

# CACTI Convection Modeling Use Case

Compiled May 2022 by Model Data RCN team

## Summary

Weighted rubric score - 51

Category - Preserve selected simulation workflow outputs

- Use Case Description
  - *COMMENTARY NOTE BY RCN TEAM: This use case was based on a proposed simulation project that has not yet been completed.*
  - High-level overview of the use case
    - Cloud, Aerosol, and Complex Terrain Interactions (CACTI) field campaign, focuses on the Argentina area, funded by DoE.
    - ARM provides high resolution large-eddy simulations (LES) during the field program. Goal is to use CACTI data to study convection initiation, model errors, physics errors.
    - Initialization will be provided by DoE ARM, observations hosted by CACTI campaign.
  - Science goals and basic workflow
    - The planned approach is that they won't need to save the run output, just the documentation, configuration, diagnostics, and some 2D fields. They don't have to save full 3D fields.
    - Goal: perform large-eddy simulations for selected events of significance using a different microphysics parameterization as well outputting additional process-based diagnostics.
    - WRF output/diagnostic files in netCDF. Simulations will be done at a 500-m and 100-m horizontal grid spacing with domain sizes of approx 750 x 860 grid cells and 2100 x 2700 grid cells, respectively.
    - Workflow: Grab ARM simulations, rerun these simulations, by changing the physics options. Test a different physics suite, compare how the LES simulations look. Second goal, investigate the tendency diagnostics, how they relate to dynamics.
- What use-case specific additional materials should be preserved and shared?
  - Data
    - Inputs to model
      - ARM LES provided by DoE -initialization data.
        - Accessible through LASSO database
        - <https://www.arm.gov/capabilities/modeling/lasso/>
        - 2 Pb total
    - Raw model output
      - Potentially to produce 180 Tb
        - Model produces 3D variables.
        - Tendency outputs

- Full vertical model resolution
- Processed model output
  - Purpose is to compare the simulations with the CACTI observational database. CACTI observations are managed by DoE, and will be used as comparison with model output.
  - 6 TB model output is the planned post-processed files to archive
    - Vertical resolution interpolated to pressure levels - There will be extensive thinning of vertical levels
      - Standard levels
    - Cloud microphysical property parameters 2D variables of importance will be archived
      - Cloud top temperature, Radar reflectivity, CAPE, etc...
      - Environmental conditions, moisture at various levels
    - 2D geospatial information is very important to help diagnose model errors.
- Software - plan to share via a GitHub repo and point to relevant public code pieces.
  - Model configuration
    - Namelists - plan is to use the same configuration as LASSO, as close as possible. Will provide details about the configuration in publications or other documentation.
  - Preprocessing code
    - Will be preprocessing LES data to pre-coarse it for coarse grain simulations.
    - No set plan for if/how to share this code yet
  - Model code
    - WRF is public model
    - The project will be modifying Tendency components of the model
    - Could put code and information about this on Github
  - Postprocessing code
    - Lots of WRF post-processing code available from the community, in [WRF MET](#).
- Other
  - Documentation
    - Haven't thought about this much. But they plan to be sharing tendency diagnostics on Github, and documenting the versions of WRF used for the project.
  - Visualizations or images
    - None, beyond typical article graphics.
- Why should these things be preserved and shared?
  - General
    - To show the validity of the results.

- Requirements of publications
  - Helpful for the community that uses LASSO. Stimulate ideas of other comparisons for physics schemes in LASSO.
- Reasons why the things listed above are important
  - Note expected/intended audience and what they expect/need
    - Are there specific people who will be using the data downstream?
      - Not initially.
    - Possible/aspirational users?
      - If funded, the project team will inform the LASSO community about these products.
  - Note any temporal considerations, such as particular products that become more/less useful over time
    - Will still be useful over time as comparison with CACTI observations. Not everybody can run simulations of this size. Should not reduce in value, unless errors are found later.
    - Other perspective is that this is comparing with a specific configuration of LASSO, so as LASSO changes, this data may be less relevant
- Broader Impacts:
  - How will output from this project be used by stakeholders?
    - Main collaborators are DoE labs and NCAR researchers.
    - NCAR has South America affinity group, which is a potential avenue for sharing results.
  - How were stakeholders involved in the data curation decision-making?
    - Not discussed
  - How will stakeholders be compensated for their participation in the data curation decision-making process?
    - Not discussed
- Do you have any concerns about misuse of your data or software? If so, what concerns do you have, and what are the reasons for those concerns?
  - None. This project is focused on cloud process, convection initiation, which are not a concern to be shared.