LASSO Breakout Session:
CACTI Deep Convection Scenario

William I. Gustafson Jr. (PNNL), Andrew M. Vogelmann (BNL), & James H. Mather (PNNL)

Via Zoom, 24 June 2020

Zoom “Webinar” Instructions (vs. “Meetings”)

- Invite links are personal to you
  - Do not share your invitation link (this can cause issues with shared audio)
- Please identify yourself by full name for your login (you can rename your posted name)
- Session attendees will be muted except when called on by the host; video will also be off
- Questions and comments can be typed into the Q&A feature at any time & are the preferred method of initiating discussion
- Virtually “raise your hand” if you would like to have a longer interaction and we will call on you
- Those using a phone to call into the session
  - *9 will raise your hand
  - *6 will mute & unmute your line once the host allows it
- This session is being recorded for later viewing

Agenda

- Annual status update (Bill Gustafson)
- Overview of the CACTI campaign (Adam Varble)
- Motivation and goals for the LASSO CACTI scenario (Bill Gustafson)
- Observational data and skill scores for the CACTI scenario (Andy Vogelmann)
- Current thoughts on the framework and model setup (Bill Gustafson)
- Open discussion

LASSO Update for Shallow Convection

William I. Gustafson Jr. & Heng Xiao (PNNL)
Andrew M. Vogelmann, Satoshi Endo, Tami Fairless, Karen Johnson (BNL)
Zhijin Li (JPL & UCLA)
Kyle Dumas, Michael Giansiracusa (ORNL)

LASSO article in BAMS

- Please cite our new BAMS article when referring to LASSO

Gustafson, W. I., A. M. Vogelmann, Z. Li, X. Cheng, K. K. Dumas, S. Endo, K. L. Johnson, B. Krishna, T. Fairless, and H. Xiao, 2020: The Large-Eddy Simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) activity for continental shallow convection. *Bull. Amer. Meteor.*, 101, E462–E479, https://doi.org/10.1175/BAMS-D-19-0065.1.

2019 LASSO Shallow Cumulus Cases

The highlights…

- 17 cases from 2019 season to be released this summer
- Changed the ECMWF forcing methodology
- Including high-frequency observation data for each case
- COGS photogrammetry-based cloud fraction added as a side product
- Reprocessing
- Putting ShCu production on hiatus to focus on additional scenarios

Year	Number of Cases	Release Date
2015	5	July 2018
2016	13	Sept. 2017
2017	30	Sept. 2018
2018	30	Sept. 2019
2019	17	Anticipated Jul./Aug. 2020
Total	95	

ECMWF Large-Scale Forcing Changes

- Data availability required us to approach ECMWF forcing differently for 2019 cases

- Evaluated two approaches for obtaining these large-scale forcings
  1. DDH forcing available via ARM that is based on the same regions as we used before
  2. Area-averaged large-scale forcing derived from ERA5 on the raw model levels

- Similar results between the two, so chose to use the DDH profiles already available from ARM

New High-Frequency Observations File

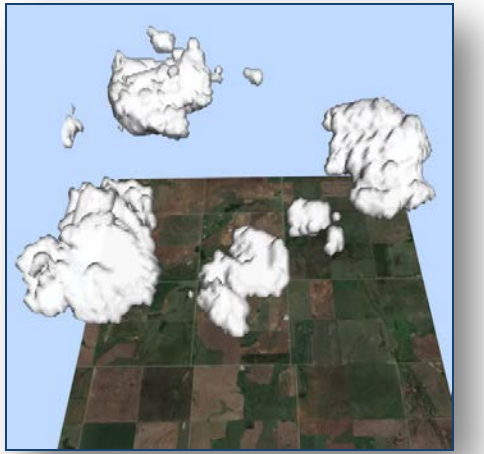
- Observations available within the data bundles are at the hourly intervals used for the skill scores
- Now, observations are available at native or near-native time resolution for LASSO days
- High-frequency observations downloadable via an optional tar file
  - About 10 MB compressed, uncompresses to about 145 MB

Obs. Type	Frequency
RWP winds	10 min.
LCL (multi site)	1 min.
Cloud fractions (from ARSCL & TSI)	1, 5, & 15 min.

