Toward the Development of Strongly Coupled Land-Atmosphere Data Assimilation for E3SM Using DART

Zhaoxia Pu[1]

Shixuan Zhang[2], **Dongze Xu**[1], **Brett Raczka**[3], **Jeffrey Anderson**[3], **Ruby Lueng**[2]

[1] *Department of Atmospheric Sciences, University of Utah*

[2] *Pacific Northwest National Laboratory*

[3] *NSF National Center for Atmospheric Research*

Coupled Land-Atmosphere-Ocean Data Assimilation for E3SM with DART for Understanding Subseasonal-to-Seasonal Predictability of Extreme Events (DOE Award: DE-SC0025198)

Zhaoxia Pu, *University of Utah* (Principal Investigator)

Shixuan Zhang, *Pacific Northwest National Laboratory* (Institutional PI)

L. Ruby Leung, *Pacific Northwest National Laboratory* (Institutional co-I)

Jeffrey Anderson, *NSF National Center for Atmospheric Research* (FFEDC PI)

E3SM Webinar

November 20, 2025

Project Overall Objectives

- **Develop an effective coupled atmosphere-land-ocean data assimilation system for E3SM**
  Build and optimize **E3SM-DART** for **land–atmosphere–ocean coupling**, advancing from standard (100 km) to high-resolution (0.25°) configurations, including **regional refinement** and **strongly coupled assimilation** of soil moisture, SSTs, and atmospheric data.

- **Establish balanced initial conditions with reduced spin-up**
  Use coupled data assimilation (CDA) to minimize initial shocks, improve consistency of atmosphere–land–ocean states, and shorten time to equilibrium, enabling more efficient high-resolution simulations.

- **Advance subseasonal-to-seasonal (S2S) predictability research**
  Employ CDA and hindcast simulations to assess the sensitivity of MJO and ENSO predictability to ensemble size, model resolution, and coupled model uncertainties, establishing standardized workflows for S2S studies.

- **Enhance prediction of weather and extremes**
  Apply improved hindcasts to study extreme precipitation anomalies and tropical cyclone genesis, with a focus on the Maritime Continent, western Pacific, and Atlantic basin, providing quantified prediction metrics and diagnostics.

This presentation emphasizes the current progress toward strongly coupled land–atmosphere data assimilation.

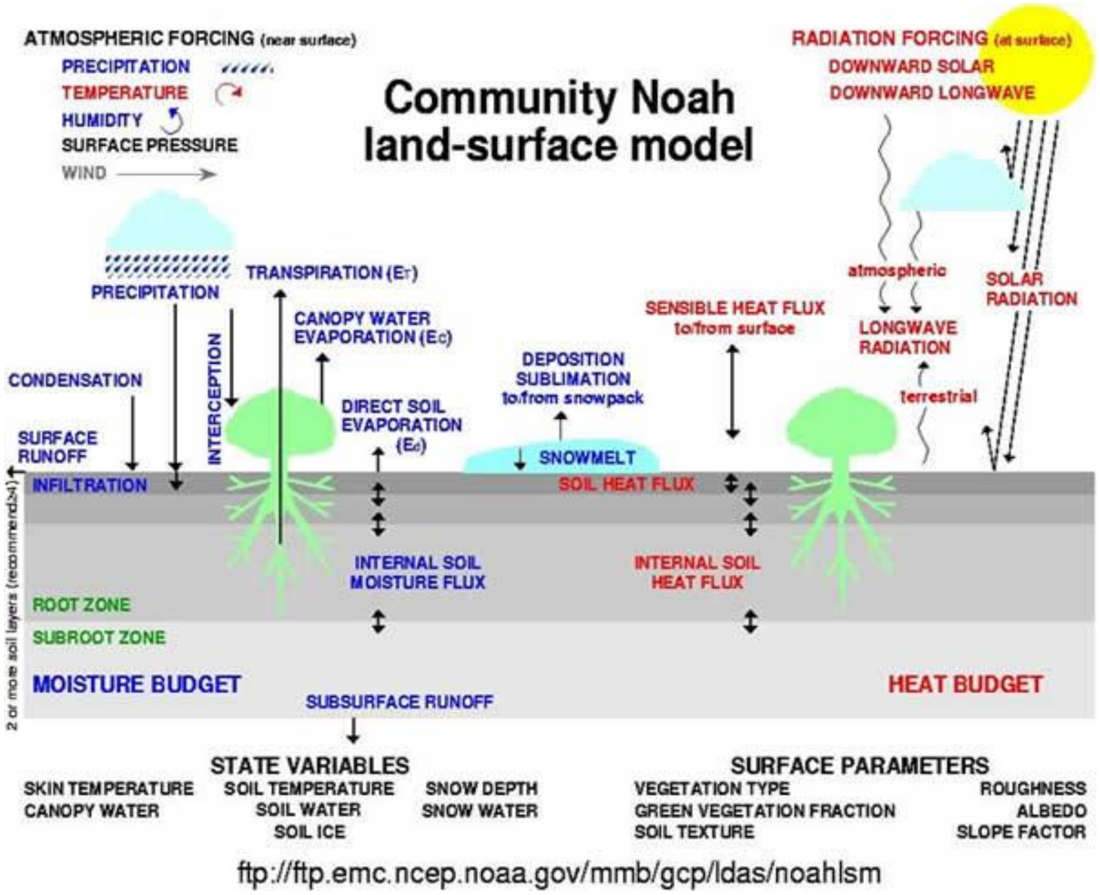
Science Background: Land-Atmosphere Coupled Data Assimilation

Land–atmosphere interactions are critical processes that influence weather, subseasonal, seasonal, and longer-term variability.

Poor representation of land-surface processes can introduce prediction biases across all timescales, leading to failures in forecasting extreme events.

Soil moisture is a key land variable that affects weather prediction and the occurrence of extreme events.

It is essential to **develop coupled land–atmosphere data assimilation** to improve prediction skill in Earth system models.

