bundles for download directly from the table, with options for delivery via FTP, THREDDS, or Globus. The latter is the recommended transfer method, and it is particularly efficient for users with access to a Globus endpoint to receive the data.

The fifth, and final, core concept is reproducibility. In addition to the basic ethical driver of scientific integrity, the LASSO implementation is designed to ensure other modelers can reproduce the LES runs, as well as generate variants of them. The ability to easily reproduce the simulations is driven by the target audience of modelers. Rarely would a single simulation from an event serve the needs of most modeling endeavors. For example, users wanting to design new parameterizations might need to test a range of tunable parameters to identify model sensitivity. Based on the core LASSO LES, they can determine which case dates and forcings they want to use and have an initial indication of how well the model should behave. This can be used as a launching point for their simulations. To encourage additional modeling with LASSO input data, we provide a conversion script to convert LASSO’s Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) LES inputs provided with the data bundles into the input format of the System for Atmospheric Modeling (SAM) (Khairoutdinov and Randall 2003). Over time, we hope to build a library of conversion scripts for additional models.

Select All (Excludes Date)

- DATE
- MEASUREMENTS
- MODEL TYPE
- OUTPUT DOMAIN SIZE
- NUMBER OF LEVELS
- LARGE SCALE FORCING
- LARGE SCALE FORCING SCALE
- INITIAL CONDITION
- SURFACE TREATMENT
- MICROPHYSICS

Submit

Overview Plots

07-20-2017

Heat Maps

Skill Score

GOES Satellite Images

**Introduction**

The LASSO Bundle Browser is designed to assist users with identifying LASSO data bundles and associated large-eddy simulations (LES) of interest for their research. Below, the plots and data table update dynamically based on user search criteria, and links within the table enable direct access to order the data bundles of the displayed simulations. Static plots associated with each data bundle can be accessed via the "Diagnostics" links in the table. More information on LASSO and the data bundles can be found at the LASSO home page. A more detailed description of the browser can be found in the LASSO documentation. [Click here to access, filter, and download the LASSO metadata table.](#)

Measurement: Cloud Fraction TSI

Date: July 20, 2017
  Model Type: WRF
  Model Type: SAM
  Output Domain Size: 7.2 km
  Output Domain Size: 14.4 km
  Output Domain Size: 25 km
  Output Domain Size: 28.8 km
  Output Domain Size: 57.6 km
  Number of Levels: 220 levels
  Large Scale Forcing: ECMWF
  Large Scale Forcing: MSDA
  Large Scale Forcing: VARANAL
  Large Scale Forcing: None
  Large Scale Forcing Scale: 16 km
  Large Scale Forcing Scale: 75 km
  Large Scale Forcing Scale: 114 km
  Large Scale Forcing Scale: 150 km
  Large Scale Forcing Scale: 300 km
  Large Scale Forcing Scale: 413 km
  Large Scale Forcing Scale: N/A
  Initial Condition: Sounding
  Surface Treatment: ECMWF
  Surface Treatment: MSDA
  Surface Treatment: VARANAL
  Microphysics: LASSO Morrison
  Microphysics: Morrison 2-MOM (M2005)
  Microphysics: Thompson

Taylor Diagram

Skill Score

Time Series

Scatter Plot

**Skill Scores**

Copy CSV Print PDF Search:

Measurement Skill

0  1

1D Cloud Skill

0  1

2D Cloud Mask Skill

0  1

Total Cloud Skill

0  1

Slide orange sliders to adjust min and max skill ranges

Date	Simulation ID	Links	Measurement Skill (Cloud Fraction TSI)	1D Cloud Skill	2D Cloud Mask Skill	Total Cloud Skill	Download
07-20-2017	2	(Diagnostics)	0.61	0.41	0.48	0.44	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	3	(Diagnostics)	0.55	0.42	0.55	0.48	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	4	(Diagnostics)	0.46	0.33	0.36	0.35	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	5	(Diagnostics)	0.53	0.35	0.42	0.38	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	6	(Diagnostics)	0.53	0.53	0.6	0.56	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	7	(Diagnostics)	0.63	0.51	0.48	0.49	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar
07-20-2017	8	(Diagnostics)	0.68	0.55	0.4	0.47	<input type="checkbox"/> Config Obs Model Tar <input type="checkbox"/> Raw Model Tar

Showing 1 to 7 of 7 entries

Previous  Next

Select All Download Options Order Data

Fig. 1. Illustration of the LASSO Bundle Browser web page, which serves the purpose of providing simple access to model behavior with dynamic queries. Once users identify simulations of interest, they can download the associated data bundles using the table at the bottom of the page. Links in the table also provide the metadata for each simulation and access to quick-look plots.

### Description of LASSO for shallow convection

What does it mean to routinely produce LES for shallow convection? While LASSO produces an “operational” product, this does not mean simulations are produced daily as in weather forecasting or as was done for the KNMI Testbed. Instead, the current approach is to process case dates for which shallow convection occurred. The typical period with shallow convection at the SGP runs from April to September from which days with shallow convection are identified and processed.

Shallow convection days are defined similarly to the criteria defined in the climatologies developed by Berg and Kassianov (2008) and Zhang and Klein (2013). Based on these criteria, an algorithm has been developed that routinely runs to automatically identify potential shallow convection days, available from the ARM archive as the “shallowcumulus” VAP ([www.arm.gov/capabilities/vaps/shallowcumulus](http://www.arm.gov/capabilities/vaps/shallowcumulus); Lim et al. 2019). LASSO uses this as guidance and applies additional criteria, such as whether critical ARM observations are available for the day, to ultimately select which days to simulate. Further, as the focus is on fair-weather shallow cumuli driven primarily by surface forcing, we seek cloud fields that are somewhat homogeneous in the surrounding region. Satellite data are used to exclude cases with pronounced large-scale heterogeneity within the several-hundred-kilometer region around the SGP. This is because the heterogeneity cannot be captured in regionally averaged large-scale forcing, which represents an average of the conditions over scales up to 413 km in our ensemble of large-scale forcings, which is the largest averaging area used for the LASSO forcings; forcings over small scales are more impacted by natural variability and therefore have increased sampling noise within the region.

Admittedly, the case selection process has a bit of subjectivity and some days with shallow convection at the SGP might be excluded. The site is located in a region where strong meteorological gradients can occur resulting in it sometimes being located on the dividing line between two different meteorological regimes throughout much of the day. Philosophically, one could argue to include such days since shallow convection is observed. However, we try to avoid days where the forcing is unlikely to be uniform across the simulated region with a goal of avoiding muddled forcing data that might be inconsistent with the localized observations. Some days are difficult to choose, as cloud development propagates over time, so regional disparities occur in some of the LASSO cases. Fully capturing regional heterogeneity might be accomplished using a nested LES approach instead of the doubly periodic lateral boundaries currently employed.