Obs. Type	Frequency
Bdy-layer thermodynamics (from AERI + Raman lidar)	10 min.
Cloud-base height (from Doppler lidars)	10 min.
LWP (from AERIoe=AERI + MWR-2C)	10 s

Clouds Optically Gridded by Stereo (COGS) Cloud Fraction

- The COGS cloud fraction product
  - Based on photogrammetry with three cameras spaced in a 12-km diameter ring (Romps and Öktem, 2018, BAMS)
  - Best for cloud fractions <0.6
  - Available for 2018 onward at SGP



- COGS substantially improves cloud mask estimates over ARSCL
  - Not impacted by insects
  - Samples a more representative volume
  - Universally measures less cloud thickness than ARSCL

ARSCL vs. COGS Cloud Fraction Comparison

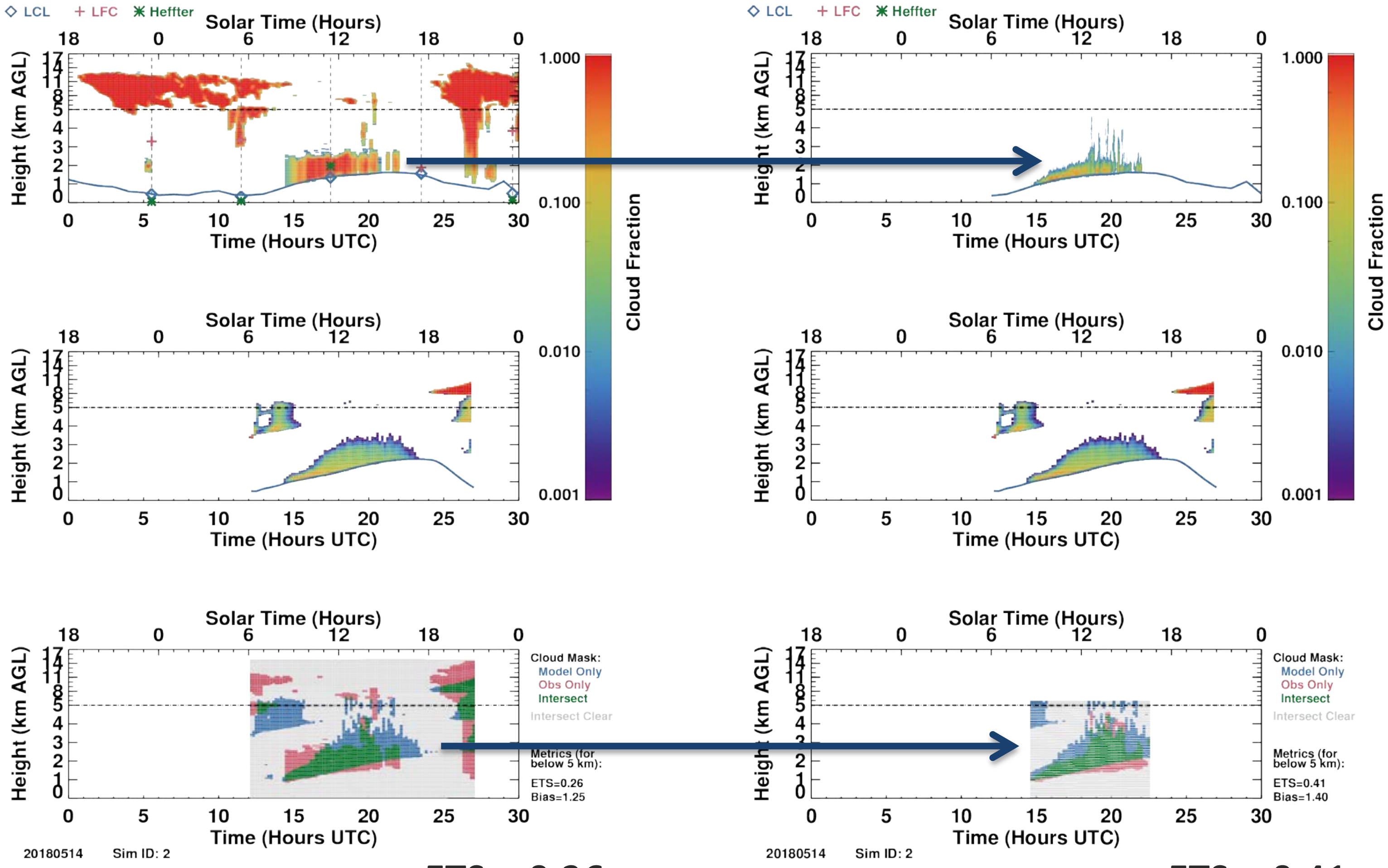
14-May-2018 at SGP, Simulation Forced by VARANAL