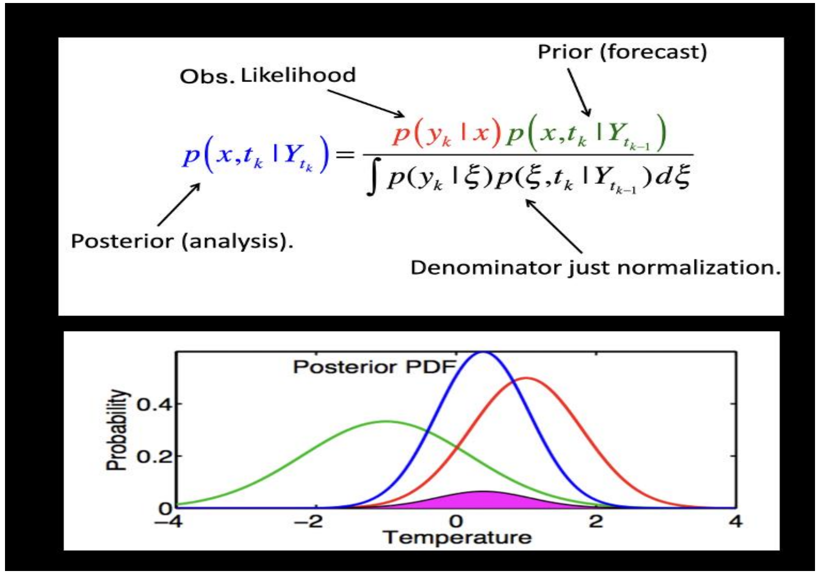


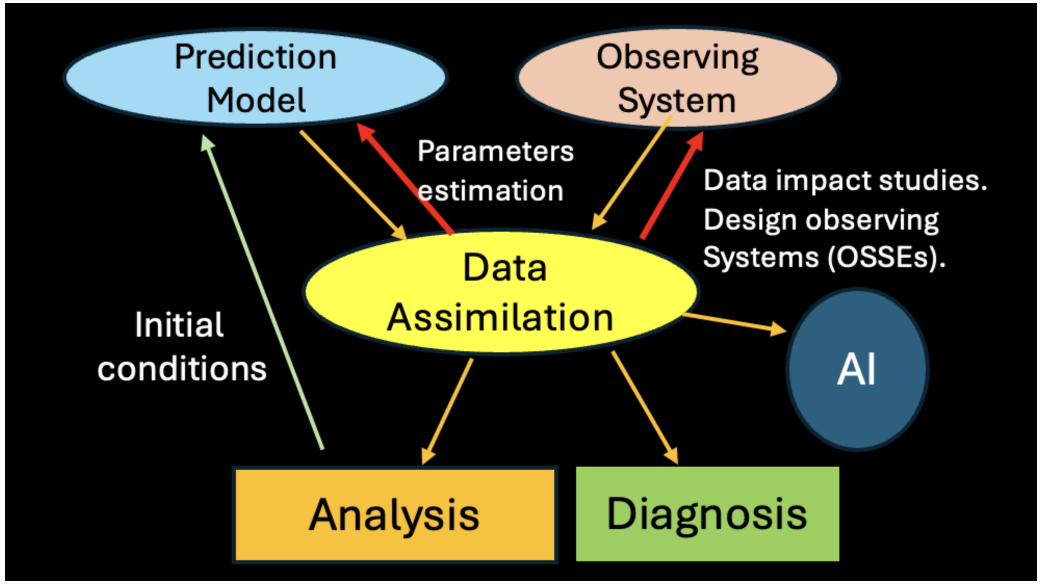
Science Rationale: Data Assimilation to Overcome Challenges with E3SM

- **E3SM integrates atmosphere–ocean–land–cryosphere systems, creating tightly coupled interactions.**

- **Initialization Challenges:**
  - ➢ Independent spin-up of components (reanalysis/climatology) introduces imbalances.
  - ➢ Stability is often prioritized over accuracy, leading to model shocks.
  - ➢ Initial drift from observed states reduces predictability

- **Role of Data assimilation (DA)**
  - ➢ Provides realistic initial conditions.
  - ➢ **Strongly coupled DA** ensures consistency across interfaces.





Science Rationale: SCDA vs. WCDA

Strongly-Coupled Data Assimilation (SCDA)

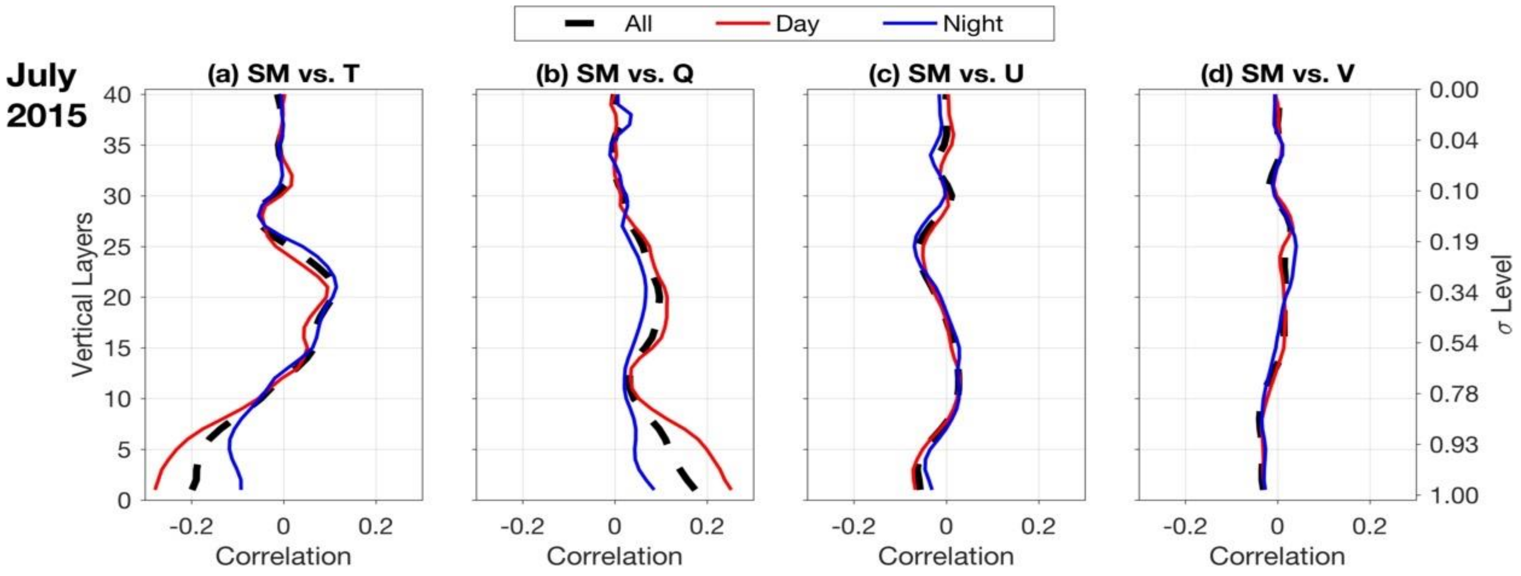
- Land and atmospheric data assimilation were performed simultaneously with the influence of each other; soil state (e.g., soil moisture, soil temperature) was added as an analysis variable.

Weakly-Coupled Data Assimilation (WCDA)

- Land and atmospheric data assimilation are performed separately and then coupled during integration.

