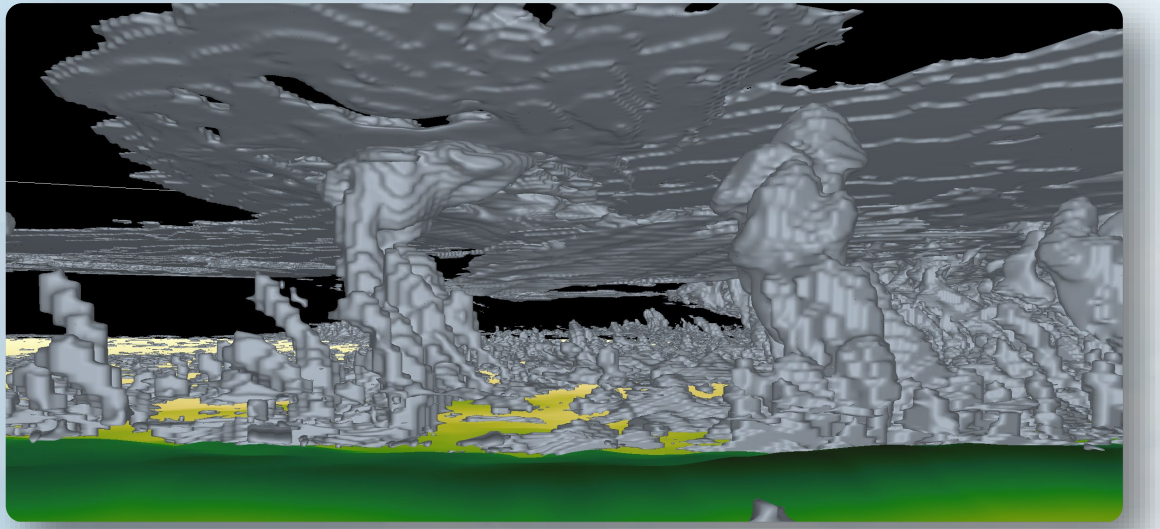
LASSO-CACTI Scenario for Deep-Convection with Large-Eddy Simulation



25 October 2022

William I. Gustafson Jr.[1], Andrew M. Vogelmann[2], Mark M. Delgado[2], Satoshi Endo[2], Eddie K. Schumann[1], Adam C. Varble[1], & Heng Xiao[1]

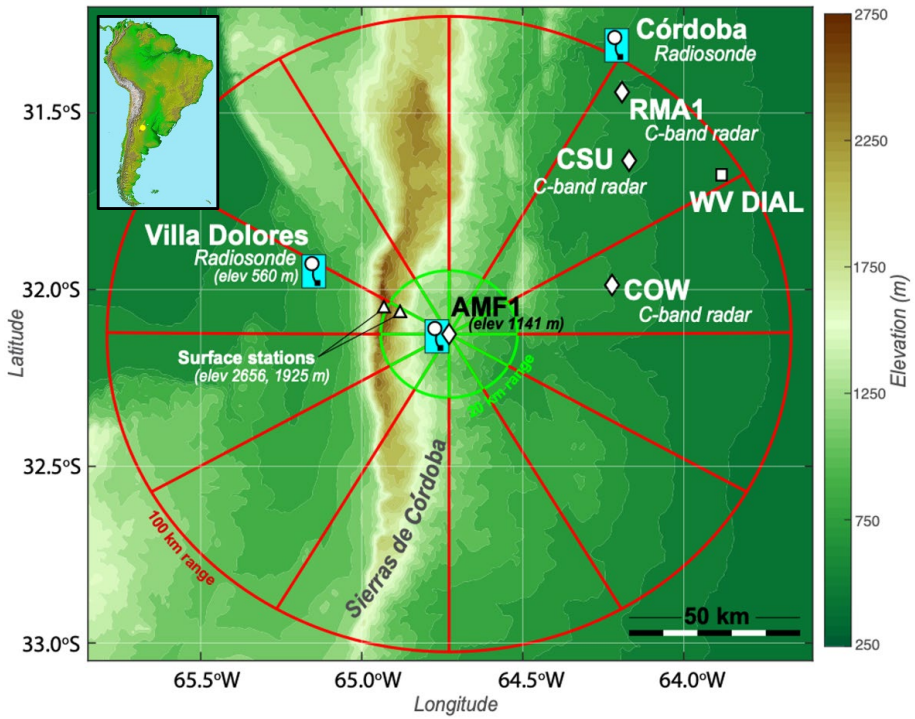
[1] Pacific Northwest National Laboratory, [2] Brookhaven National Laboratory

https://www.arm.gov/capabilities/modeling/lasso
https://discourse.arm.gov/c/lasso/

What is LASSO-CACTI?

- The CACTI field campaign occurred in 2018–2019 in Argentina with a focus on convective clouds and their transitions such as upscale growth
- LASSO = LES ARM Symbiotic Simulation and Observation
- LASSO adds value to ARM observations by using libraries of high-resolution modeling to bridge scale gaps and add context to observations
- The data is starting to flow…
  - Beta release made available May 2022
  - Full release coming soon in early 2023

