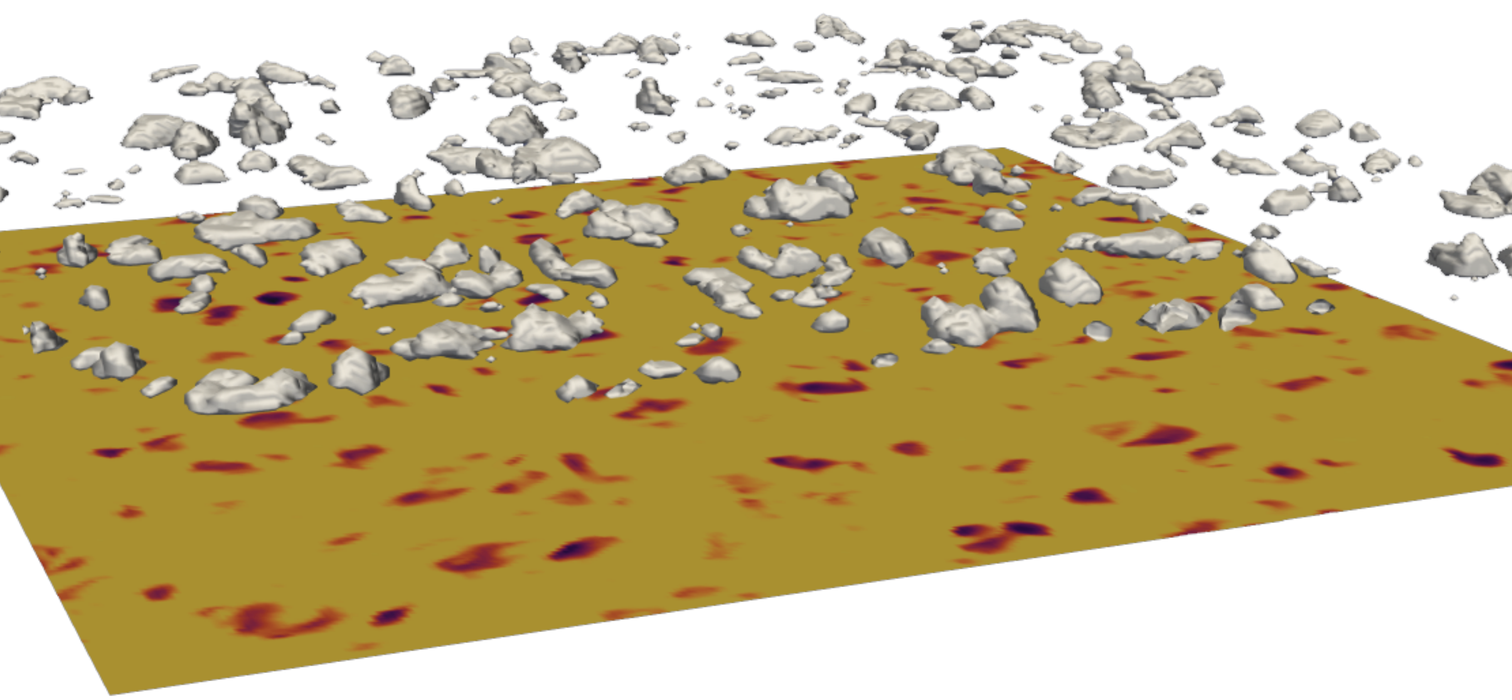
Shallow Convection in Convenient Packages