Lin and Pu (2019 & 2020)

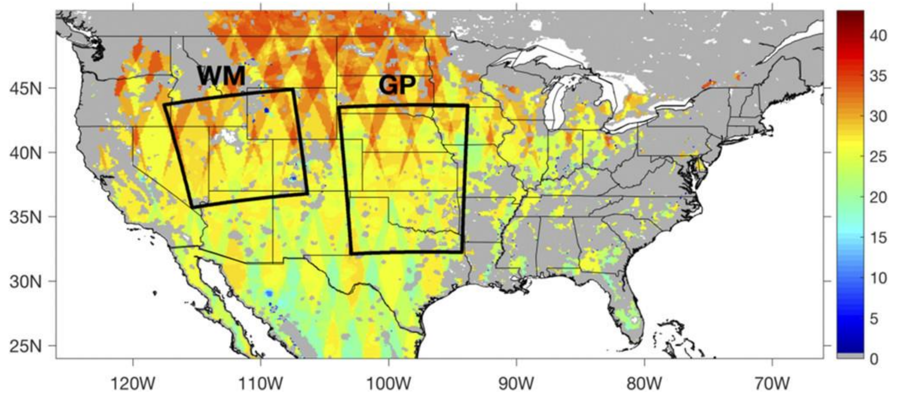
Covariances between soil and atmospheric states in a strongly coupled land-atmosphere data assimilation with WRF-Noah



Liu and Pu (2019); Lin and Pu (2018)

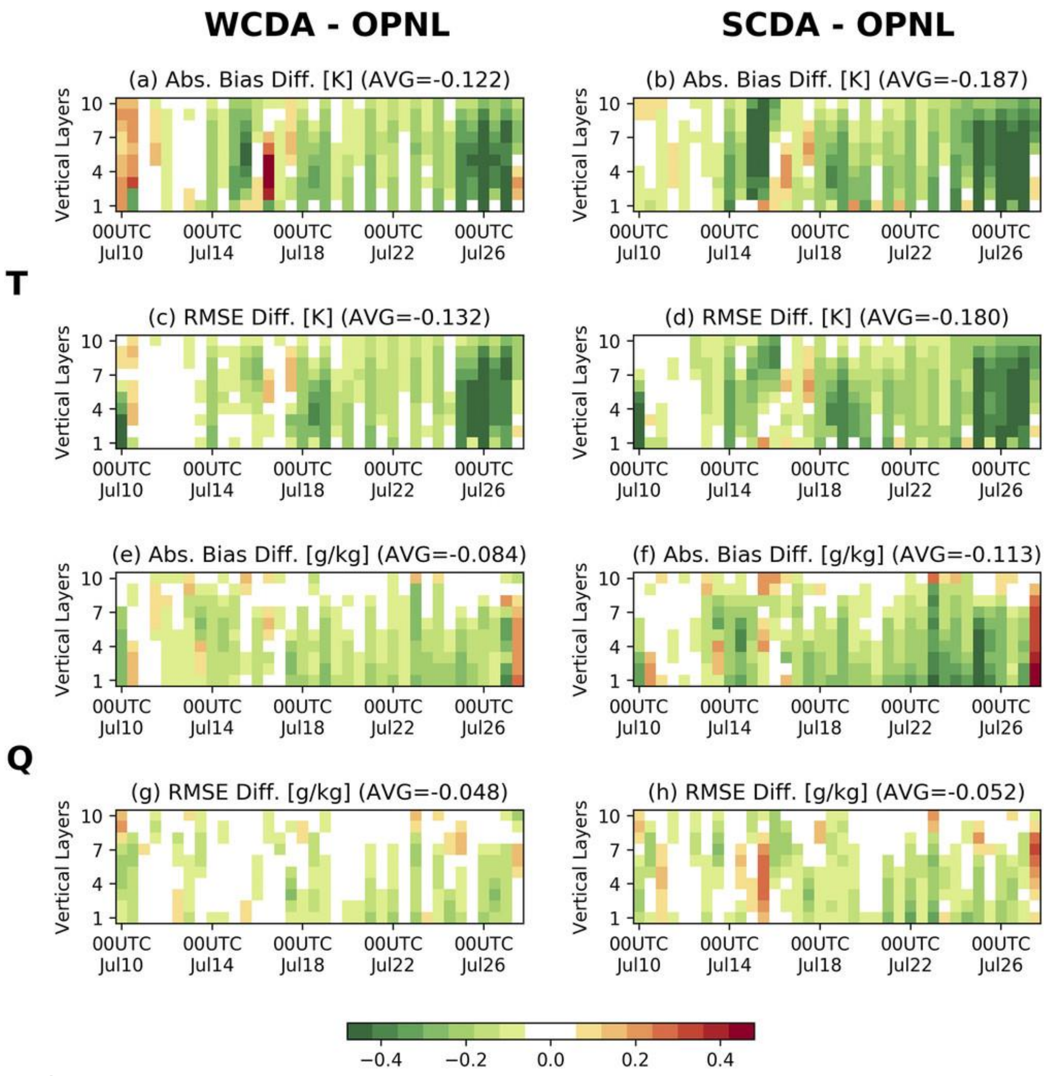
Science Rationale: Impacts of Strongly Coupled Land–Atmosphere Data Assimilation on Numerical Weather Prediction

(Liu and Pu 2019, JGR-A); Lin and Pu 2018, JAMC; Lin and Pu 2019, JAMC; Lin and Pu 2020, MWR)