Map of CACTI Deployment in Argentina



Science drivers guide scenario design

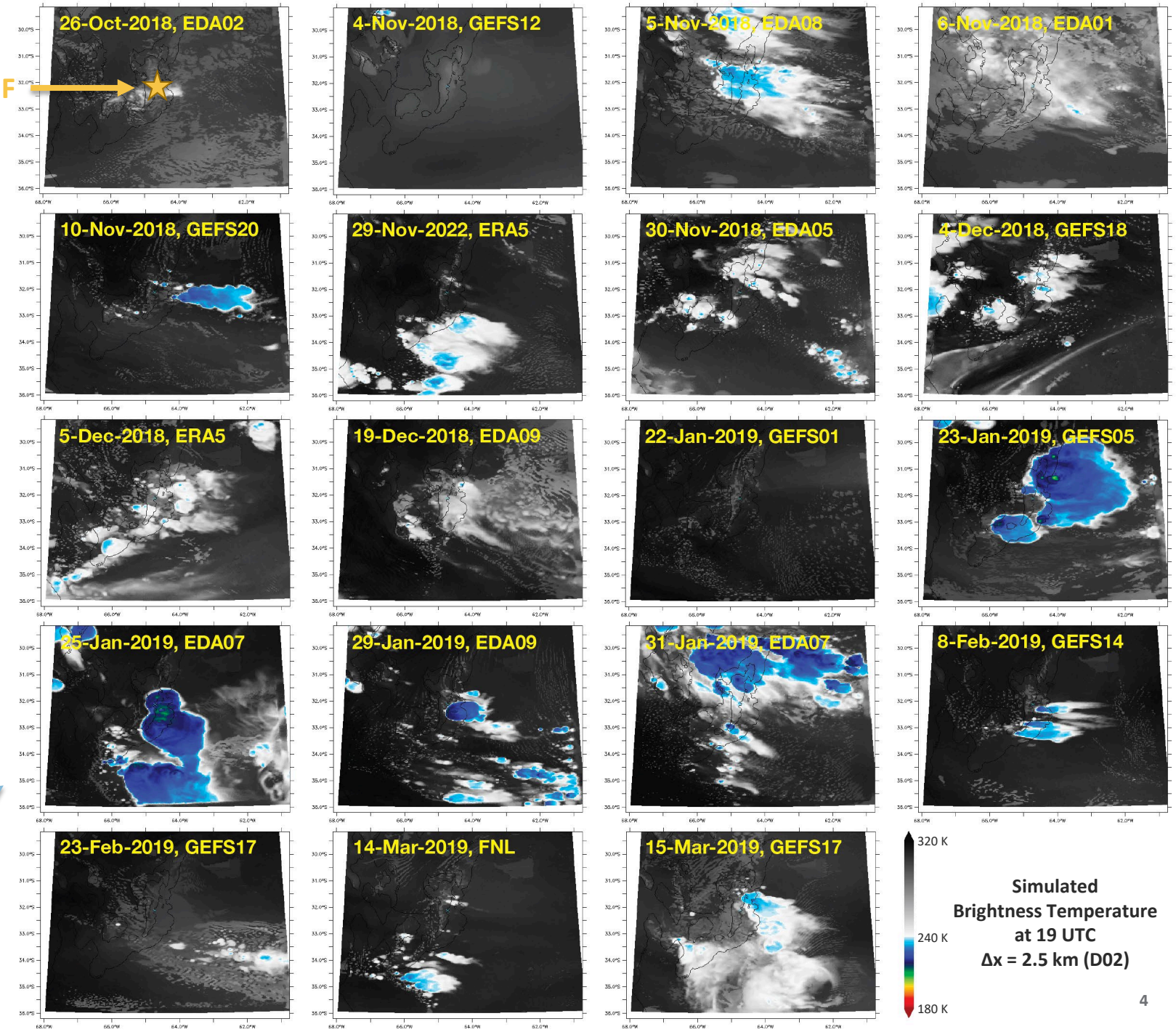
- Convective cloud dynamics
  - e.g., thermal-like structures, updraft strength, and entrainment; the relationship to critical features like updraft and downdraft mass fluxes, vertical transport, and the shallow-to-deep convective transition
  - Convection-environment interactions, e.g., cold pools
  - Convective drafts in turbulent flow
- Microphysics-dynamics interactions
  - Especially in the context of cloud-scale eddies and smaller-scale turbulence

- Science drivers chosen to balance relevant science with computational capacity
  - LES resolution governed by cloud core requirements
  - Domain size determines portion of lifespan simulated
  - Limiting ensembles to mesoscale simulations with the potential for a small number of LES ensemble members for specific cases

Case dates target a selection of convective behavior

- Chosen days have convection form and grow within view of ARM's scanning radar
- Identified 20 case dates meeting our criteria
  - Convection ranges from short-lived convection to large MCSs
  - See Vogelmann et al. poster, #66, for list of dates