Once cases have been selected, forcing data for the LES are compiled from a range of sources to construct an ensemble, as detailed in sidebar “Large-scale forcings used by LASSO.” Each of the three different forcing sources uses different methods to blend observations to estimate the large-scale meteorological conditions around the SGP, providing variability in the estimated large-scale-forcing input, with each method being a plausible representation of reality. We further select three spatial scales from the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Multiscale Data Assimilation (MSDA) sources to account for heterogeneity in the meteorology around the region. In total, the operational ensemble currently consists of eight members, including one member that uses no large-scale forcing. This latter member only uses the initial conditions combined with the time-varying surface fluxes, described next; the tendencies from the large-scale forcings are zero for this member. Profiles for the LES initial conditions are identical for each member and are taken from the morning radiosonde at 1200 UTC [0600 local standard time (LST)].

In addition to large-scale atmospheric forcing, LASSO uses time-varying, spatially homogeneous surface fluxes obtained from the ARM Variational Analysis (VARANAL) product (Xie et al. 2004). VARANAL includes fluxes derived from the network of Eddy Correlation Flux Measurement (ECOR) and Energy Balance Bowen Ratio (EBBR) stations situated around the SGP. The data from these stations are averaged using a method that takes into account the spatial density of the flux measurements to produce a single flux value every 30 min that is representative of the region, the “Zhang” method described in Tang et al. (2019) and Zhang et al. (2001). The same regionally averaged surface fluxes are used for all LESs, regardless of the forcing scales provided for the large-scale-forcing data.

Using the observed surface fluxes for the LES lower boundary condition can insulate the simulations from radiative errors arising from any incorrectly simulated clouds, which would

impact the land surface fluxes. The prescribed surface fluxes replace an interactive land model, simplifying the model initialization and avoiding issues with handling land heterogeneity and uncertainty in the soil moisture state. On the downside, not using an interactive lower boundary precludes physical processes related to land–atmosphere interactions, such as surface shading by clouds, which can dampen formation of cloud water content (Xiao et al. 2018), or shading-induced changes to the Bowen ratio, which can impact boundary layer growth and relative humidity (Zhang and Klein 2013).

The operational configuration uses the WRF Model and a domain configuration chosen via a prototype phase where sensitivity tests were run (Gustafson et al. 2017). The domain configuration uses 100-m horizontal grid spacing and a domain extent of 25 km with doubly periodic lateral boundaries. The chosen grid spacing has been successfully used for multiple cloud simulation intercomparisons (e.g., Siebesma et al. 2003; Stevens et al. 2001), but may not be sufficient for all users. For example, if users desire to calculate entrainment associated with the shallow clouds, the LASSO simulations can serve as a guide from which to select case dates with “good” forcings; then, these users could generate new simulations with higher resolution to meet their research needs. The vertical grid spacing is 30 m up to 5-km height and stretches to 300 m near the model top near 15 km.

The specific version of WRF used for LASSO derives from the DOE Fast-Physics System Testbed and Research (FASTER) Project and includes LES-specific modifications such as extra outputs for traditional LES statistics, for example, domain-average fluxes and cloud details (Endo et al. 2015), along with an overall update to WRF, version 3.8.1. Full details regarding the model setup can be found in the “Description of the LASSO data bundles product” (Gustafson et al. 2019a).

An important challenge has been putting together an optimal suite of observations for model evaluation. By providing a core suite of observations to the users, they are saved from needing to go through the available options and deal with any quality-control issues in the data. In building the dataset, a balance is needed between the observations used to constrain the model initial conditions and forcings versus the observations used for model evaluation. Additionally, one needs to find ways to deal with sampling issues to account for spatial heterogeneity and differences in measurement methodologies, for example, cloud fraction measured via a time series versus instantaneously in space. For the most part, we decided to use less frequent measurements as part of the data assimilation process, such as the radiosondes, and high-frequency measurements for evaluation, such as from the Doppler lidars. This was motivated by the fact that the three-dimensional variational analysis used by the MSDA only assimilates data every 3 h. Thus, the MSDA cannot take advantage of the high sampling frequencies of the Raman and Doppler lidars, while the four-times-daily radiosonde launches are too infrequent to contribute much toward evaluating boundary layer growth compared to the subhourly mid–boundary layer thermodynamics measured by the Raman lidar. The high-frequency measurements also better reflect the scales of motion and variability simulated by the LES.

To better capture the spatial variability around the SGP, ARM established four “extended facilities” located approximately 50 km from the SGP Central Facility (Fig. 2). Each of these extended facilities includes surface meteorological instrumentation, a Doppler lidar, a three-channel microwave radiometer (MWR), and at three of the stations, an atmospheric emitted radiance interferometer (AERI). The Doppler lidars are used within LASSO to generate a regional estimate of cloud-base height and future development could extend their use to characterize boundary layer vertical velocity variance. The ultimate goal for the hyperspectral infrared measurements from the AERIs is to obtain regional estimates of the boundary layer temperature for the lowest couple kilometers of the atmosphere and, when combined with MWR data, retrievals of liquid water path (LWP) (Turner and Löhnert 2014) based on the AERI Optimal Estimation (AERIOe) VAP.