Obs. from KAZRARSCL | Obs. from COGS

Observed Cloud Fraction

LES Cloud Fraction

Observations vs. LES Mask



ETS = 0.26

ETS = 0.41

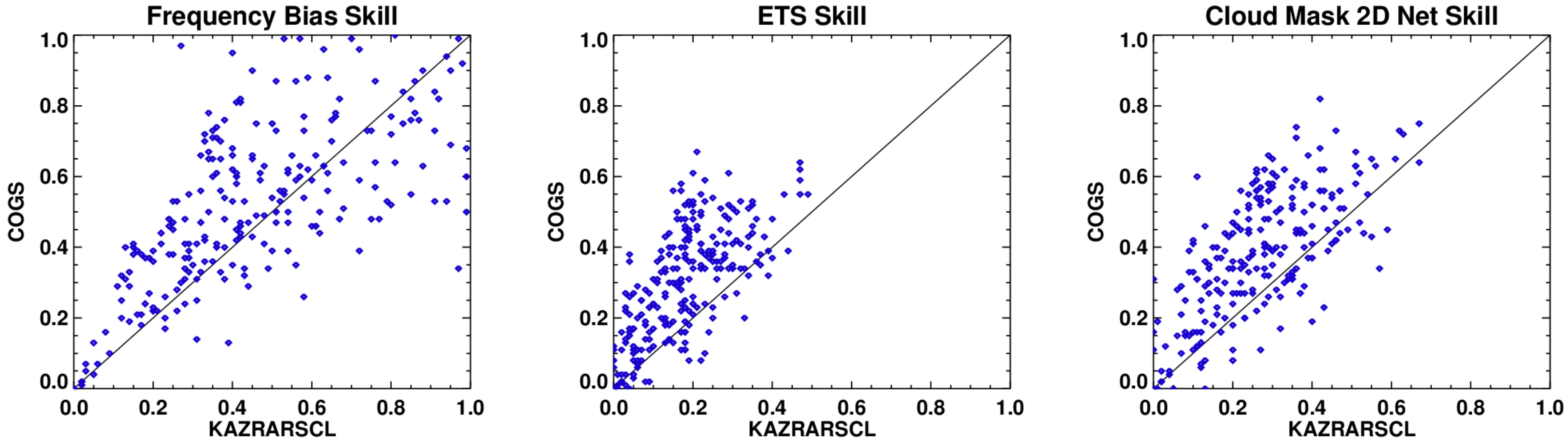
Reduced insect clutter

LES matches better

Releasing a LASSO-COGS Evaluation Product

- LASSO LES verifies much better against COGS than KAZRRARSCL

LASSO Time-Height Cloud-Mask Metrics Using COGS vs. KAZRARSCL, 2018 Cases



- Releasing an evaluation product for LASSO 2018 & 2019 days that includes:
  - Hourly COGS data for cloud-top heights <6 km
  - LASSO skill scores vs. COGS

Reprocessing of Previously Released Cases

- Bugs have been found in both the VARANAL and ECMWF data used for LES forcings
  - 2018 VARANAL forcing is highly impacted, particularly in May, plus a 2015 case; issue is related to incorrectly handling missing data when generating VARANAL
  - 2015–2018 ECMWF forcing had an incorrect multiplier for the handling of vertical velocity
    - Before-and-after comparisons show impact on LES is not systematic and fixed runs are not necessarily better
    - Quick fix for ECMWF forcing users: $w_{fixed} = w_{original}*10$

- Will be reprocessing impacted cases and releasing them later this year
  - Will take this opportunity to include the high-freq. obs. file in the 2018 cases

Looking Toward LASSO's Future

- Putting shallow-convection scenario into a hiatus after 2019 cases
  - The COVID shutdown cost 1–2 months of data for 2020
  - The library of 97 cases has reached a critical mass for research
  - Would like to free up budget to broaden LASSO to other regimes

- Plan to complete CACTI simulations in 2021
  - Expect to have the observation processing done by end of 2020
  - LES will likely need to extend into 2021—it will be slow
- Budgeting to start working on an ACE-ENA scenario in FY21
  - Primary focus on maritime-cloud precipitation processes
  - Open to suggestions, but will not be putting a lot of thought into this until after we get CACTI further along in development
  - Seeking a good methodology for CCN(SS) profiles from observations

Background on the CACTI Field Campaign

Adam Varble, PNNL

Slides for this section available in a separate file.