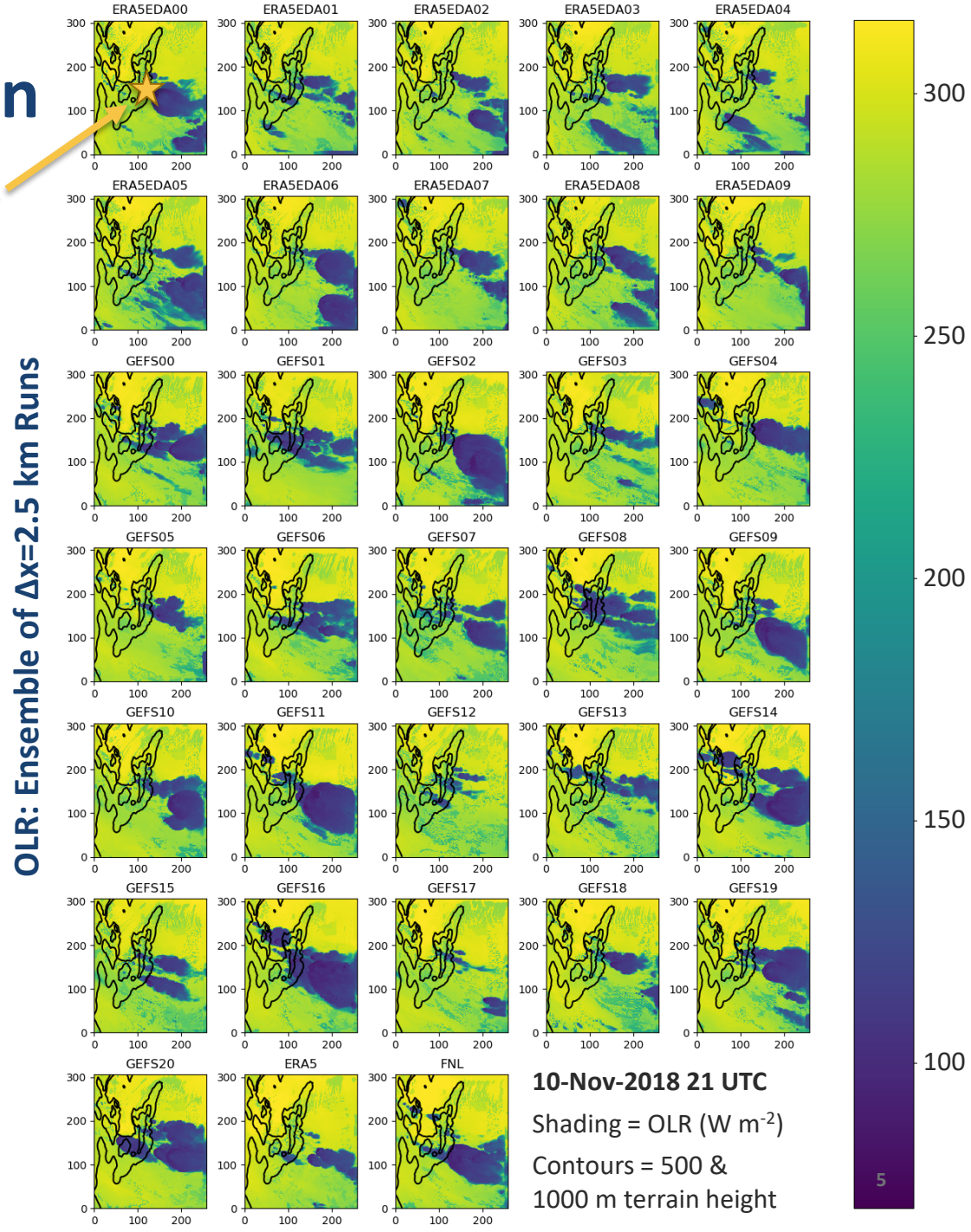
Some of the mesoscale simulations for different case dates, all plotted at the same time of day, 19 UTC



Mesoscale ensembles for case selection and LES boundary condition choices

AMF

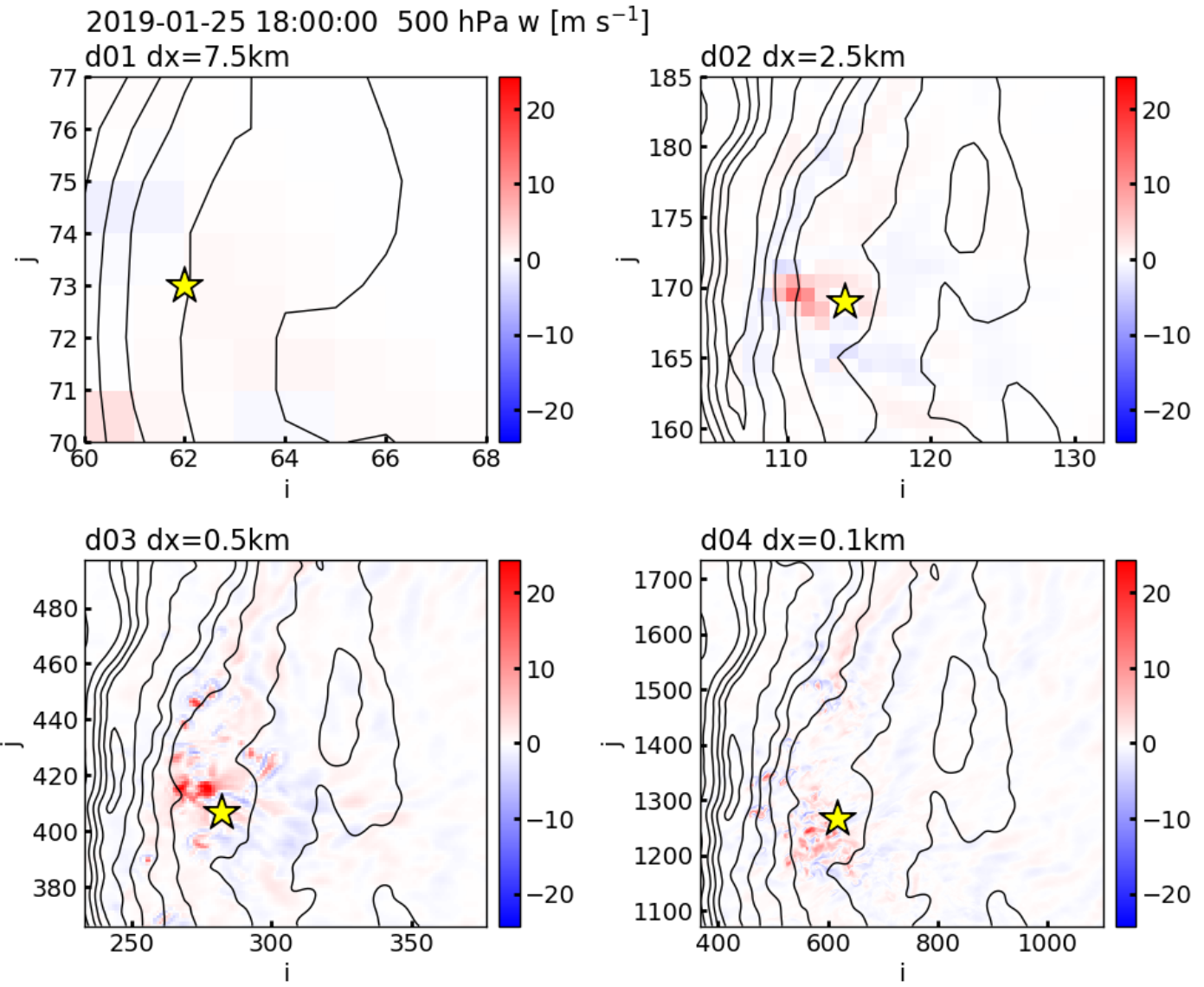
- Mesoscale ensembles run for each case date (example for 10-Nov-2018 at right)
  - 33 ensemble members based on ERA5, ERA5 Ensemble, FNL, and GFS Ensemble
  - Nested down to 2.5 km grid spacing
  - Best performing ensemble members identified based on cloud comparison to GOES-16 IR data
  - Down-selected ensemble members get final vetting using bulk CSPAR2 statistics, e.g., 20 dBZ echo-top height