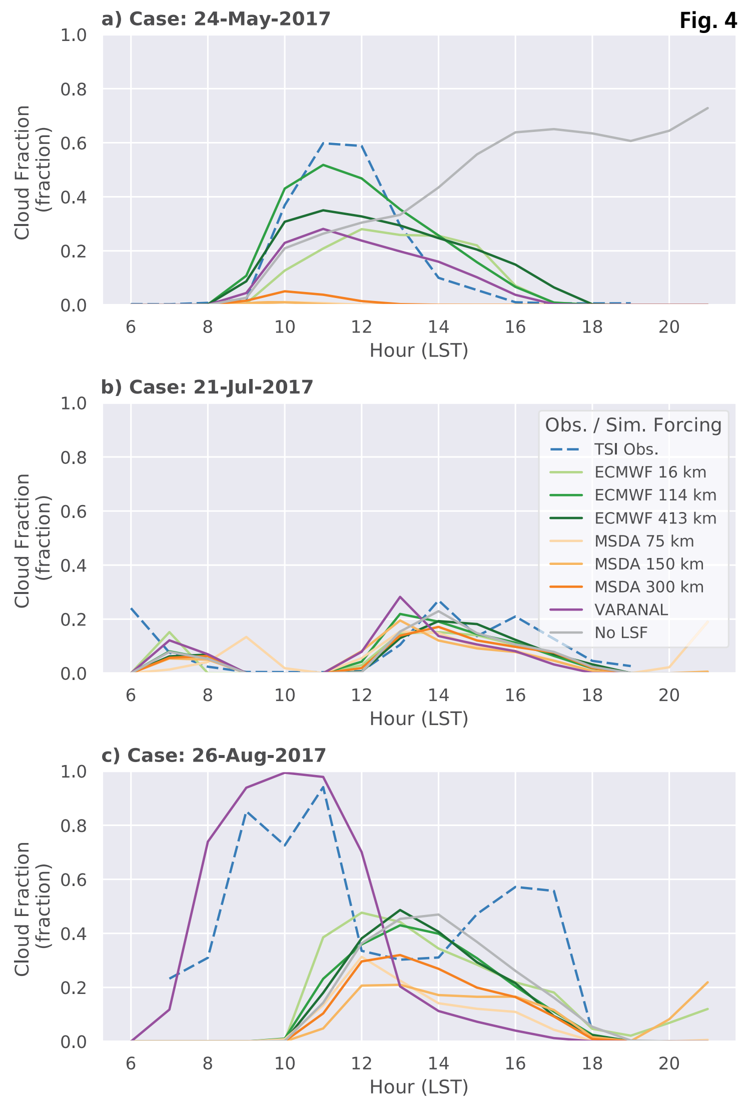
LASSO Bundles Observations and Simulations



Adapted from "The Large-Eddy Simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) Activity for Continental Shallow Convection," by **William I. Gustafson Jr.** (Pacific Northwest National Laboratory), **Andrew M. Vogelmann, Zhijin Li, Xiaoping Cheng, Kyle K. Dumas, Satoshi Endo, Karen L. Johnson, Bhargavi Krishna, Tami Fairless,** and **Heng Xiao**. Published online in *BAMS*, April 2020. For the full, citable article see DOI:10.1175/BAMS-D-19-0065.1.

Even the best-instrumented supersite cannot observe the atmosphere sufficiently to infer everything researchers want to know. In particular, researchers continue to struggle to connect localized observations to scales relevant for model parameterization development. For example, most observations are point based whereas models provide volumetric representations of the atmosphere. A synergistic combination of observations and modeling provides a holistic view of the atmospheric system. Observations provide real-world evaluations of the model, while modeling provides a dynamical context for localized measurements to bridge the gap to larger scales relevant to typical model parameterizations.

Large-Eddy Simulation (LES) ARM Symbiotic Simulation and Observation (LASSO) blends supersite observations with modeling. We have initially implemented LASSO to combine an LES with measurements from U.S. Department of Energy's



(DOE) Atmospheric Radiation Measurement's (ARM) Southern Great Plains (SGP) atmospheric observatory in Oklahoma for shallow convection conditions.

LASSO creates "data bundles" combining information including 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. Considerable effort has gone into facilitating the use of the data bundles, which are freely available to the community.

To date, three categories of users use data bundles in different ways: observationalists, modelers, and theoreticians. Early uses of LASSO LES output included proxies for cloud fields to improve cloud–radar scan strategies, evaluation of planetary boundary layer parameterizations in a single-column model, and theoretical investigation of aerosol–cloud interactions in shallow cumuli. LASSO forcing data is also used in test-bed mode with next-generation forecast models in development of parameterizations.

The broad target audience for LASSO presents a challenge to clearly communicate how to effectively use the information. For example, some retrievals only work in certain situations; observational uncertainty can be difficult to convey and often requires context. 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.