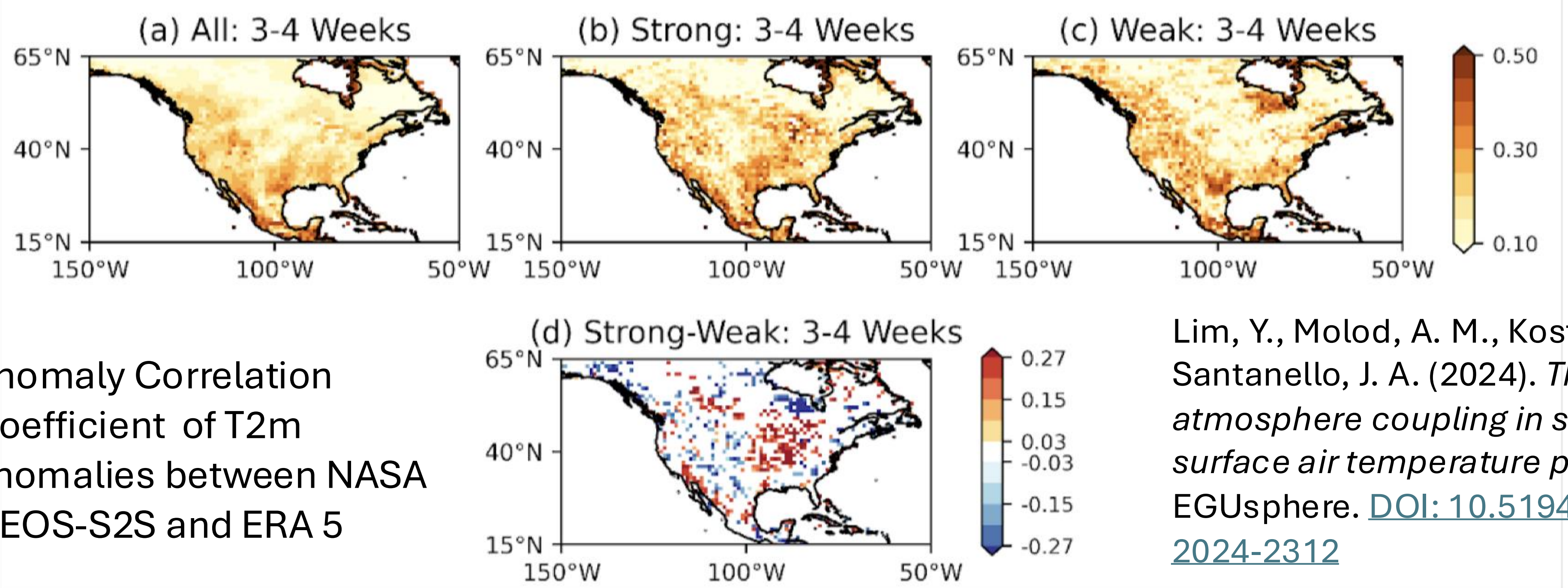


SMAP data 1 to 27 Jul 2016.

- Soil moisture has a significant influence on near-surface atmospheric conditions (e.g., 2-meter temperature and humidity) and boundary layer structures.
- Assimilating land data—especially soil moisture—improves near-surface weather prediction.
- Strongly coupled land–atmosphere data assimilation outperforms weakly coupled data assimilation.



Science Rationale: Land-Atmosphere Coupling and data assimilation for S2S Prediction



Anomaly Correlation Coefficient of T2m anomalies between NASA GEOS-S2S and ERA 5

Lim, Y., Molod, A. M., Koster, R. D., & Santanello, J. A. (2024). *The role of land-atmosphere coupling in subseasonal surface air temperature prediction*. EGUsphere. DOI: 10.5194/egusphere-2024-2312

- Assimilating land data—especially soil moisture—improves S2S forecasts of atmospheric variables.
- Strong land–atmosphere coupling acts as a memory mechanism, enhancing persistence and predictability in weeks 3–4 of forecasts.
- Different models/regions (e.g., GEOS, NorCPM, CESM setups) consistently show improved skill when land state is well-initialized, especially in high-coupling areas.

Guo et al. 2011; Lim, Y. et al. (2024); Nair et al. (2024)

Project progress: EAMv3-DART Ensemble Data Assimilation

Infrastructure and Workflow

(Shixuan Zhang, PNNL)

- Successfully integrated and validated the **DART** data assimilation system within **E3SMv3**, forming the **EAMv3–DART** workflow.

- System configured with **ne30pg2** grid, **80 vertical levels**, and **F20TR atmosphere-only** setup. **EAMv3** provides atmospheric state variables through its dynamical core and physics (convection, clouds, turbulence, radiation, aerosol interactions).

- **DART** assimilates conventional and satellite observations using QC, forward operators, and the **Ensemble Adjustment Kalman Filter (EAKF)**.

- Produces analyses every **6 hours**, feeding back to EAMv3 for the next forecast cycle. A **fully automated** HPC workflow handles ensemble initialization, parallel forecasts, assimilation, job scheduling, and data exchange.

- Demonstrated **stable cycling** and **excellent scalability** on DOE leadership-class supercomputers.

Project progress: EAMv3-DART Ensemble Data Assimilation

Infrastructure and Workflow

(Shixuan Zhang, PNNL)

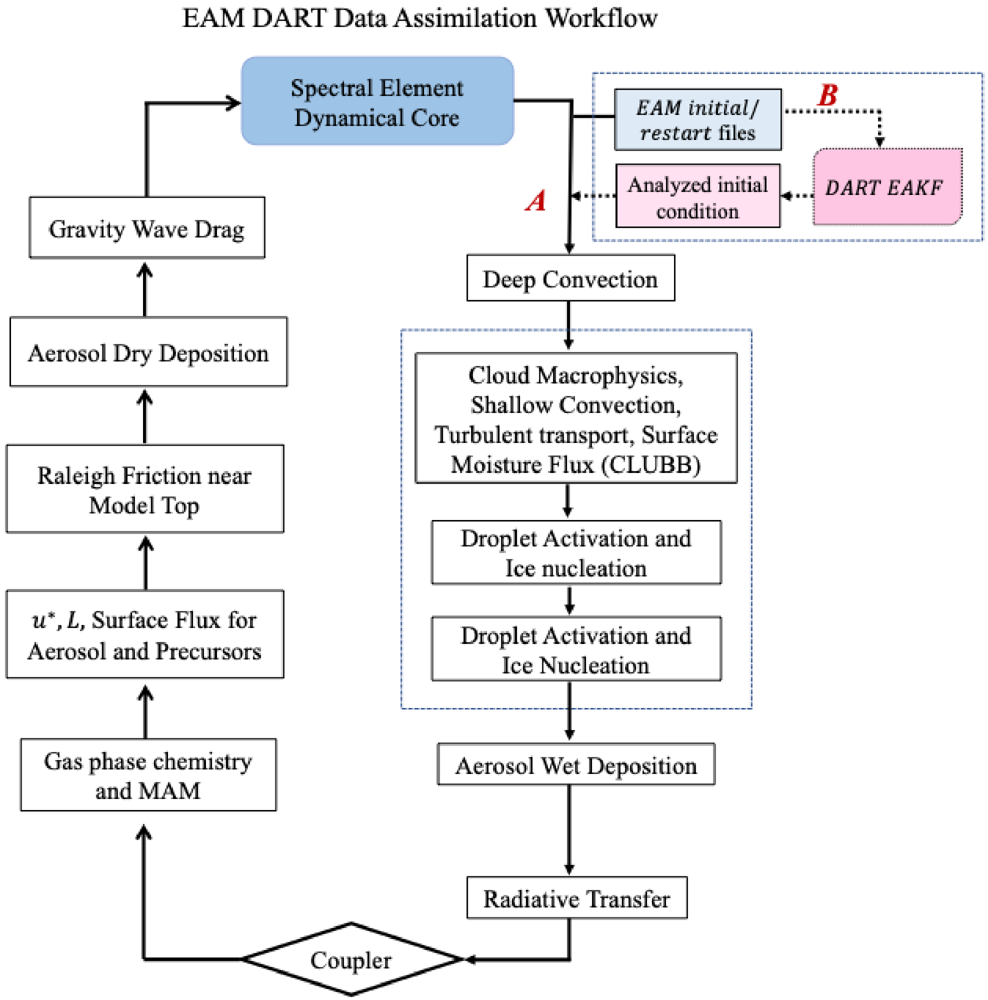
Table 1: DART configuration for atmospheric data assimilation in E3SM.