LES domains

- "Ndown" from D02 to D03
- Nesting permits starting domains at different times to save resources and smooth spin-up process
  - D03 starts at 6 UTC
  - D04 can start at 12 UTC
- Primary LES run based on best-performing mesoscale ensemble member(s)
  - Some additional LES runs for testing other BCs or physics
- ~25 h wall time per model hour on 7168 cores of Cumulus-2

Comparison of 500 hPa Vertical Velocity at Each Grid Spacing



Modeling stages to achieve Δx=100 m

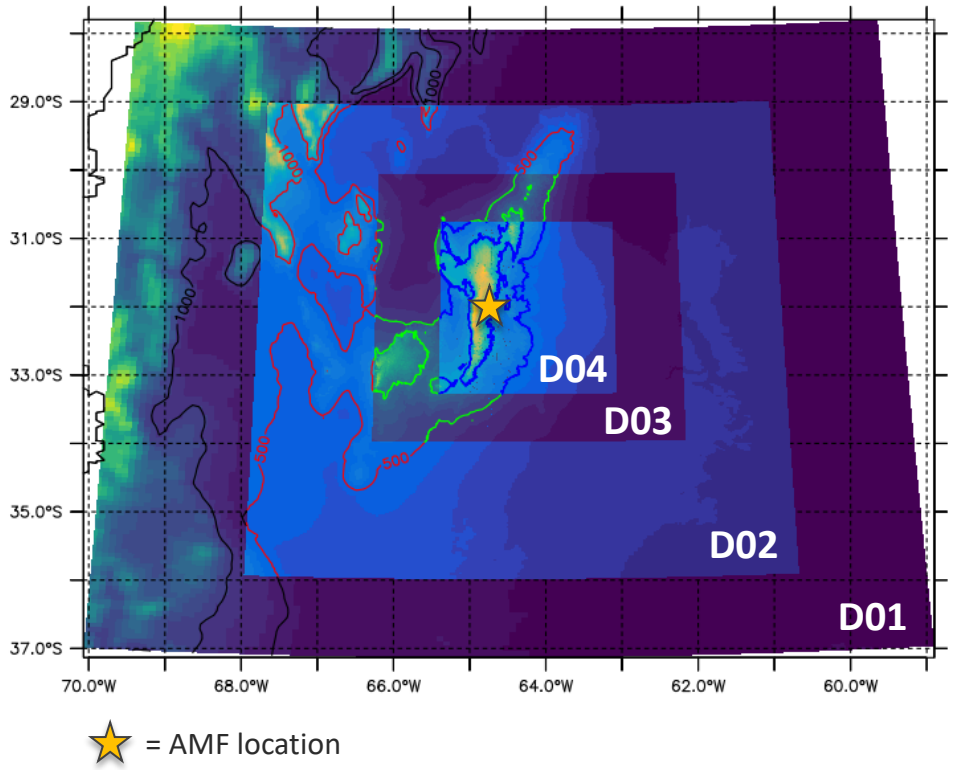
- **Stage 1:** Mesoscale ensembles with Δx=7.5 & 2.5 km
  - For selecting boundary conditions and case selection
- **Stage 2:** LES setup with Δx=500 & 100 m
  - For selected cases, some with several LES per case
- **Stage 3:** Post-process data to simplify usage

Domain Sizes

Δx =	D01 7.5 km	D02 2.5 km	D03 500 m	D04 100 m
N_x	130	258	750	2145
N_y	136	306	865	2775

Meso LES

WRF Model Domains
Δx = 7.5 km, 2.5 km, 500 m, & 100 m



Model output strategy

- Mesoscale ensembles have 33 members (20 dates), LES have 2–3 per case (9 dates)
- We expect the total dataset for the scenario to exceed 1 PB, possibly approaching 2 PB

WRF Model Data

Category	Domain(s)	Δx	Frequency	Period	Purpose
Meso	D01, D02	7.5 km, 2.5 km	15 min.	0–24 UTC	Full model state and diagnostics
Bridge	D03	500 m	15 min.	6–24 UTC	Full model state and diagnostics
LES	D04	100 m	5 min.	12–24 UTC	Full model state and diagnostics
Restart	D03 and D04		30 min.		Enable users to do restarts

We want to make it as easy as possible for users, but…

- Using these runs will be non-trivial due to the data size!
- Raw output sizes
  - Mesoscale ensemble for D02
    - ~325 GB per ensemble member
    - >100 TB for full set of cases and members
  - LES runs for D04
    - Raw output >35 TB per run
    - >1 PB raw model output for 10 cases & 2 LES/case
  - Subsets add to above sizes