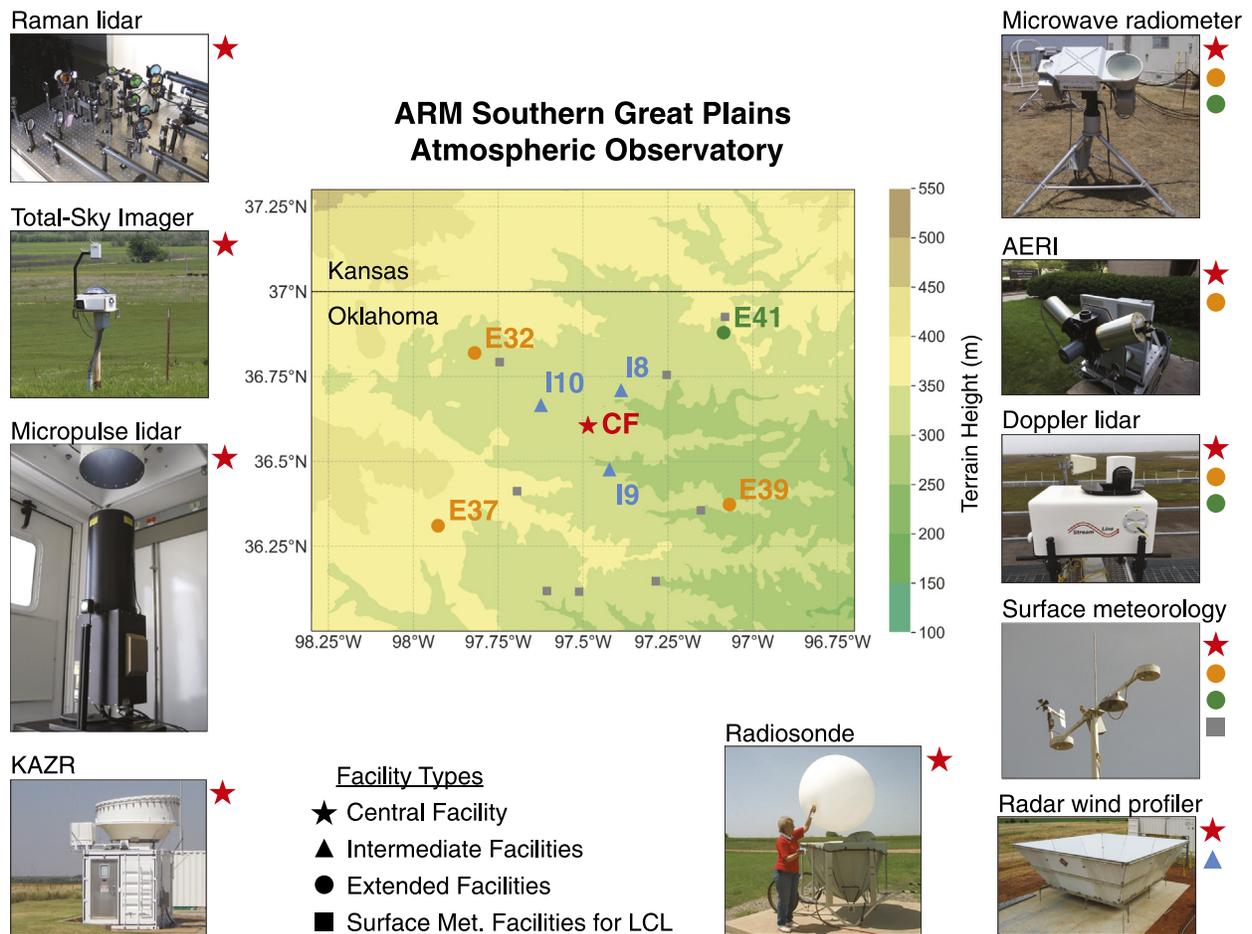


Fig. 2. The ARM Southern Great Plains atmospheric observatory consists of a regionally distributed set of instrumentation from which LASSO draws a suite of measurements particularly relevant for shallow convection. The instruments shown in this figure produce measurements used within LASSO for model input and skill scores. The shape of the symbols on the map indicate the type of facility and the colors indicate particular combinations of available instruments at each site, as indicated by the symbols next to each image. Site identification numbers are provided for the main ARM instrument clusters. The surface meteorology facilities used to calculate LCLs are from a combination of ARM and Oklahoma Mesonet sites. Images are courtesy of the DOE ARM user facility.

Observation of a time–height cloud mask is compared with the LES cloud field. The locations of cloud layers in a vertical column over the Central Facility are obtained from ARSCL, which combines data from active remote sensors to produce an objective determination of hydrometeor height (Clothiaux et al. 2000). The algorithm uses the Ka-band ARM zenith-pointing radar (KAZR) combined with micropulse lidar (MPL) and ceilometer data.

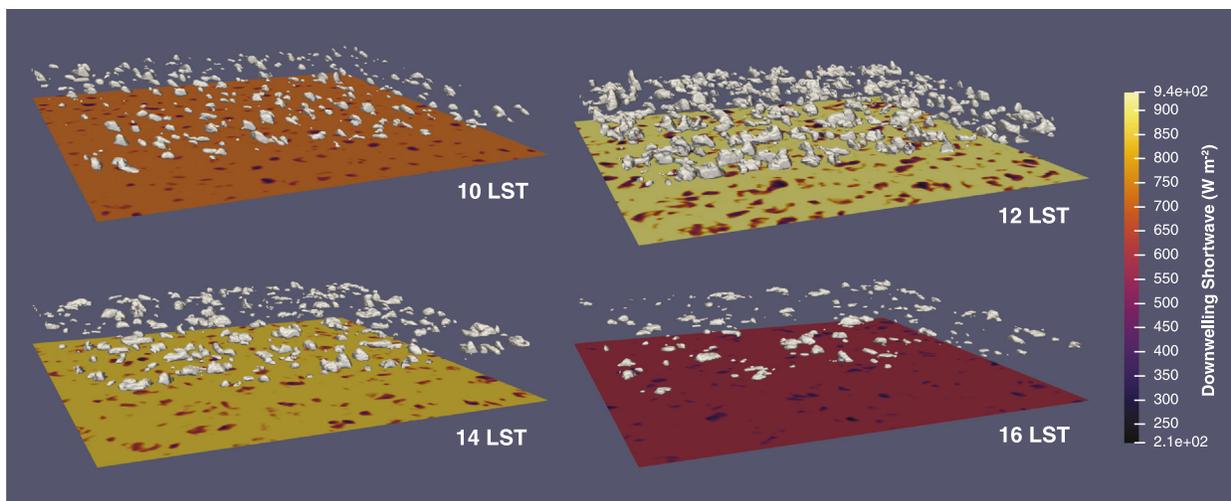
### Available LASSO data and general LES behavior

The production of LASSO data bundles began for the 2015 summer season and has been run to present. A total of 78 case days spanning 2015–18 are available to date for researchers to download, with the number continuing to grow. The LASSO modeling and observations have evolved slightly over time and earlier years include sensitivity simulations that were part of the testing done during development, such as the domain size, grid spacing, microphysics parameterization, and WRF versus SAM for the LES model. In total, 1172 data bundles are available for download for the 4 years of cases, with the current approach producing eight ensemble members per case date. Details are captured in the “Release history and change log description” in appendix A of Gustafson et al. (2019a). All of the example data in this paper come directly from the LASSO dataset, and more detailed descriptions of the methodology

behind applying the measurements can be found in “Description of the LASSO data bundles product” (Gustafson et al. 2019a).

Given that the current LASSO implementation targets shallow convection, many of the LASSO simulations have similar behavior with differences coming from the timing and amount of cloud. Figure 3 shows what a common cloud field looks like from the LES over the course of the day, which in this case is from a simulation driven by the VARANAL forcing on 30 August 2017. The simulated shallow cumuli form between 0900 and 1000 LST, peak around 1200 LST, and then decay throughout the remainder of the afternoon, with little-to-no cloud organization evident on this day. Overall, the LES cloud field forms roughly 2 h too early compared to the observed clouds, but the cloud fraction magnitude is roughly correct compared to the cloud fraction from the TSI (not shown). Other simulations in this day’s ensemble range from having very little shallow cloud much of the day to significantly over generating cloud in the afternoon.