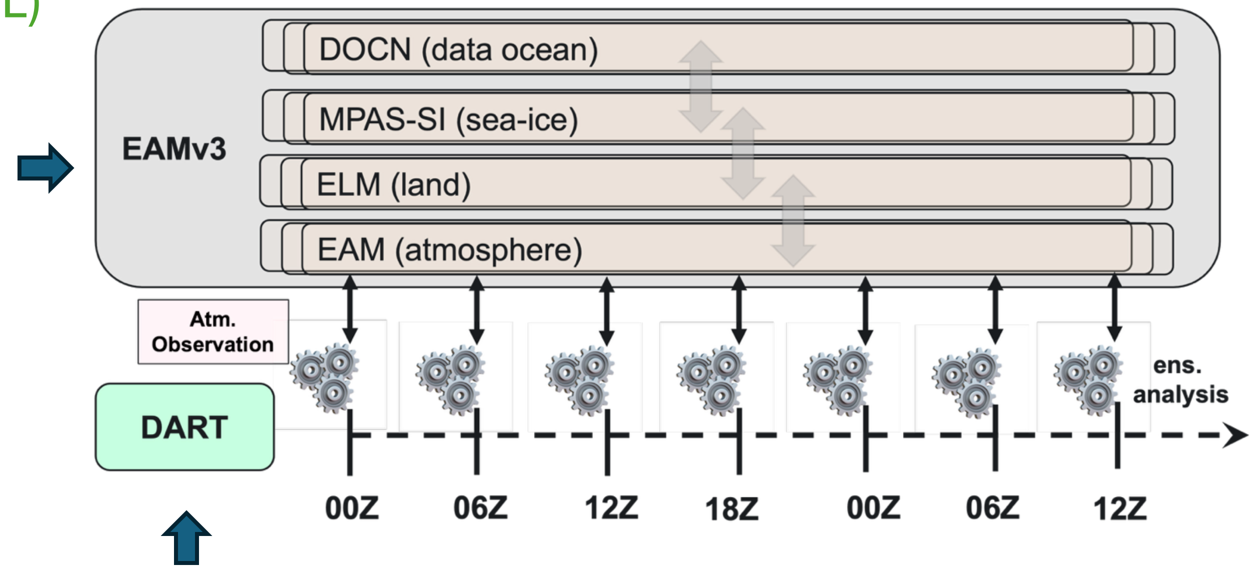
Component	Configuration
Method	DART Ensemble Adjustment Kalman Filter (EAKF; Anderson 2009)
State variables	U, V, T, Q, CLDLIQ, CLDICE, PS
Ensemble size	40
Covariance inflation	Adaptive inflation, spatially and temporally varying (Gharamti, 2018)
Localization function	Gaspari and Cohn (1999)
Horizontal half-width	0.2 radians (≈1000 km)
Vertical half-width	200 hPa
Observation error variance	Derived from NCEP PREPBUFR reports
Observations	Conventional and satellite observations

Project progress: EAMv3-DART Ensemble Data Assimilation

Infrastructure and Workflow

(Shixuan Zhang, PNNL)





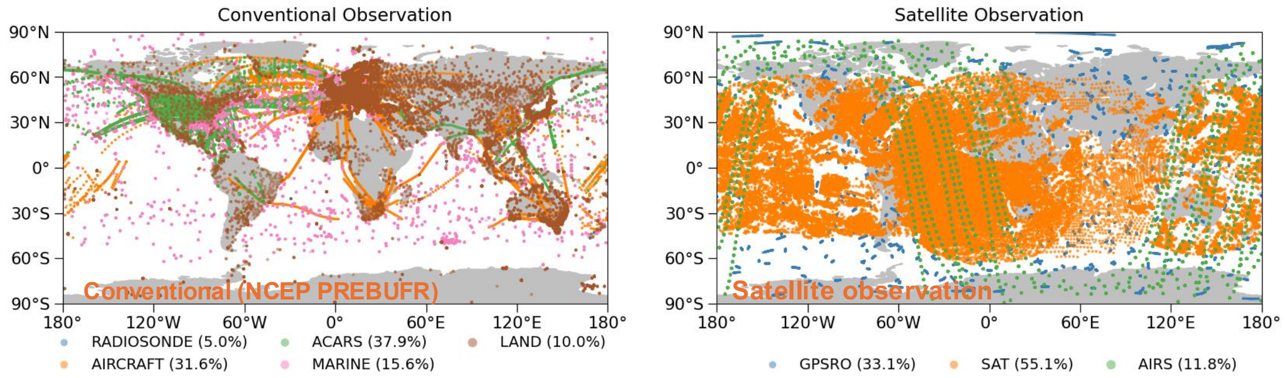
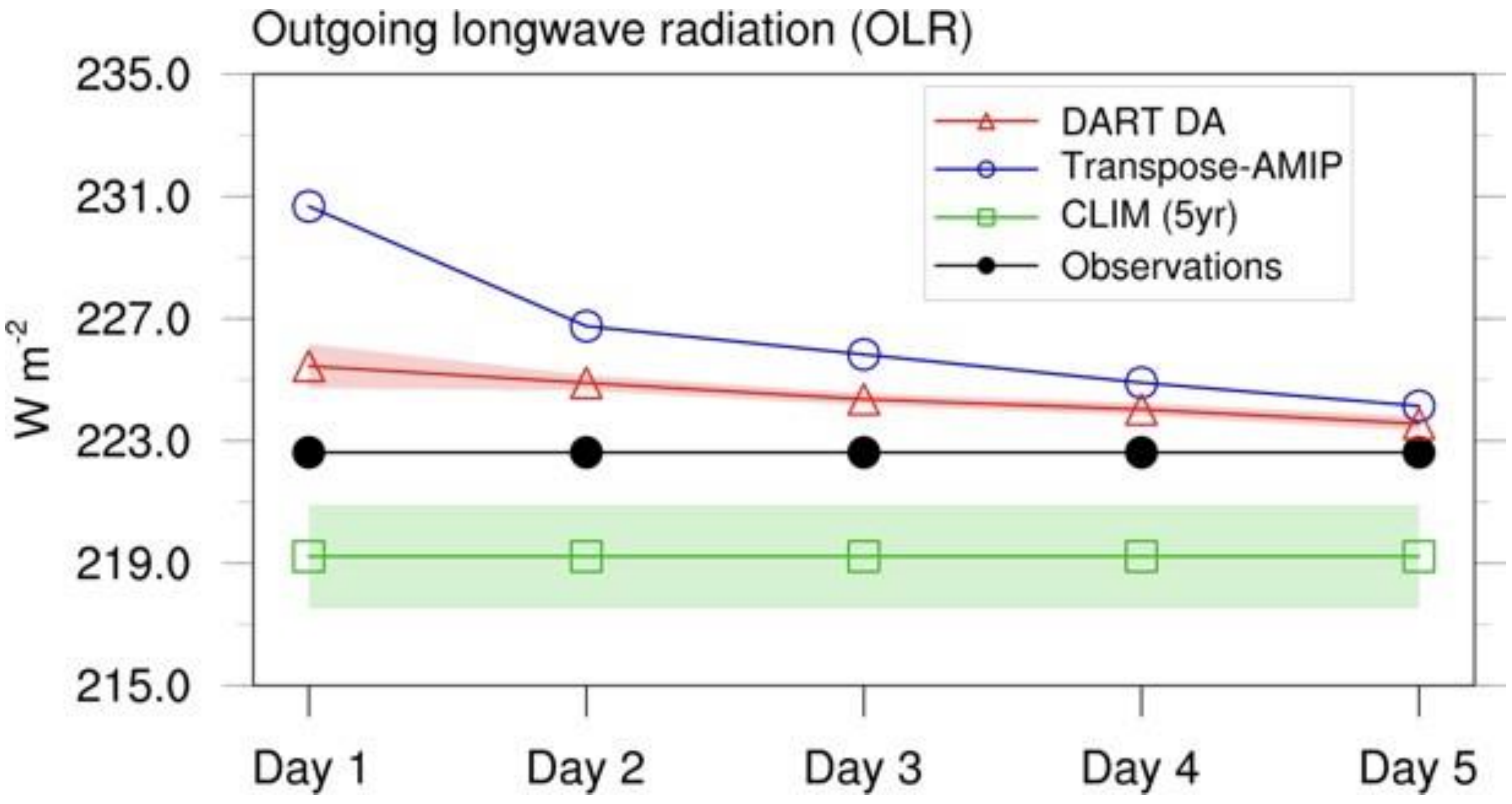