Rough File Sizes for Each Domain

Δx =	D01 7.5 km	D02 2.5 km	D03 500 m	D04 100 m
N_x	130	258	750	2145
N_y	136	306	865	2775
Snapshot Size	0.6 GB	2.8 GB	19 GB	171 GB

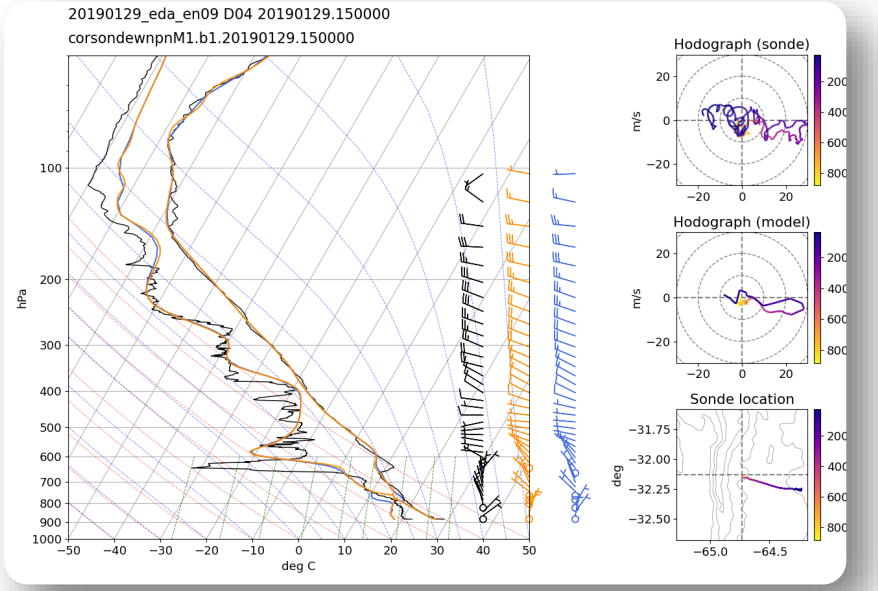
Subsets generated in post-processing

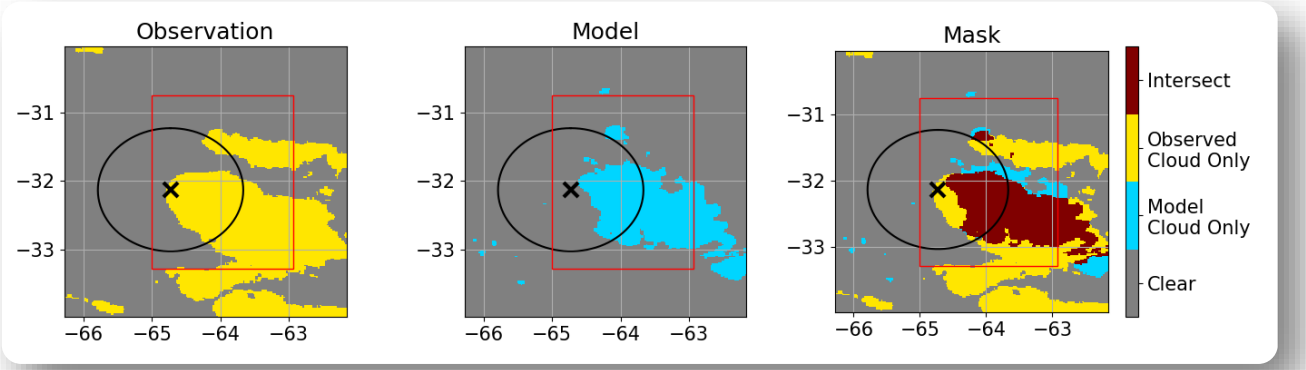
- Goal of reducing file sizes for users not needing whole raw files
- Extra diagnostics provided, e.g., LWP, CAPE, destaggered winds, heights, pressure
- Variable subsets grouped by theme in separate files*:
  - Static data, constant in time like terrain height, 0.1 GB
  - Meteorological state, 28 GB (with staggered variables interpolated to cell centers)
  - Meteorological state for staggered variables, 8 GB
  - Cloud data, 2 GB
  - Surface data, 0.4 GB
  - Boundary layer data, 5 GB
  - Radiation data, 0.2 GB
  - Aerosol data, 4 GB
  - Tendency data, 10 GB (e.g., microphysics tendencies & process rates)
  - Tracer data, 8 GB
- Subsets available on different height coordinates
  - Height above ground level
  - Height above sea level
  - Pressure levels
  - Raw model levels

*** File sizes given are per output time for a typical D04 subset file on raw model levels. Note that a wrfout_d04 is 171 GB.**

Multiscale Observational Datasets

- **Regional: Satellite-based**
  - Sources
    - VISST: IR brightness temperatures (11.2 µm channel)
  - Application
    - Time-dependent areal coverage of the convective cores
- **Local: Scanning C-Band Radar-based**
  - Sources
    - CSAPR-2/Taranis
  - Applications
    - Locate AMF-storm position within the LES grid
    - Time series of surface rain rates, and of radar echo-top heights (varied dBZ)
- **Point Measurements**
  - Sondes (ARM & RELAMPAGO)