Motivations and Goals of the New LASSO CACTI Scenario

William I. Gustafson Jr., PNNL

Approaching the CACTI Scenario

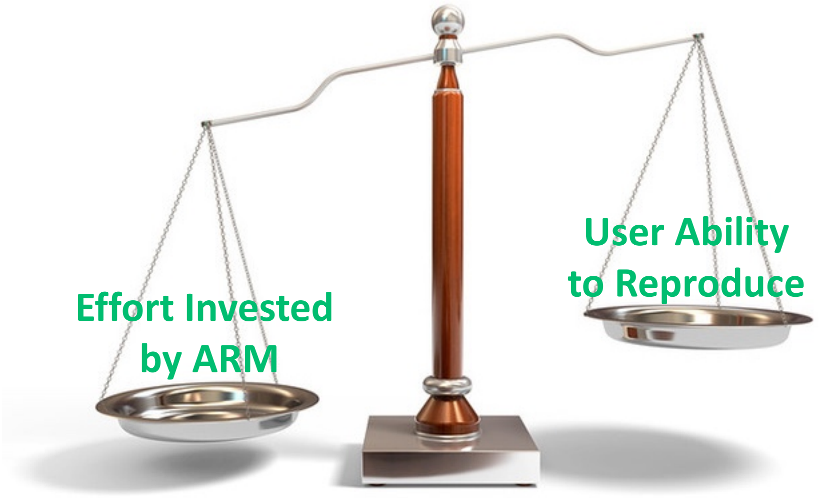
The highlights…

- LASSO is geared toward adding value to ARM's observations
- Deep convection is much more variable than diurnal, surface-driven shallow convection
- Deep convection requires a larger domain than shallow convection
- General storm features are visible in km-scale simulations, but many details require LES
- Overall approach follows thinking from the LASSO Expansion Workshop report

https://www.arm.gov/publications/programdocs/doe-sc-arm-19-023.pdf

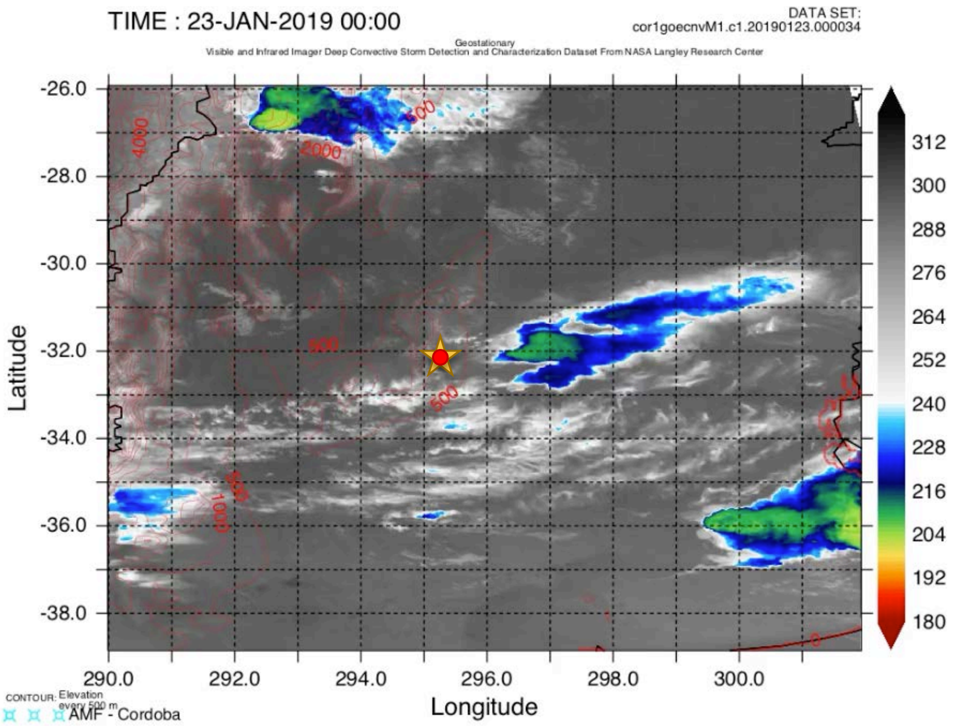
Adding Value to ARM's Observations

- ARM is primarily an observation-based facility
- LASSO modeling is meant to enhance users' abilities to use ARM's observations
- LASSO is not intended to provide research "tied in a bow," but we do want to make it easier for researchers to do impactful work
- Designing for deep convection changes the balance of what we will do since the runs are so much more expensive