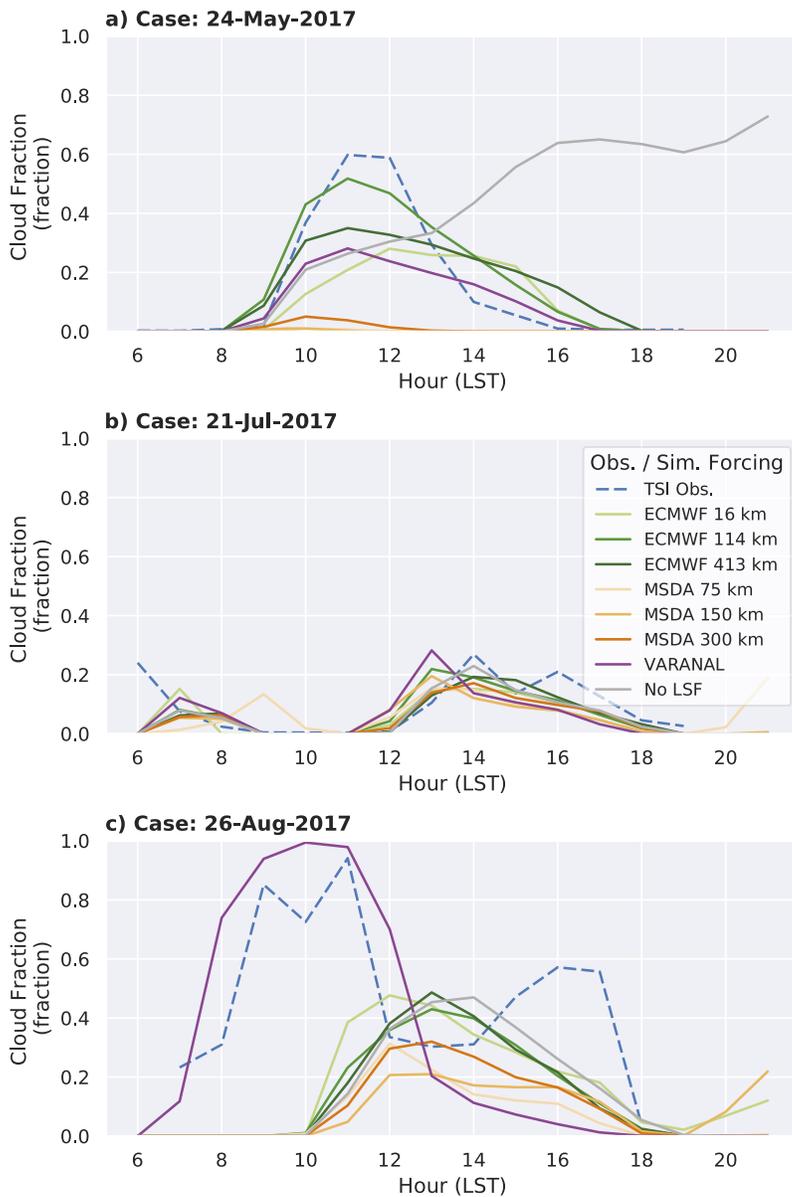
The variability within a given day’s ensemble can differ noticeably from day to day. In a perfect case, all ensemble members would converge on the observed conditions; for example, on 21 July 2017 all members have similar cloud fractions (Fig. 4b). However, a more common



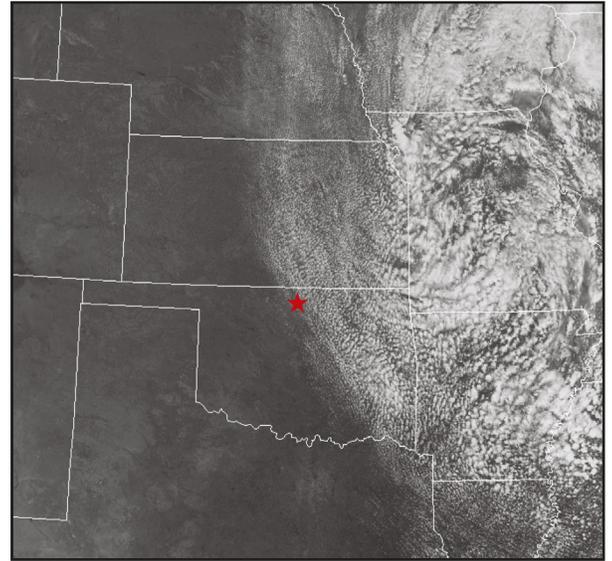
**Fig. 3.** Renderings of cloud water isosurfaces ( $10^{-6} \text{ kg}^{-1}$ ), every 2 h, show the diurnal evolution of a cloud field from a simulation forced by the VARANAL large-scale forcing on 30 Aug 2017. Cloud shadows can be seen in the surface downwelling shortwave radiation (colors;  $\text{W m}^{-2}$ ).

situation is where one or more members deviate. For example, on 24 May 2017 (Fig. 4a), most simulations follow the typical diurnal cycle for shallow cumuli, yet this is not what happens in reality. Only the ECMWF 114-km forcing captures the early peak in cloud fraction and a couple simulations produce almost no cloud. Interestingly, the simulation using no large-scale forcing differs substantially by having an increasing cloud fraction almost the entire day. This is an example of how important the large-scale forcing can be throughout the day. On 24 May 2017, the SGP location lies on the western edge of the shallow convection region (Fig. 5a), which impacts how the region is averaged when calculating the large-scale forcing—cloudy and noncloudy locations get mixed into a single forcing that blurs the observed heterogeneity. In contrast, the cloud field is more uniform around the SGP on 21 July 2017 (Fig. 5b) when the ensemble members behave more uniformly.

The 26 August 2017 case is another interesting situation where most ensemble members generate a typical shallow cumuli diurnal cycle, yet one member is a clear outlier in terms of cloud fraction (Fig. 4c). The simulation forced by the VARANAL does a much better job capturing the midlevel clouds during the morning that are likely due to impacts of a residual moisture layer generated by nearby deep convection. Time–height cross sections of cloud



a) GOES Visible, 24 May 2017, 17:45 UTC



b) GOES Visible, 21 July 2017, 19:45 UTC

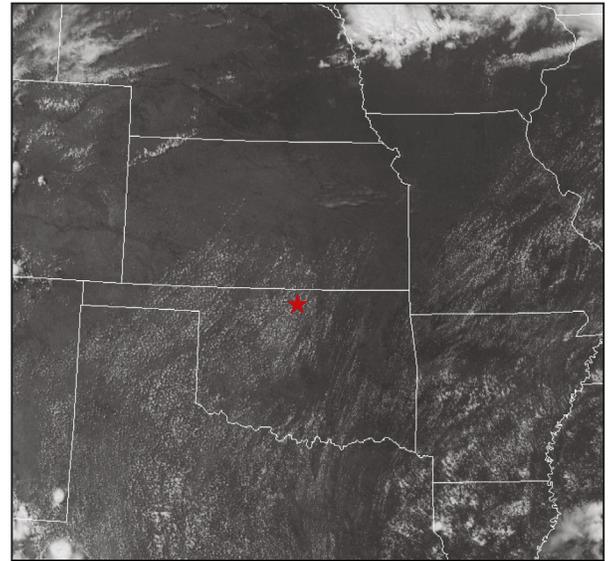


Fig. 4. Comparison of cloud fraction from the TSI (dashed blue) with the LES ensemble for 3 days from 2017.

Fig. 5. GOES visible images from (a) 1745 UTC 24 May and (b) 1945 UTC 21 Jul 2017. These are cropped images from aviationweather.gov, which are also available from the LASSO Bundle Browser. The SGP Central Facility is indicated by the red stars.

fraction derived from the ARSCL product show a cloud layer between about 1.5 and 4 km that is distinct from the shallow convection that forms later in the day (Fig. 6). The domain-averaged LES cloud fraction for the VARANAL-forced simulation also forms layered clouds in the morning that are separate from the surface-driven clouds. However, none of the other ensemble members contain this type of cloud feature, demonstrating the value of the ensemble approach.