How LASSO is implemented for shallow convection is based on five core concepts. The first is that individual case studies are insufficient to gain robust understanding. LASSO relates a library of shallow convection days. Second, model inputs should not be fine-tuned, case-by-case, to make the model output match observations. For each shallow convection case date, we produce an ensemble of LES using a range of forcing scales and sources. Third, the data need to be easily usable. Fourth, users should be able to easily find and retrieve simulations. The data bundles are composed with care and provide skill scores for each simulation. The fifth and final core concept is ensuring that other modelers can reproduce the LES runs, as well as generate variants.

LASSO described

While LASSO produces an "operational" product, it differs from weather forecasting in that simulations are limited to days with shallow convection, typically April to September for the SGP facility. Selection of days is guided by an algorithm identifying potential for shallow convection as well as by availability of critical ARM observations. A final requirement for selection is the presence of somewhat homogeneous cloud fields in the surrounding region. We try to avoid days where the forcing is unlikely to be

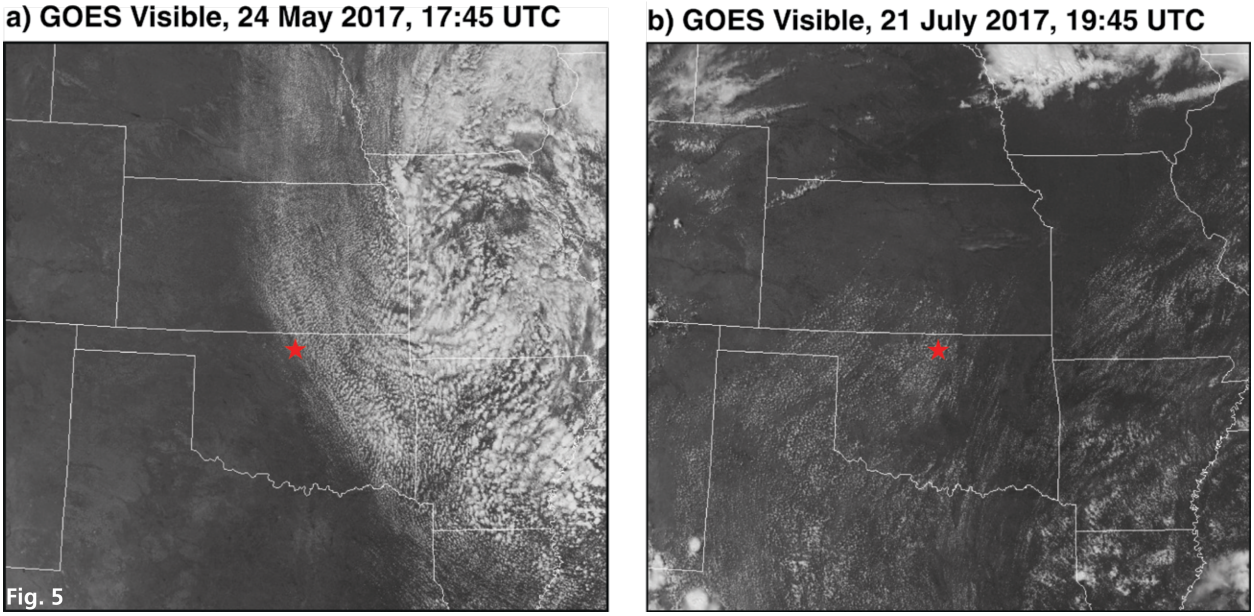
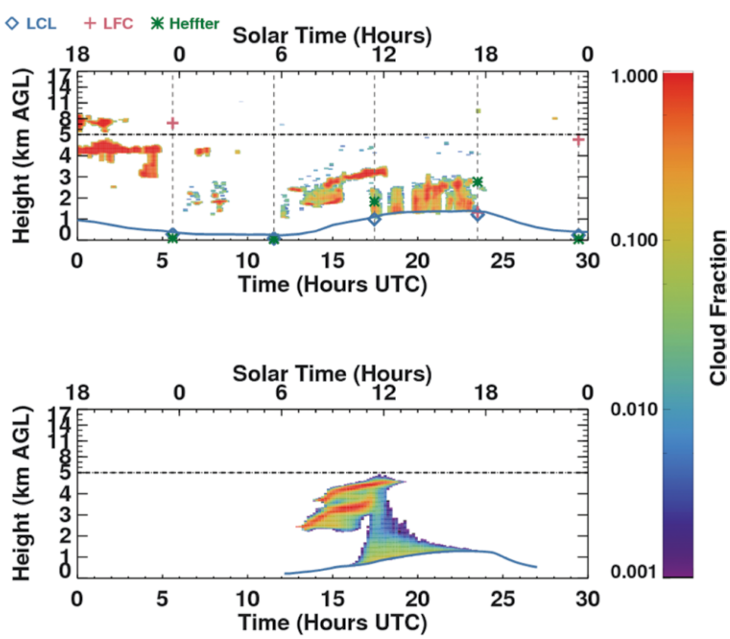