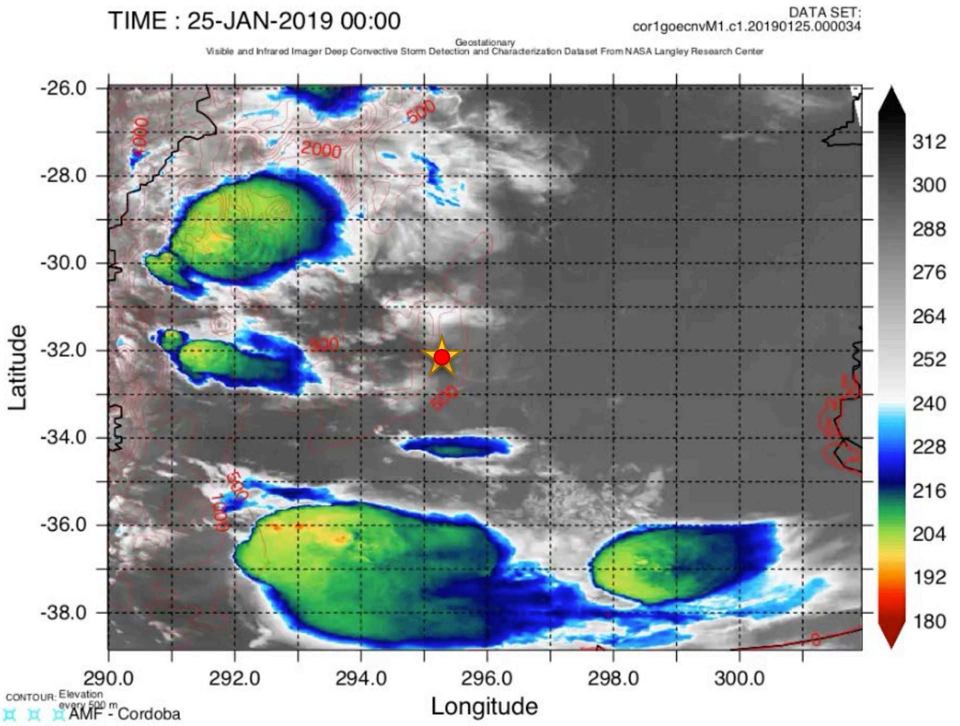


Deep Convection is More Variable (and Bigger) than Shallow

- Deep convection more easily takes different forms
- A statistically representative library is not feasible—argues for a targeted approach



IR Brightness Temperature Image (degrees_Kelvin)

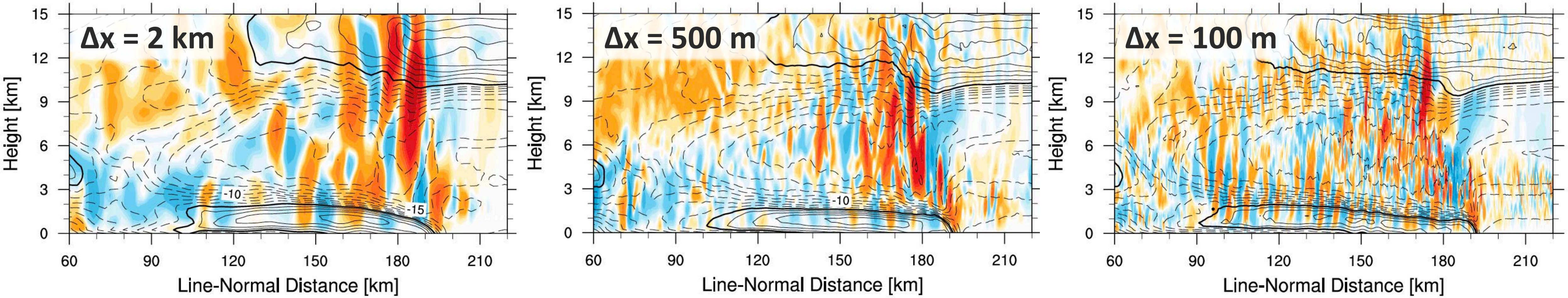


IR Brightness Temperature Image (degrees_Kelvin)