Figure 7 further highlights a sampling of the range of ARM instrumentation combined with the LES in the LASSO data bundles. Shown are the diurnal cycles of percentiles for the 30 case days from 2017 for observations and the associated 240 simulations, which are segmented by large-scale forcing. Four variables have been chosen that highlight different aspects of the overall behavior of the simulations and a sample of ARM observational capabilities relevant to LASSO. The lifting condensation level (LCL; Fig. 7a) is a regional average benefiting from the 13 surface meteorological sites in a 60 km × 60 km region around the SGP (see Fig. 2). All simulations are initialized using the radiosonde profile from the Central Facility at 0600 LST,

## Cloud Fraction from ARSCL vs. LES Forced by Variational Analysis, 26 August 2017

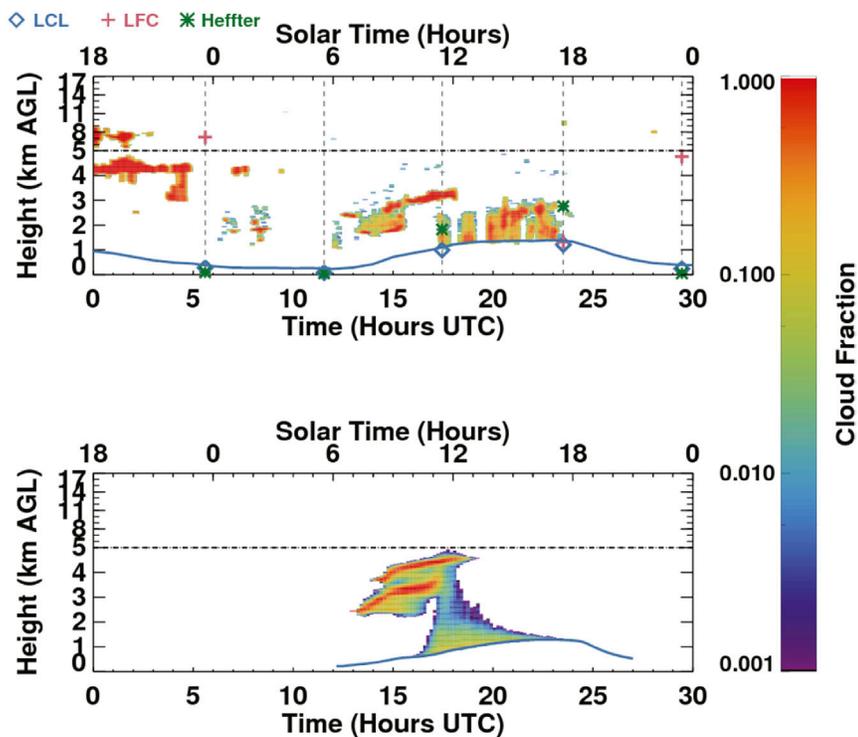


Fig. 6. An example of comparison plots from the data bundles showing (top) the ARSCL-derived time–height cloud fraction and (bottom) the LES cloud fraction for the ensemble member forced by the VARANAL for 26 Aug 2018. The ARSCL observations show cloud conditions during the night prior to the LES initialization at 1200 UTC combined with the cloud fraction for the simulated daytime period. The ARSCL cloud fraction is based on vertically pointing lidar and radar that is sampled in time, whereas the LES cloud fraction is an instantaneous domain average. Note the variable vertical resolution used in the plot: below 5 km AGL uses an expanded scale and above 5 km AGL uses a compressed scale. The blue lines in each panel indicate the LCL. In the top panel, red plus signs indicate the level of free convection and green stars indicate the boundary layer height as determined by the Heffter method computed from sonde data.

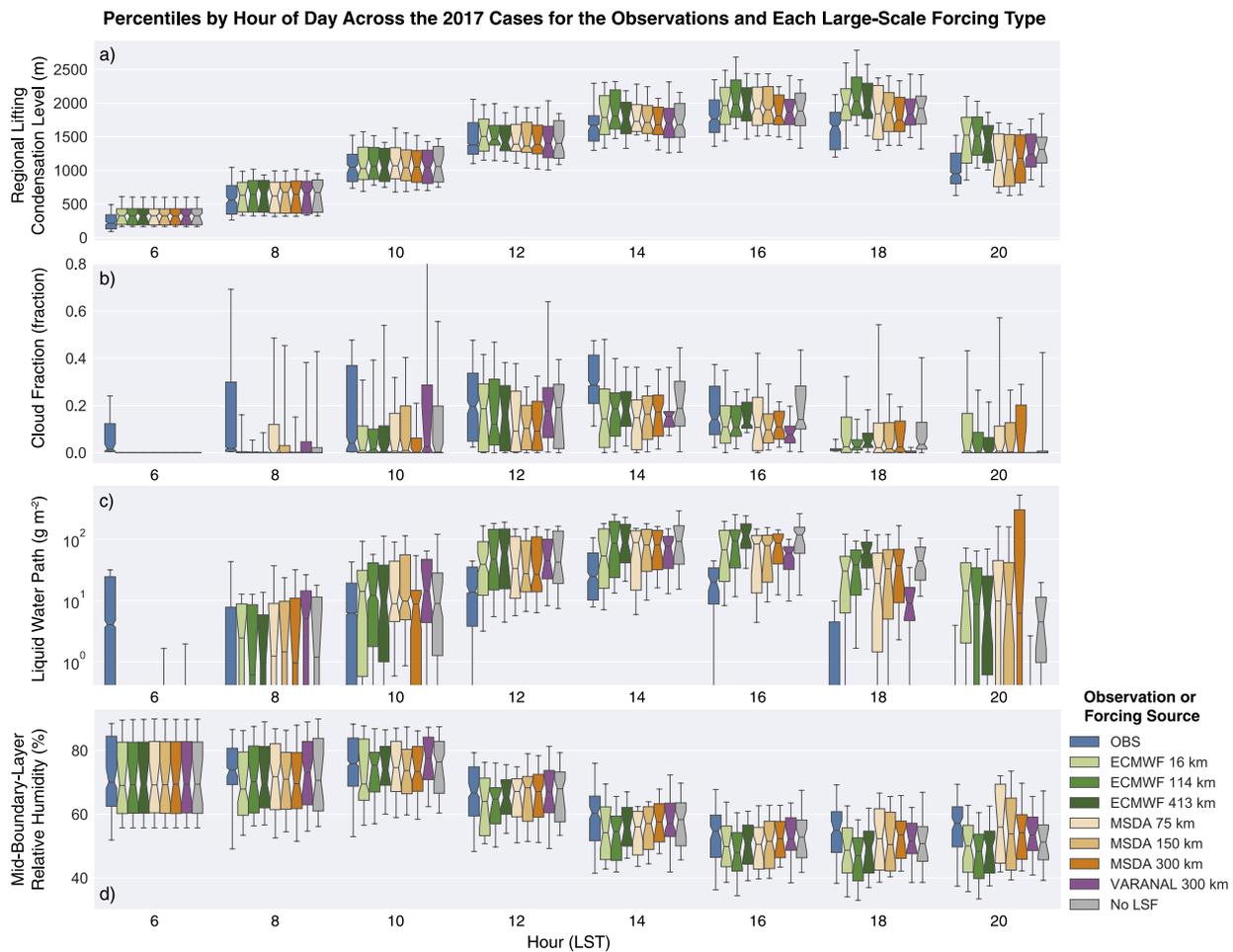
which on average has a higher LCL than the regional average estimated from the surface stations. Throughout the day, the different forcings evolve and the ECMWF simulations tend to maintain a high LCL bias while the MSDA simulations end the day closer to the regional average. Clouds are more impacted by model spinup than the LCL, so the first couple simulated hours of cloud properties should be treated as a spinup period when they are more uncertain. The cloud fraction (Fig. 7b) shows a near-zero median until 1000 LST, after which shallow clouds develop, peak in midafternoon, and then decay going into the evening. For much of the day, the simulations tend toward a lower cloud fraction than observed until the observed clouds fully decay around 1800 LST. The MSDA median cloud fraction drops similarly, but it is more likely to retain clouds longer into the evening than observed, with the 75th percentile near 0.1 at 1800 and 2000 LST compared to the observed value near 0. The ECMWF members also retain clouds more than observed, but not to such the extent of the MSDA members.