Skill scores to evaluate simulations

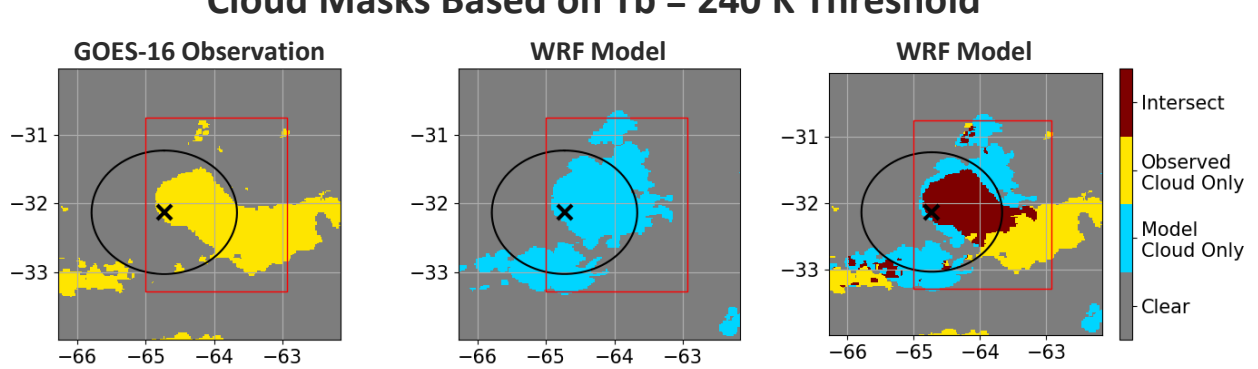
Satellite brightness temperatures (Tb) for convective area development

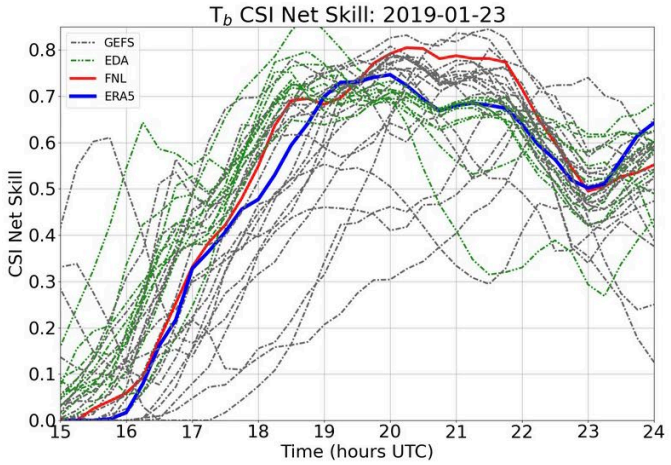
Radar echo-top heights for local convective intensity

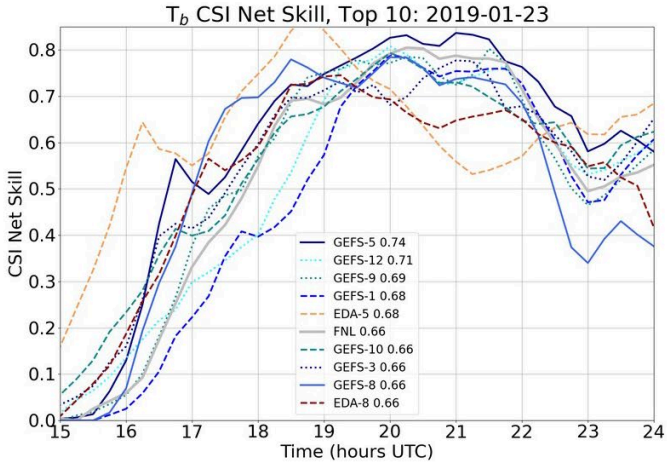
Web pages with statistics for beta release simulations

17:00 2019-01-23

Skills:
Freq Bias: 0.72
CSI: 0.37
CSI Net: 0.51







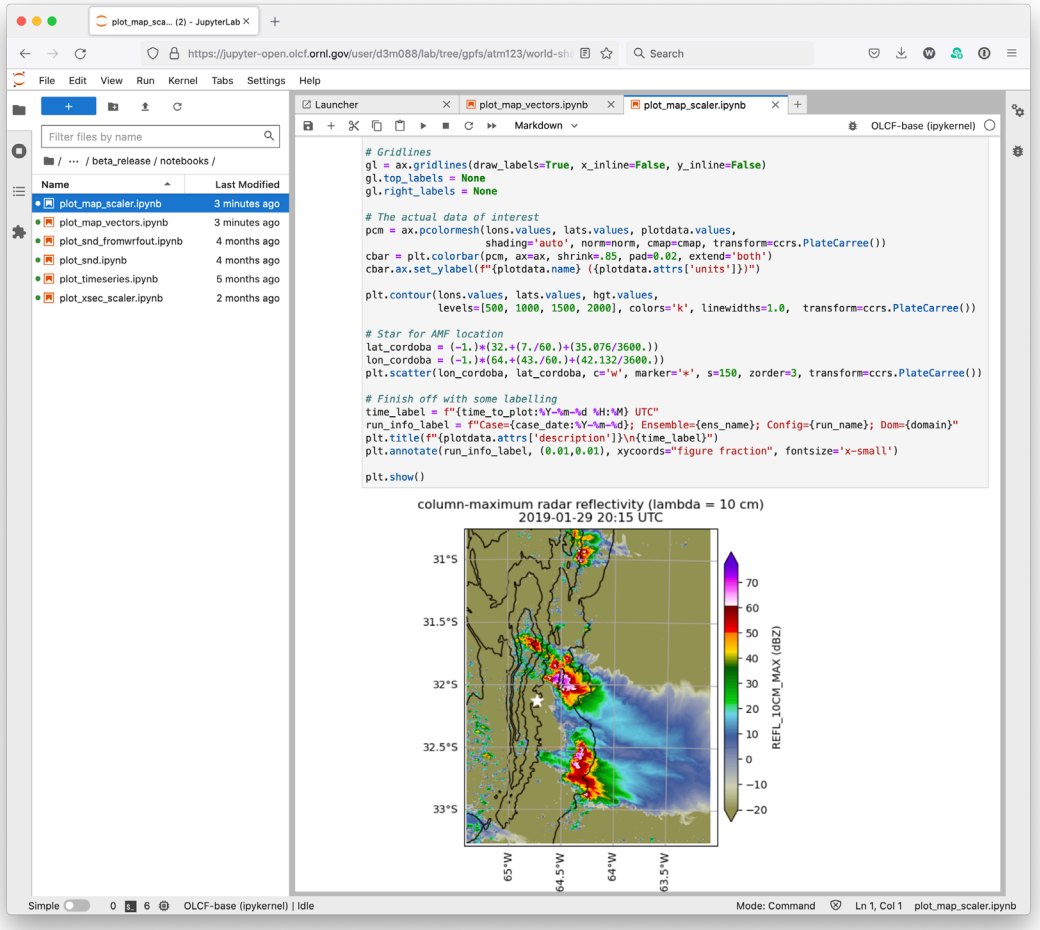
Scoring based on Critical Success Index, Frequency Bias, RMSD