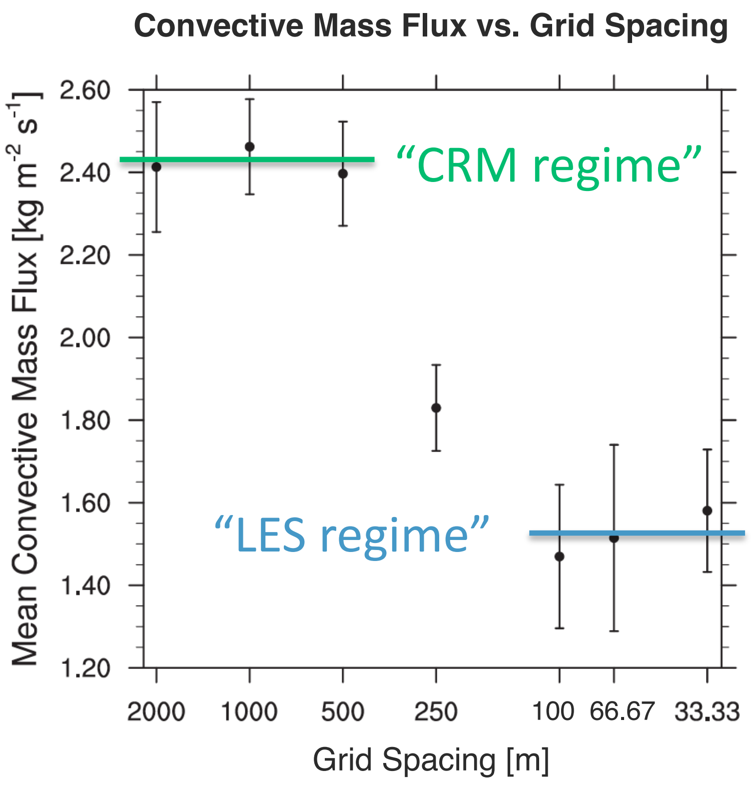
Science Drivers Dictate Resolution Choice

- Kilometer-scale simulations are "off-the-shelf" for many parts of the globe, but there is still value in them, particularly as an ensemble
- Many deep convective signatures are evident with km-scale grid spacing, but high-priority science requires seeing the bubbles

Vertical (colors) & Horizontal (contours) Velocities



Many convective statistics start to converge for Δx ≲ 200 m



- At least for this idealized storm simulation, around 100 m grid spacing is the sweet spot
- Still investigating cost implications versus available computing resources

Figure from Lebo and Morrison (2015, MWR)

Big-Picture Science Drivers from the Expansion Workshop

- Convective cloud dynamics (e.g., thermal-like structures, updraft strength, and entrainment) and the relationship to critical features like updraft and downdraft mass fluxes, vertical transport, and the shallow-to-deep convective transition
- Cold pool interactions with the surrounding environment and convective drafts in turbulent flow
- Microphysics-dynamics interactions, especially in the context of cloud-scale eddies and smaller-scale turbulence

Prioritizing Science Drivers

- Life cycle and behavior of initiating and initially isolated deep convective cells
  - Convective initiation, e.g., what determines timing, scale, and location of initiation?
  - Early upscale growth, e.g., how do internal storm dynamics vs. the environment control growth?

- Implications of science drivers on configuration choices
  - Both initiation & growth require resolving thermals within convective cores
  - Both require frequent sampling to identify evolution of statistics and motion

Questions…

- What science priorities would you rank highest?
- How would you use CACTI simulations generated by ARM?

Observational Data and Skill Scores for the CACTI Scenario

Andrew M. Vogelmann, BNL

Slides for this section available in a separate file.