Complementing these measurements is the AERIOe retrieval of in-cloud LWP (Fig. 7c). Prior to 1200 LST the observations show a large variability in the observed LWP from day to day, with LWP increasing to several tens of grams per square meter from 1200 to 1600 LST, followed by smaller values later in the afternoon. The simulations mirror this behavior for much of the day, with a slight high bias during midday, and a tendency for clouds to linger too long in the late afternoon.

The last highlighted measurement is the mid–boundary layer relative humidity (RH) from the Raman lidar, determined from the mid–boundary layer temperature and moisture retrievals (Fig. 7d). This analysis shows that the ECMWF forcing has a tendency toward lower RH than the other forcings.

### Future of LASSO

The LASSO framework has been under development since 2015 and has reached a state of routine production. At this point, the model configuration and core set of observations have been established, and we expect the format of the data bundles to be stable for the shallow



**Fig. 7.** Percentiles by hour shown as box-and-whisker plots for the observations and LES results across the 30 case dates from 2017. Variables shown are (a) the LCL (m), (b) cloud fraction (fraction), (c) in-cloud LWP ( $\text{g m}^{-2}$ ), and (d) mid–boundary layer relative humidity (%). Observations are in blue and the large-scale forcings for the LES are shown by color as follows: different ECMWF forcing scales in green, different MSDA forcing scales in orange, VARANAL in purple, and no large-scale forcing (LSF) in gray. Simulated values are domain averages except for LWP, which is in-cloud only. Observed LCL values are estimated from 13 nearby surface meteorological stations and spatially averaged. Observed cloud fraction is from the TSI retrieval of opaque cloud fraction. Observed LWP is based on a hybrid retrieval methodology using the MWR and AERI data. The observed boundary layer relative humidity is retrieved using the Raman lidar and represents approximately the value at the middle of the boundary layer. The top and bottom of the boxes indicate the 25th and 75th percentiles, the middle bar is the median, and the whiskers indicate the 10th and 90th percentiles. Notches in the bars indicate median confidence intervals based on bootstrap sampling.

convection scenario. Changes over time could include additional observations as they become available and the model code may be updated as new WRF versions are released. For at least the next couple years, LASSO will continue to be run for shallow convection to fill out a large library of available cases. Current information on available cases can be found at the LASSO website ([www.arm.gov/capabilities/modeling/lasso](http://www.arm.gov/capabilities/modeling/lasso)), along with data bundles for download via the LASSO Bundle Browser (<https://adc.arm.gov/lassobrowser>).

The intention from the beginning of LASSO has been for it to be a portable framework for use at more than one ARM facility. Now that the shallow convection scenario is operational and much of the infrastructure needs are being implemented, the ARM facility is considering additional scenarios where LASSO can add value to ARM’s observations. Examples suggested during the LASSO Expansion Workshop, held in May 2019 (Gustafson et al. 2019b), include maritime clouds at ARM’s East North Atlantic atmospheric observatory in the Azores, Arctic clouds during the Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAIC)

field campaign (Shupe et al. 2018), deep convection during the Cloud, Aerosol, and Complex Terrain Interactions (CACTI) field campaign (Varble et al. 2018), and clear-air turbulence and boundary layer transitions at the SGP. Of these four scenarios, it has been decided to develop the CACTI scenario starting in 2020, with the other scenarios to be developed over time.

We note that each scenario offers potential to enhance ARM's capabilities in different areas, and synergies exist between the scenarios. For example, the clear-air turbulence scenario would contribute to understanding the environment leading up to shallow convection. An open question surrounding this scenario involves how to handle the land–atmosphere interactions. If an interactive land model were used, this implies the use of a downscaling approach with nests instead of the doubly periodic lateral boundaries currently used for the shallow convection scenario. Lessons learned developing the CACTI scenario, which will use an interactive land model and nested domains, will be valuable when developing the clear-air turbulence scenario. Likewise, efforts to better understand profiles of cloud condensation nuclei for a maritime scenario will help address similar data needs in the Arctic scenario. Overall, enacting LASSO for these new scenarios would entail determining the optimal set of observations to couple with modeling, identifying one or more satisfactory large-scale forcings for the given region, reconfiguring the model to meet the scenario-specific science drivers, and modifying the skill scores to be more appropriate for the given meteorological regime.

The ultimate utility of the LASSO framework will be determined by how researchers use it. We have been encouraged by the early adopters who have begun using LASSO in creative ways, and we look forward to seeing how such uses evolve over the coming years. We are also excited to see additional related interests in the starting of new observation–model endeavors, such as through the recently funded Ruisdael Observatory project (<http://ruisdael-observatory.nl/>). This project will expand upon CESAR with additional observation sites in the Netherlands plus accompanying ambitious modeling efforts.

**Acknowledgments.** Funding has been provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research, via the Atmospheric Radiation Measurement facility. We acknowledge the advice from external members of the LASSO Advisory Team: Maike Ahlgrimm (Deutscher Wetterdienst), Chris Bretherton (University of Washington), Graham Feingold (NOAA ESRL), Chris Golaz (LLNL), David Turner (NOAA ESRL), Minghua Zhang (Stony Brook University), and James Mather (ARM Technical Director).

We gratefully acknowledge the large number of ARM infrastructure team members outside of the authors that it takes to conduct an activity such as LASSO. People have contributed in multiple capacities as follows—for contributing new or custom-processed observations: David Turner, Laura Riihimaki, Tim Shippert, K. Sunny Lim, Virendra Ghate, Jonathan Helmus, and Rob Newsom; for proving ECMWF-based forcing data: Maike Ahlgrimm; for processing of VARANAL forcing data: Shaocheng Xie and Shuaiqi Tang; for providing data archive plus computing software and hardware support: Robert Records, Michael Giansiracusa, Jitu Kumar, Lynn Ma, and Aifang Zhou; for management support and working on the SGP reconfiguration and computing facilities that enabled LASSO: James Mather, Jennifer Comstock, and Giri Prakash; and for communication support: Hanna Goss, Rolanda Jundt, Robert Stafford, and Stacy Larsen.