Fig. 6 Cloud Fraction from ARSCL vs. LES Forced by Variational Analysis, 26 August 2017



uniform across the simulated region with a goal of avoiding muddled forcing data that might be inconsistent with the localized observations.

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 eight-member ensemble of large-scale forcing datasets, using

Within LASSO, complementary observations combine 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 observed cloud fraction is included from the total-sky imager (TSI, based on hemispheric images), alongside shallow-cloud fraction from ARM's Active Remote Sensing of Clouds (ARSCL) Value-Added Product, 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. Fig. 4: Comparison of cloud fraction from the TSI (dashed blue) with the LES ensemble for 3 days from 2017. Fig. 5: Cropped GOES visible images, from aviationweather.gov, also available from the LASSO Bundle Browser. The SGP Central Facility is indicated by the red stars. Fig. 6 is an example of comparison plots from the data bundles: (top) The ARSCL-derived cloud fraction 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 sampled in time, whereas the LES cloud fraction (bottom) is an instantaneous domain average. Note above 5 km AGL the vertical scale is compressed. The blue lines in each panel indicate the lifting condensation level. In the top panel, red "plus" signs indicate the level of free convection and green stars indicate the boundary layer height computed from sonde data using the Heffter algorithm.

a range of forcing scales and sources. The large-scale forcings come from three sources: two based on numerical weather prediction and a third using a forecast product as a background field that is optimally adjusted to be consistent with observations. 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, whereas with LASSO, the ensemble is produced after reality has happened: 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 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. Finally, one ensemble member uses only the LES initial conditions taken from the radiosonde at 0600 local standard time, combined with surface fluxes; the tendencies from the large-scale forcings are zero for this member.

LASSO uses observed surface fluxes for the LES lower boundary condition, which simplifies the model initialization but precludes physical processes related to land–atmosphere interactions. The operational configuration uses the Weather Research and Forecasting (WRF) Model with a 25-km domain of 100-m horizontal grid spacing and 30-m vertical spacing up to 5-km height stretching to 300 m near the model top at around 15 km.

An important challenge has been putting together an optimal suite of observations for model evaluation. A balance is needed between the observations used for model evaluation and observations used to constrain the model initial conditions and forcings.

To better capture spatial variability, ARM established four facilities approximately 50 km from the SGP Central Facility. Each includes surface meteorological instrumentation, a Doppler lidar (for cloud-base heights and eventually boundary layer vertical velocity variance), a three-channel microwave radiometer, and at three of the stations an atmospheric emitted radiance interferometer to obtain regional estimates of boundary layer temperature and retrievals of liquid water path.

Data and general LES behavior

The production of LASSO data bundles has run from the summer of 2015 to the present, with 78 case days through 2018 available for download. In total, 1,172 data bundles are available so far, with the current approach producing 8 ensemble members per case date.

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. The variability within a given day's ensemble can differ noticeably. In a perfect case, all ensemble members would converge on the observed conditions. However, commonly one or more members deviate. For example, on 24 May 2017 most simulations followed the typical diurnal cycle for shallow cumuli, yet this did not happen in reality. Interestingly, the simulation using no large-scale forcing differed substantially by having an increasing cloud fraction almost the entire day. This was an example of how important the large-scale forcing can be throughout the day.

Future of LASSO

At this point, the model configuration and core set of observations have been established, and we expect the data bundle format to be stable for the shallow convection scenario. In addition to the available cases, the 2019 season is being processed, and additional shallow convection seasons may be run in the future to expand the library of available cases.

The LASSO concept is meant to be adaptable, and additional scenarios will be implemented for other sites and phenomena. Now that the shallow convection modeling process is operational, LASSO will begin adding value to other ARM observations, such as for the Cloud, Aerosol, and Complex Terrain Interactions (CACTI) field campaign studying deep convection in Argentina.