LASSO-CACTI Beta Release Contents

- Ensembles of ~33 members for 20 case dates for mesoscale domains (Δx = 7.5 & 2.5 km)
- Three LES simulations (Δx = 500 & 100 m), ask if you want access to others
- Included files
  - Input data and run directories
  - Raw model output for wrfout and wrfrst files
  - Example subset files for EDA09 member on 29-Jan-2019 and code to do one's own subsets
- Animations for GOES-16 infrared and visible data for the 20 case dates
- Skill scores versus Tb and ETH data (see previous slide)
- Example Jupyter notebooks for plotting
- Beta documentation
  - https://discourse.arm.gov/t/lasso-cacti-beta-release-documentation/118
- Contents are evolving as we near the full release

Accessing the beta release…

- Still working out details for accessing LASSO-CACTI via the Bundle Browser and traditional ARM methodologies
- Files currently reside on ARM's Cumulus-2 cluster
- Two methods for access
  - ARM's Jupyterlab server
  - Interactive logins and job submissions on Cumulus-2
- One account works for both methods
  - https://www.arm.gov/capabilities/computing-resources



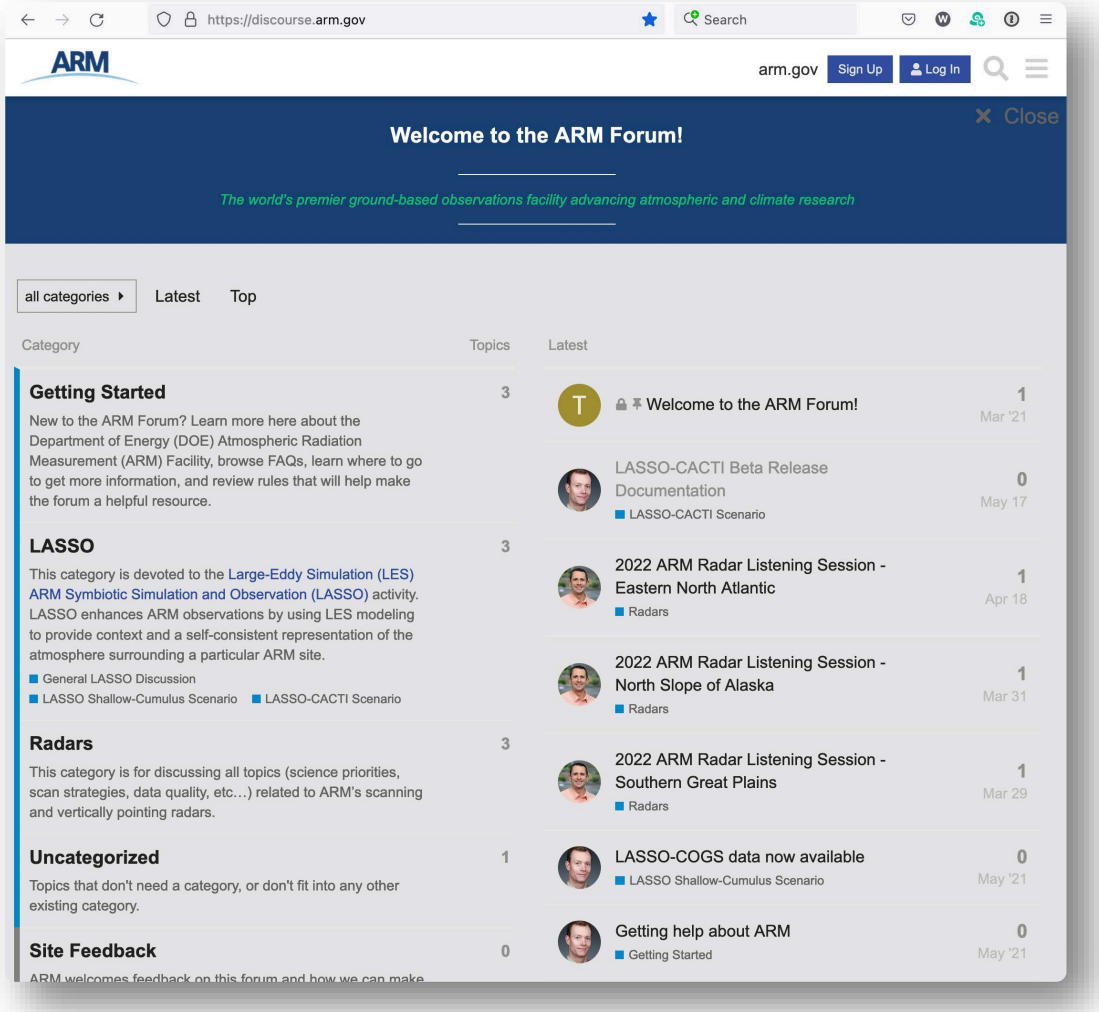
The production version release

- Hope to release the LASSO-CACTI dataset in early 2023
- What's left to do?
  - Working with ARM Data Center (ADC) to make ARM's infrastructure able to handle 1 PB of model data
    - Storage and recall of files, e.g., filename violations, quantity of datastreams, bundling of files with htar on HPSS
    - Hierarchical DOIs
  - Post-processing of raw model output into subsets and files with appropriate filenames
  - ADC working to develop a modified Bundle Browser to ease user discovery
  - Writing documentation—using a dynamic, online format similar to ReadTheDocs
  - Anticipate some kinks to work through to ease staging of data and user access on Cumulus/Jupyterlab
  - Still seeking improvement for ~4 of the LES case dates
    - Impression is a general bias toward weak convection

Simultaneously starting prototyping for LASSO-ENA…

Join the community! Online forum for LASSO, etc.