Portions of the work were performed at 1) Pacific Northwest National Laboratory (PNNL)—Battelle Memorial Institute operates PNNL under contract DEAC05-76RL01830, 2) Oak Ridge National Laboratory (ORNL)—UT-Battelle, LLC operates ORNL for the DOE under contract DE-AC05-00OR22725, 3) Brookhaven National Laboratory, and 4) the Jet Propulsion Laboratory and University of California, Los Angeles, with the latter two via subcontracts through PNNL. Computation has been provided by 1) the ARM Data Center Computing Facility, 2) the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC05-00OR22725, 3) the National Energy Research Scientific Computing Center, a DOE Office of Science user facility supported under Contract DE-AC02-05CH11231, and 4) PNNL Research Computing.

## References

- Angevine, W. M., J. Olson, J. Kenyon, W. I. Gustafson, S. Endo, K. Suselj, and D. D. Turner, 2018: Shallow cumulus in WRF parameterizations evaluated against LASSO large-eddy simulations. *Mon. Wea. Rev.*, **146**, 4303–4322, <https://doi.org/10.1175/MWR-D-18-0115.1>.
- ARM, 2017: LASSO data bundles: 36°36'18.0"N, 97°29'6.0"W: Southern Great Plains central facility (C1). ARM Data Archive, accessed September 2019, <https://doi.org/10.5439/1342961>.
- Berg, L. K., and E. I. Kassianov, 2008: Temporal variability of fair-weather cumulus statistics at the ACRF SGP site. *J. Climate*, **21**, 3344–3358, <https://doi.org/10.1175/2007JCLI2266.1>.
- Brown, A. R., and Coauthors, 2002: Large-eddy simulation of the diurnal cycle of shallow cumulus convection overland. *Quart. J. Roy. Meteor. Soc.*, **128**, 1075–1093, <https://doi.org/10.1256/003590002320373210>.
- Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.*, **39**, 645–665, [https://doi.org/10.1175/1520-0450\(2000\)039<0645:ODOCHA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2).
- Endo, S., and Coauthors, 2015: RACORO continental boundary layer cloud investigations: 2. Large-eddy simulations of cumulus clouds and evaluation with in situ and ground-based observations. *J. Geophys. Res. Atmos.*, **120**, 5993–6014, <https://doi.org/10.1002/2014JD022525>.
- , and Coauthors, 2019: Reconciling differences between large-eddy simulations and Doppler lidar observations of continental shallow cumulus cloud-base vertical velocity. *Geophys. Res. Lett.*, **46**, 11 539–11 547, <https://doi.org/10.1029/2019GL084893>.
- Firl, G., L. Carson, L. Bernardet, and D. Heinzeller, 2019: Global Model Test Bed Single Column Model (SCM) user and technical guide v3.0. Developmental Testbed Center, 31 pp., <https://dtcenter.org/GMTB/v3.0/scm-cpp-guide-v3.pdf>.
- Glenn, I. B., G. Feingold, J. J. Gristey, and T. Yamaguchi, 2020: Quantification of the radiative effect of aerosol-cloud-interactions in shallow continental cumulus clouds. *J. Atmos. Sci.*, in press., <https://doi.org/10.1175/JAS-D-19-0269.1>.
- Gristey, J. J., G. Feingold, I. B. Glenn, K. S. Schmidt, and H. Chen, 2020: Surface solar irradiance in continental shallow cumulus cloud fields: Observations and large-eddy simulation. *J. Atmos. Sci.*, **77**, 1065–1080, <https://doi.org/10.1175/JAS-D-19-0261.1>.
- Gustafson, W. I., A. M. Vogelmann, X. Cheng, S. Endo, B. Krishna, Z. Li, T. Toto, and H. Xiao, 2017: Recommendations for the implementation of the LASSO workflow. DOE/SC-ARM-17-031, DOE Atmospheric Radiation Measurement Climate Research Facility, 62 pp., [www.arm.gov/publications/programdocs/doe-sc-arm-17-031.pdf](http://www.arm.gov/publications/programdocs/doe-sc-arm-17-031.pdf); <https://doi.org/10.2172/1406259>.
- , and Coauthors, 2019a: Description of the LASSO data bundles product. DOE/SC-ARM-TR-216, DOE Atmospheric Radiation Measurement Climate Research Facility, 122 pp., [www.arm.gov/publications/tech\\_reports/doe-sc-arm-tr-216.pdf](http://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-216.pdf); <https://doi.org/10.2172/1469590>.
- , A. M. Vogelmann, and J. H. Mather, 2019b: Science drivers and proposed modeling approaches for future LASSO scenarios. DOE/SC-ARM-19-023, DOE Atmospheric Radiation Measurement Climate Research Facility, 114 pp., [www.arm.gov/publications/programdocs/doe-sc-arm-19-023.pdf](http://www.arm.gov/publications/programdocs/doe-sc-arm-19-023.pdf); <https://doi.org/10.2172/1569273>.
- Khairoutdinov, M. F., and D. A. Randall, 2003: Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *J. Atmos. Sci.*, **60**, 607–625, [https://doi.org/10.1175/1520-0469\(2003\)060<0607:CRMOTA>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0607:CRMOTA>2.0.CO;2).
- Leijnse, H., and Coauthors, 2010: Precipitation measurement at CESAR, the Netherlands. *J. Hydrometeorol.*, **11**, 1322–1329, <https://doi.org/10.1175/2010JHM1245.1>.
- Li, Z. J., S. Feng, Y. Liu, W. Lin, M. Zhang, T. Toto, A. M. Vogelmann, and S. Endo, 2015a: Development of fine-resolution analyses and expanded large-scale forcing properties: 1. Methodology and evaluation. *J. Geophys. Res. Atmos.*, **120**, 654–666, <https://doi.org/10.1002/2014JD022245>.
- , J. C. McWilliams, K. Ide, and J. D. Farrara, 2015b: A multiscale variational data assimilation scheme: Formulation and illustration. *Mon. Wea. Rev.*, **143**, 3804–3822, <https://doi.org/10.1175/MWR-D-14-00384.1>.
- , X. P. Cheng, W. I. Gustafson, and A. M. Vogelmann, 2016: Spectral characteristics of background error covariance and multiscale data assimilation. *Int. J. Numer. Methods Fluids*, **82**, 1035–1048, <https://doi.org/10.1002/flid.4253>.
- Lim, K.-S. S., and Coauthors, 2019: Long-term retrievals of cloud type and fair-weather shallow cumulus events at the ARM SGP site. *J. Atmos. Oceanic Technol.*, **36**, 2031–2043, <https://doi.org/10.1175/JTECH-D-18-0215.1>.
- Löhnert, U., and Coauthors, 2015: JOYCE: Jülich Observatory for Cloud Evolution. *Bull. Amer. Meteor. Soc.*, **96**, 1157–1174, <https://doi.org/10.1175/BAMS-D-14-00105.1>.
- Morris, V. R., 2005: Total sky imager handbook. Tech. Rep. TR-017, DOE Atmospheric Radiation Measurement Climate Research Facility, 12 pp., [www.arm.gov/publications/tech\\_reports/handbooks/tsi\\_handbook.pdf](http://www.arm.gov/publications/tech_reports/handbooks/tsi_handbook.pdf).
- Neggers, R. A. J., A. P. Siebesma, and T. Heus, 2012: Continuous single-column model evaluation at a permanent meteorological supersite. *Bull. Amer. Meteor. Soc.*, **93**, 1389–1400, <https://doi.org/10.1175/BAMS-D-11-00162.1>.
- , P. J. Griewank, and T. Heus, 2019: Power-law scaling in the internal variability of cumulus cloud size distributions due to subsampling and spatial organization. *J. Atmos. Sci.*, **76**, 1489–1503, <https://doi.org/10.1175/JAS-D-18-0194.1>.
- Newsom, R. K., D. D. Turner, and J. E. M. Goldsmith, 2013: Long-term evaluation of temperature profiles measured by an operational Raman lidar. *J. Atmos. Oceanic Technol.*, **30**, 1616–1634, <https://doi.org/10.1175/JTECH-D-12-00138.1>.
- Oue, M., P. Kollias, K. W. North, A. Tatarevic, S. Endo, A. M. Vogelmann, and W. I. Gustafson, 2016: Estimation of cloud fraction profile in shallow convection using a scanning cloud radar. *Geophys. Res. Lett.*, **43**, 10998–11 006, <https://doi.org/10.1002/2016GL070776>.
- Randall, D. A., and D. G. Cripe, 1999: Alternative methods for specification of observed forcing in single-column models and cloud system models. *J. Geophys. Res.*, **104**, 24 527–24 545, <https://doi.org/10.1029/1999JD900765>.
- Sarna, K., and H. W. J. Russchenberg, 2017: Monitoring aerosol–cloud interactions at the CESAR Observatory in the Netherlands. *Atmos. Meas. Tech.*, **10**, 1987–1997, <https://doi.org/10.5194/amt-10-1987-2017>.
- Schalkwijk, J., H. J. J. Jonker, A. P. Siebesma, and F. C. Bosveld, 2015: A year-long large-eddy simulation of the weather over Cabauw: An overview. *Mon. Wea. Rev.*, **143**, 828–844, <https://doi.org/10.1175/MWR-D-14-00293.1>.
- Shupe, M., and Coauthors, 2018: The Multidisciplinary Drifting Observatory for the Study of Arctic climate (MOSAic) atmosphere science plan. DOE/SC-ARM-18-005, DOE Atmospheric Radiation Measurement Climate Research Facility, 31 pp., [www.arm.gov/publications/programdocs/doe-sc-arm-18-005.pdf](http://www.arm.gov/publications/programdocs/doe-sc-arm-18-005.pdf).
- Siebesma, A. P., and Coauthors, 2003: A large eddy simulation intercomparison study of shallow cumulus convection. *J. Atmos. Sci.*, **60**, 1201–1219, [https://doi.org/10.1175/1520-0469\(2003\)60<1201:ALESIS>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2).
- Skamarock, W. C., and Coauthors, 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>.
- Stevens, B., and Coauthors, 2001: Simulations of trade wind cumuli under a strong inversion. *J. Atmos. Sci.*, **58**, 1870–1891, [https://doi.org/10.1175/1520-0469\(2001\)058<1870:SOTWCU>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<1870:SOTWCU>2.0.CO;2).
- , and Coauthors, 2016: The Barbados Cloud Observatory anchoring investigations of clouds and circulation on the edge of the ITCZ. *Bull. Amer. Meteor. Soc.*, **97**, 787–801, <https://doi.org/10.1175/BAMS-D-14-00247.1>.
- Tang, S. Q., S. C. Xie, M. H. Zhang, Q. Tang, Y. Y. Zhang, S. A. Klein, D. R. Cook, and R. C. Sullivan, 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. *J. Geophys. Res. Atmos.*, **124**, 3301–3318, <https://doi.org/10.1029/2018JD029689>.