Figure 2. Observations (conventional + satellite) assimilated into EAMv3-DART during 1-day period, available at d345000 | DOI: 10.5065/JG1E-8525

Influence of EAMv3-DART Analysis on the Quality of Numerical Simulations

Reducing "initial shock" on the first day by data assimilation, as observed in the Transpose-AMIP simulation.

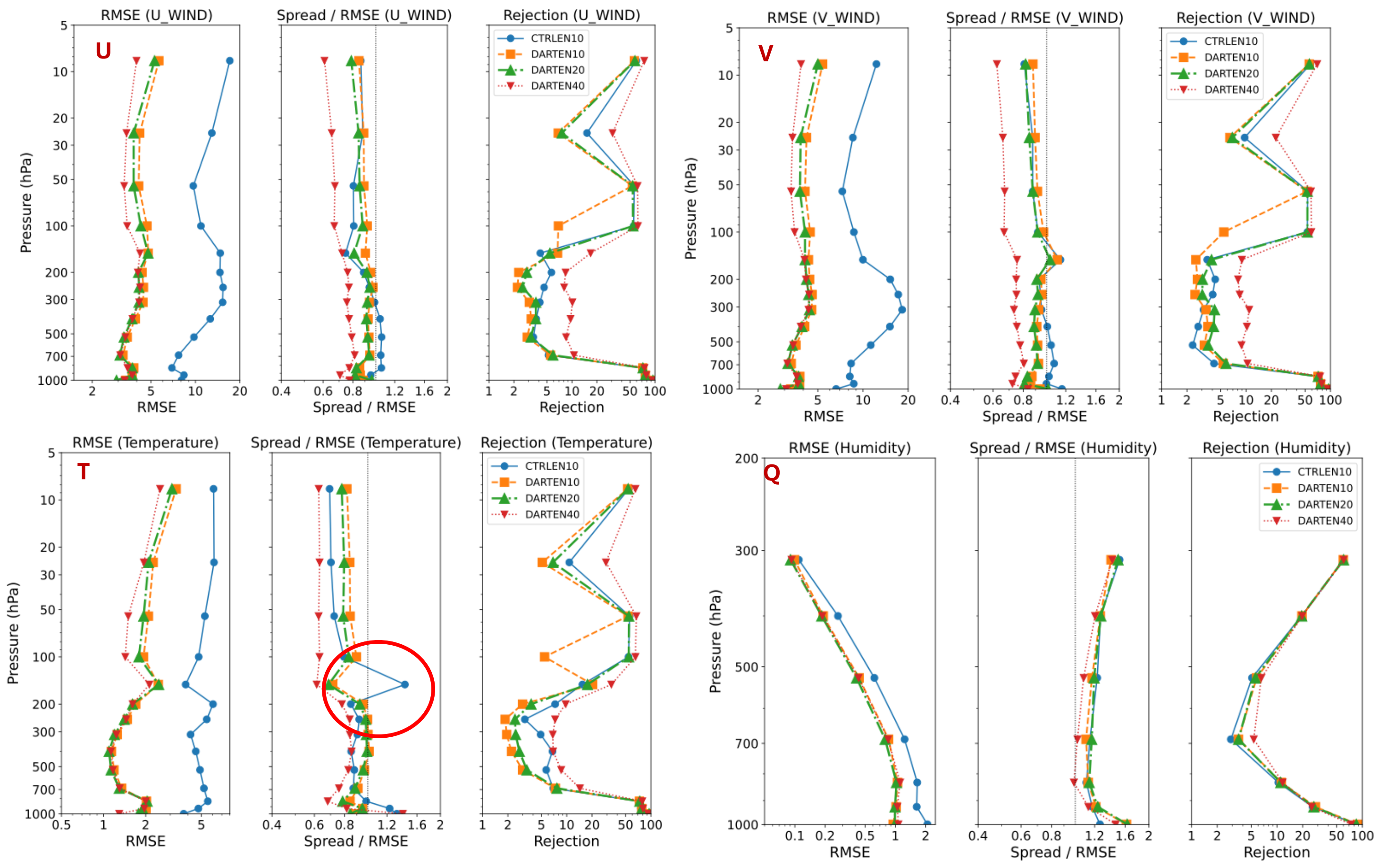


20-60N, 150E-60W
Simulations for Jan 2011

Shixuan Zhang et al. (2024)

Quality of EAMv3-DART Analysis

EAMv3-DART DA experiments for December 2011



Profiles of RMSE (first panel), Spread/RMSE ratio (second panel) and percentage data rejection rate (third panel) for U, V, T and Q derived from the EAMv3-DART DA experiments