- Check out the online forum for LASSO: https://discourse.arm.gov/
  - For user support, discussing scenario development, and related topics around LASSO and ARM
  - Aiming for it to be an online resource for LASSO information and support
  - Other ARM topics besides LASSO are also possible—ask us if you would like a category added, e.g., for a field campaign or value-added product
- Email LASSO PIs (Bill & Andy) at lasso@arm.gov
- LASSO-CACTI posters this week
  - Gustafson et al., #67
  - Vogelmann et al., #66



Extra Slides

Input data

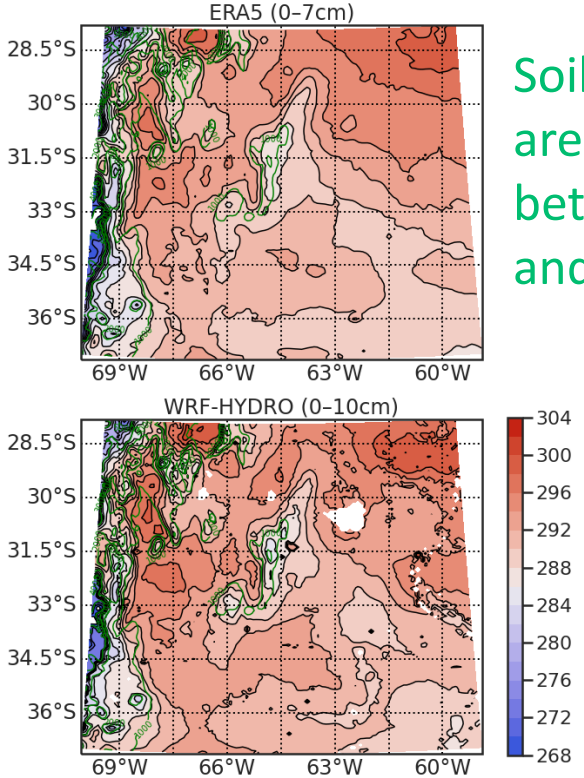
- Using MERIT DEM data for terrain elevation (Yamazaki, GRL, 2017)
  - Raw data at 3”, Δx ~ 90 m at equator
  - Smoothing for model stability using ~1 km spatial scale
- Soil initialization with WRF-Hydro to establish a spun-up soil state consistent with WRF physics
  - Continuous run from August 2018 using Δx=2.5 km
  - Driven by ERA5

Soil comparison for 26-Oct-2018 0 UTC:

ERA5 →

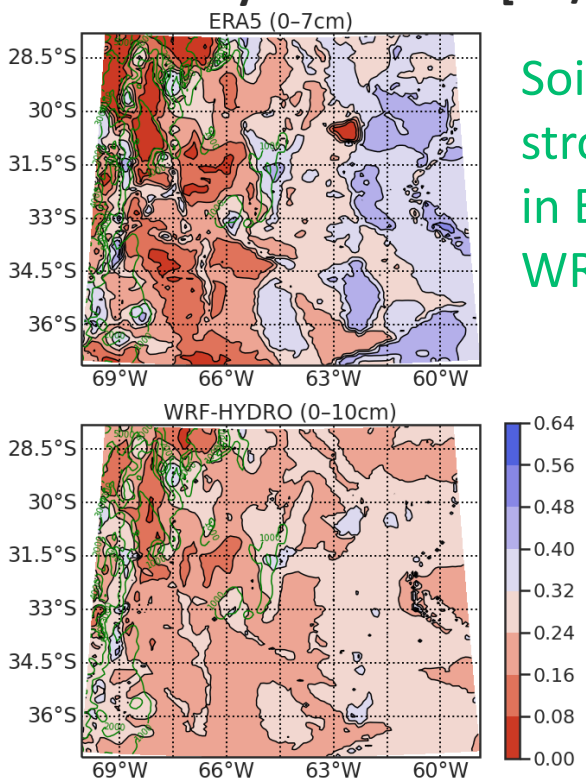
WRF-Hydro →

Topmost Soil Layer Temperature [K]



Soil temperatures are similar between ERA5 and WRF-Hydro

Topmost Soil Layer Moisture [m³/m³]



Soil moisture has stronger gradient in ERA5 versus WRF-Hydro

Model physics configuration

- Basic physics setup is a derivative of WRF's "CONUS" physics configuration

Physics Option	Number	Name
mp_physics	28	Aerosol-Aware Thompson Microphysics, Aerosol data from GEOS-5 model
cu_physics	6	Modified Tiedtke Cumulus (only Δx=7.5 km)
ra_lw_physics	4	RRTMG Longwave Radiation
ra_sw_physics	4	RRTMG Shortwave Radiation
bl_pbl_physics	2	Mellor-Yamada-Janjic TKE PBL (only Δx=7.5& 2.5 km)
km_opt	2	1.5 Order TKE SGS (only Δx=500 & 100 m)
sf_surface_physics	2	Noah Land Model