- Turner, D. D., and U. Löhnert, 2014: Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based atmospheric emitted radiance interferometer (AERI). *J. Appl. Meteor. Climatol.*, **53**, 752–771, <https://doi.org/10.1175/JAMC-D-13-0126.1>.
- , and R. G. Ellingson, Eds., 2016: The Atmospheric Radiation Measurement (ARM) Program. Meteor. Monogr., No. 57, Amer. Meteor. Soc., 504 pp.
- , R. A. Ferrare, L. A. H. Brasseur, and W. F. Feltz, 2002: Automated retrievals of water vapor and aerosol profiles from an operational Raman lidar. *J. Atmos. Oceanic Technol.*, **19**, 37–50, [https://doi.org/10.1175/1520-0426\(2002\)019<0037:AROWVA>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0037:AROWVA>2.0.CO;2).
- van Laar, T. W., V. Schemann, and R. A. J. Neggers, 2019: Investigating the diurnal evolution of the cloud size distribution of continental cumulus convection using multiday LES. *J. Atmos. Sci.*, **76**, 729–747, <https://doi.org/10.1175/JAS-D-18-0084.1>.
- Varble, A., and Coauthors, 2018: Cloud, aerosol, and complex terrain interactions (CACTI) science plan. DOE/SC-ARM-17-004, DOE Atmospheric Radiation Measurement Climate Research Facility, 57 pp., [www.arm.gov/publications/programdocs/doe-sc-arm-17-004.pdf](http://www.arm.gov/publications/programdocs/doe-sc-arm-17-004.pdf).
- Xiao, H., L. K. Berg, and M. Huang, 2018: The impact of surface heterogeneities and land-atmosphere interactions on shallow clouds over ARM SGP site. *J. Adv. Model. Earth Syst.*, **10**, 1220–1244, <https://doi.org/10.1029/2018MS001286>.
- Xie, S. C., R. T. Cederwall, and M. H. Zhang, 2004: Developing long-term single-column model/cloud system-resolving model forcing data using numerical weather prediction products constrained by surface and top of the atmosphere observations. *J. Geophys. Res.*, **109**, D01104, <https://doi.org/10.1029/2003JD004045>.
- Zhang, M. H., and J. L. Lin, 1997: Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *J. Atmos. Sci.*, **54**, 1503–1524, [https://doi.org/10.1175/1520-0469\(1997\)054<1503:CVAOSD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<1503:CVAOSD>2.0.CO;2).
- , J. L. Lin, R. T. Cederwall, J. J. Yio, and S. C. Xie, 2001: Objective analysis of ARM IOP data: Method and sensitivity. *Mon. Wea. Rev.*, **129**, 295–311, [https://doi.org/10.1175/1520-0493\(2001\)129<0295:OAOAID>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0295:OAOAID>2.0.CO;2).
- Zhang, Y. Y., and S. A. Klein, 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. *J. Atmos. Sci.*, **70**, 1297–1315, <https://doi.org/10.1175/JAS-D-12-0131.1>.