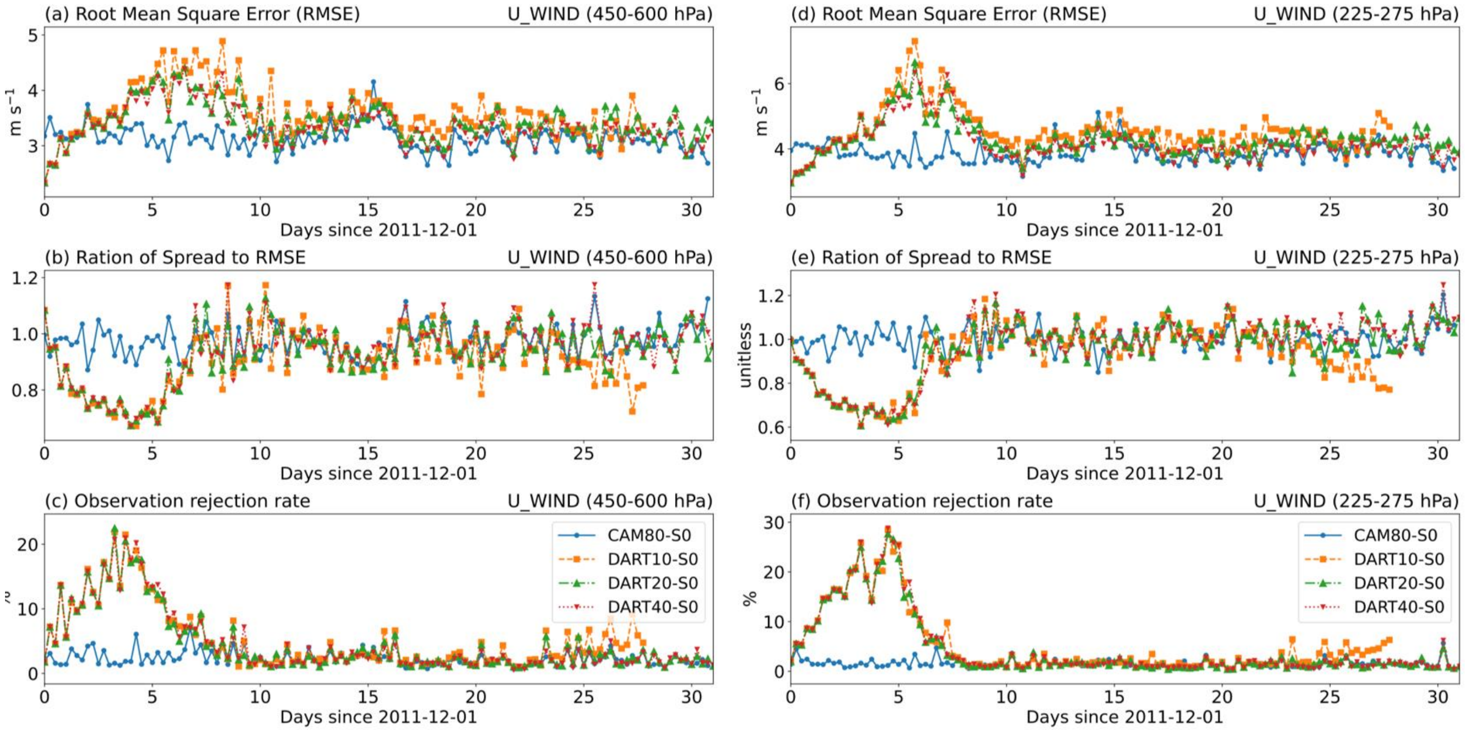
Quality of EAMv3-DART DA with varying ensemble sizes: Comparison with CAM6-DART

Larger ensembles improve DA performance, particularly at upper model levels

Overall, EAMv3–DART DA performance is **comparable** to CAM6–DART

Shixuan Zhang et al. (2025)

Quality of EAMv3-DART Analysis: Quick Sprin-up and Improved Performance



Time series of the zonal wind component (U WIND) during 1–31 December 2011 at 450-600 hPa (left column) and 225–275 hPa (right column).

EAMv3–DART quickly spins up (within ~7–10 days)

•The system stabilizes after the initial spin-up.

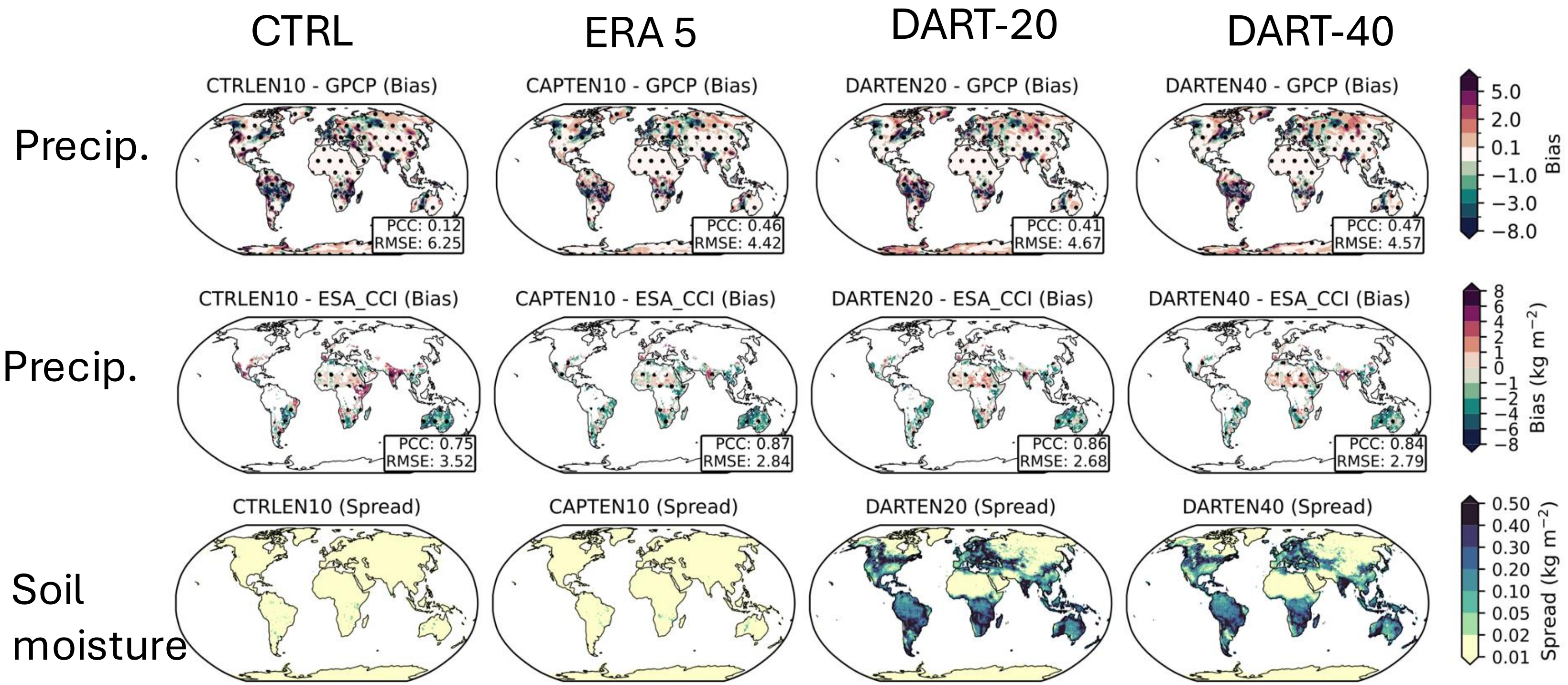
•Bigger ensembles greatly improve accuracy and reliability.