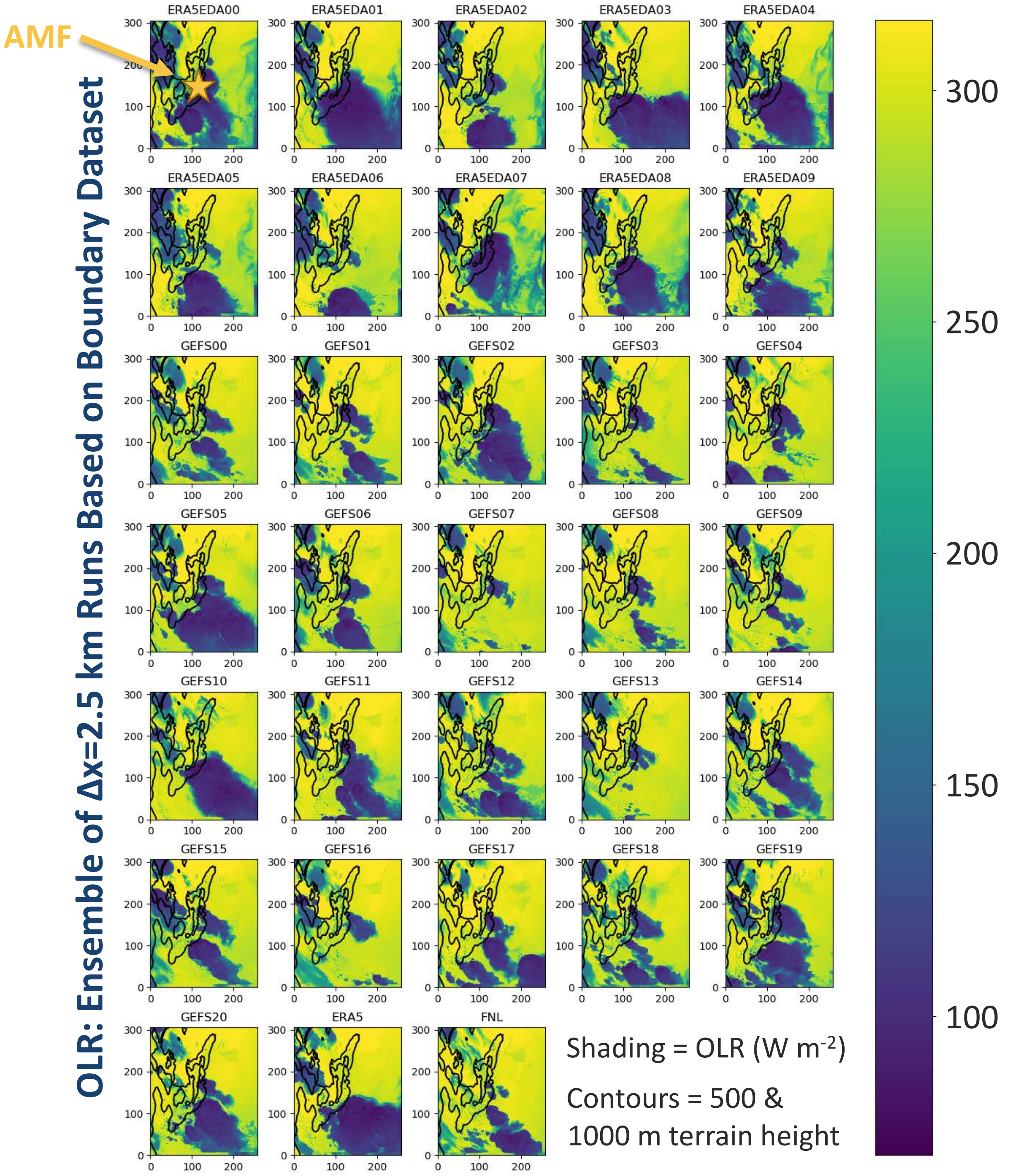
Current Thinking for the CACTI Scenario Framework and Model Setup

William I. Gustafson Jr., PNNL

Ensembles

- We used ensembles of forcings to generate ensembles of LES for the shallow convection
- For deep convection we will use boundary-condition ensembles for mesoscale runs to identify a small subset to use for the LES
- Envision km-scale ensembles to be useful for understanding dynamical sensitivities and GCM comparisons

- Ensemble example at right showing OLR
  - 25-Jan-2019 20 UTC, Δx = 2.5 km
  - Different boundary condition dataset for each ensemble member
    - ERA5-EDA for 10 members
    - GEFS for 21 members
    - ERA5
    - FNL
    - (Considering JRA-55 if somebody has a Vtable for it)