The ultimate utility of the LASSO framework will be determined by how researchers use it. We have been encouraged by the creativity of early adopters and look forward to seeing how uses evolve. We are also excited to see additional related interests in the starting of new observation–model endeavors.

BAMS: What would you like readers to learn from this article?

Karen Johnson (Brookhaven National Laboratory): *Historically, meteorological modelers and observationalists have not interacted a great deal. LASSO brings long-term observations and Large Eddy Simulation modeling together to bridge the gap between point observations and the scales needed to inform large-scale models.*

William Gustafson (Pacific Northwest National Laboratory): *The DOE ARM program has such a wealth of measurements, and at the same time, DOE is developing a new and improved climate model. LASSO is one way to help marry the two together to add value for researchers working with both sets of data. It attempts to get past limitations sometimes imposed by certain scales.*

Andrew Vogelmann (Brookhaven National Laboratory): *Heavily instrumented supersites routinely acquire observations at scales that match the grid spacing of large-eddy simulation (LES) models. With the maturity of LES modeling and the computational power now available, large ensembles of simulations are possible to use this model-observation synergy to build a statistically meaningful understanding of atmospheric processes.*

BAMS: What is a distinctive characteristic of the LASSO approach to teaming observations with models?

Zhijin Li (JPL, CalTech, and UCLA): *Even the best-instrumented supersite cannot observe the atmosphere sufficiently to infer everything researchers want to know. The combination of LES and observations produces a "data bundle" that represents many more temporal and spatial scales and variables than observations can provide. The data bundles are a new type of "almost-complete observations." They are imperfect, but their uncertainties are provided.*

BAMS: What are some of the specific problems addressed by this bundling?

ZL: *They can have a range of applications, for example, testing instrument algorithms, evaluating large-scale model parametrization schemes, multiscale interaction dynamics, etc.*

AV: *Accurate representation of cloud processes in NWP and climate models has stubbornly remained a leading challenge requiring us to "up our game" through the development of new approaches such as this.*

WG: *Marrying the LES more tightly with observations seemed like a great way to help the atmospheric community move forward and make progress improving the models.*

BAMS: What research topics led you to this project?

AV: *After I became interested in the challenges of accurately remotely sensing shallow convection cloud properties, I led an aircraft-borne cloud sampling campaign to acquire in situ cloud and environmental properties to evaluate retrievals and understand the fascinating physical processes involved. We then faced the challenge of modeling case studies from the campaign, which set the stage for the next-generation step of the routine observation-evaluated simulations described in this paper.*

ZL: *My research has been focused on integrating observations into a small-scale resolving model during the past two decades, using data assimilation.*

BAMS: What surprised you the most about what you have learned so far from LASSO?

ZL: *In some cases LES performed surprisingly well, while others were surprisingly bad. We learned a great deal through those cases. Personally, I am very surprised that so many different types of recourses converge for LASSO so that we can rely on them to succeed. I am just humbled.*

WG: *This article is the result of years combining input from a large number of people with many different specialties. Through leading the LASSO project I have had to learn to speak many different "languages" to understand all the details and make a coherent data package.*

BAMS: What have been the biggest challenges you encountered in LASSO?

ZL: *LES represents a wide range of scales. The challenge arises because observations are very inappropriate to constrain all scales. The conventional data assimilation schemes used by meteorological centers do not work well for LES.*

WG: *We have been trying to put together a standardized data bundle that would make it easy for researchers to compare simulations from different cases spanning years. However, instrumentation changes from year to year, which means we continually have to adapt. Sometimes this presents itself as a new opportunity, such as a new photogrammetric cloud fraction product. Other times, instruments malfunction or are replaced. A switch from a two-channel to a three-channel microwave radiometer in theory could offer improved results, but in reality led to years of calibration issues.*

AV: *The challenge of bringing together high-resolution model simulations and observations in this manner is not for the faint of heart. Conceptually, it looks straightforward on a flow diagram, but implementation is a whole new beast. This work would not have been possible without the dedicated network of instrumental, operational, and computational groups from ARM. It takes a village.*

WG: *I have come to really appreciate the help from so many people who make LASSO happen—all of them critical.*