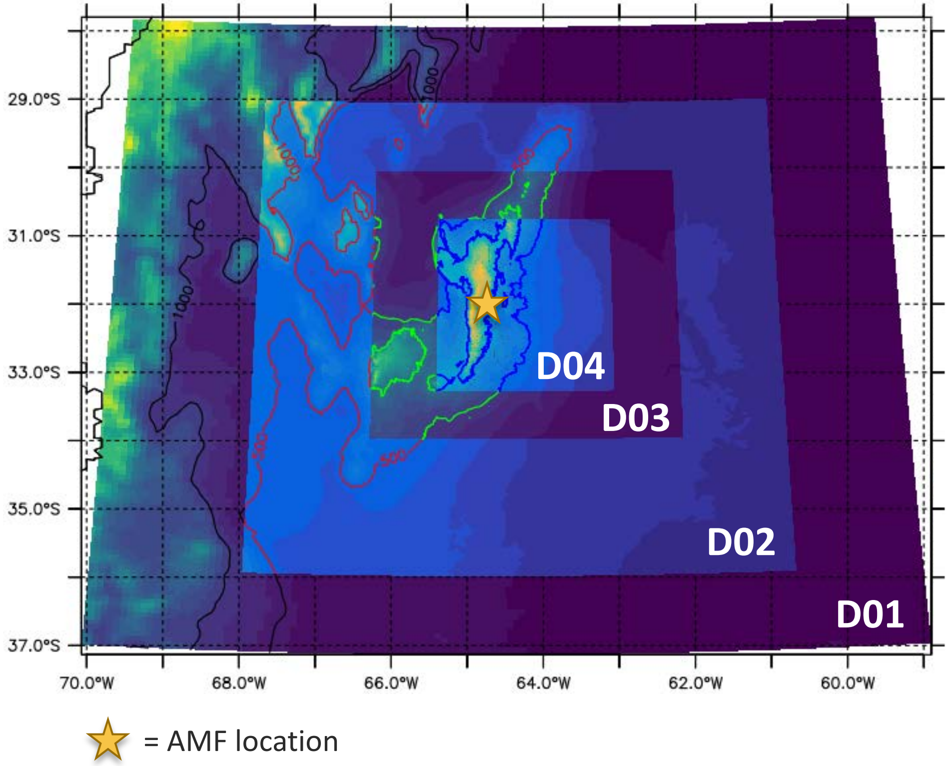


What might the CACTI LES look like?

- Leaning toward a hero-run configuration:
  big domain + high resolution + frequent output
- Nature runs using nested domains
- Inner domain location and size might differ between cases

- Example domain setup at right:
  - 4 grids from 7.5 km down to 100 m
  - D03, Δx = 500 m
    - 751 x 866 grid cells = 375.5 x 433 km^2
  - D04, Δx = 100 m
    - 2146 x 2776 grid cells = 214.6 x 277.6 km^2

Δx = 7.5 km, 2.5 km, 500 m, & 100 m



Number of LES Per Case is TBD

- Need to balance cost vs. simulation count
  - What is more valuable?
    - Multiple simulations per case with fewer cases
    - One or two simulations per case with twice as many cases
  - ARM is buying more computing power later this year, ~8k new cores to add to existing 4k

- Cost of example domain
  - Current ShCu case: wall time = 21.5 h on 500 cores per simulation
  - Estimated cost of 214.6 x 277.6 km^2 domain: wall time ≅ 2–3 weeks on 4000 cores per simulation
    - Assumes we integrate LES for 12 h, dx=100 m, 180 levels
    - Implies we could do max of ~10 cases per year with 12k cores and 2–3 LES per case

Model Physics

- No convection parameterization for dx < 5 km
- Currently comparing P3-2ice with Thompson
  - Is it worth releasing LES for both?
- Interactive Noah soil model initialized from reanalysis
- RRTMG radiation
  - What sort of detailed radiation output would be required?

Raw Model Output

Rough File Sizes for Each Domain

Δx =	7.5 km	2.5 km	500 m	100 m
N_x	130	258	750	2145
N_y	136	306	865	2775
Snapshot Size	0.5 GB	1.9 GB	15 GB	140 GB

- Outer, km-scale domains
  - Snapshots every 10 min. (Δx=2.5 km: 24 h * 6 outputs/h * 1.9 GB/output = 274 GB)
- LES domain(s)
  - Are users interested in the Δx=500 m domain?
  - Snapshots every 5 min. during the entire simulation (12 h * 12 outputs/h * 140 GB/output = 20 TB)
  - Snapshots every 1 min. during periods of interest (4 h * 60 outputs/h * 140 GB/output = 33 TB)
  - Should ARM provide sub-1 min. output for a subset of fields, e.g., 10 s?
- What variables are desired beyond the standard met. fields?
  - Microphysical process rates
  - Radiative heating profiles
  - Radar reflectivity
- Restart files once per hour (maybe more frequently depending on timing)
- Does this quantity of data cause you angst?

(10 cases * 2 LES/case * 46 TB/LES = 0.9 PB w/o sub-minute subsets, restarts, and outer domains)

Post-Processed Model Output

- Values used to compute skill scores will be provided
- Many fields can be easily computed from model output—should ARM leave this to the user? Suggestions for what ARM should provide?
  - LCL, CIN, CAPE?
  - What sort of “statistics” file should we provide?

- Should ARM make available model output by variable?
- Should ARM make it possible to select the snapshot frequency for download?

Open Discussion

Raise your hand to speak and we will call on you.
Comments and questions can also be posted via the Q&A feature.

Feel free to contact Bill and Andy at lasso@arm.gov at any time as well.

CACTI Scenario Discussion Questions

- Are there additional science drivers we should consider? Particularly, those that would change our approach?
- What science priorities would you rank highest?
- What would you need to use the km-scale simulations and LES generated by ARM?
- How would you weight the balance of ensemble members (forcing and physics) vs. number of cases?
- Value of “intermediary” grid spacings? Would you use them?
- How long should sub-minute output be done and at what frequency?
- Any special variables that should be output?
- Where would you work with this output? How would you deal with 30-50+ TB?
- How would you like ARM to provide download options?

Additional feedback welcomed at lasso@arm.gov