•Observation rejection rates drop as the system stabilizes.

•Overall performance is comparable to a trusted DA system (CAM6 80 ensemble members): E3SM with 20–40 ensemble members using EAMv3-DART can generate high-quality atmospheric initial conditions for forecasts and reanalyses.

Subseasonal-to-Seasonal Predictability in E3SM Using the EAMv3-DART System

Bias of precipitation and soil moisture ensemble spread - January 1, 2012



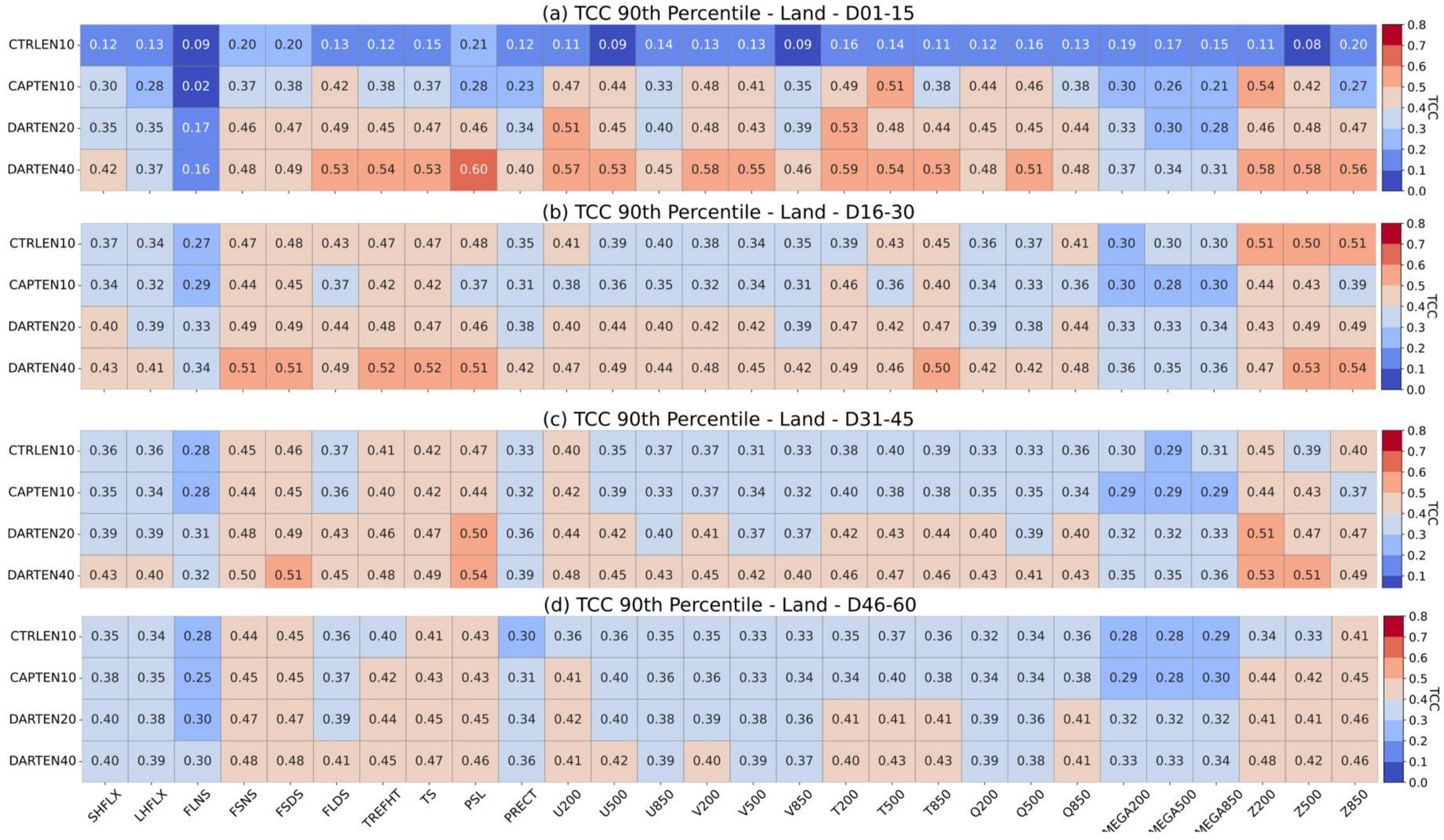
Biases: relative to GPCP or ESA CCI
PCC: pattern correlation coefficients

Shixuan Zhang et al. (2025)

Initialization strongly affects the realism of precipitation and soil moisture

- **CTRLEN10** (climatological initialization) shows the **largest spatial errors** in precipitation and soil moisture and exhibits **poor agreement** with GPCP and ESA CCI.
- **CAPTEN10** (ERA5-atmosphere-only initialization) improves large-scale precipitation patterns, raising correlations with observations.
- **DARTEN20/40** further improve regional precipitation realism and maintain **strong agreement** with ESA CCI.
- DART experiments show **greater, more spatially variable soil-moisture ensemble spread**, especially in the tropics and monsoon regions—unlike the weak, uniform spread in CTRLEN10 and CAPTEN10.

Subseasonal-to-Seasonal Predictability in E3SM Using the EAMv3-DART System



Temporal correlation coefficient (TCC) of the 90th-percentile anomalies for a suite of atmospheric and land variables over global land regions, evaluated for forecast lead times from days 1–15 through days 46–60. Higher values indicate better subseasonal predictive skill.

Shixuan Zhang et al. (2025)

Better initialization leads to better subseasonal-to-seasonal (S2S) forecast skill

Both CAPTEN10 and DART experiments outperform CTRLEN10 across almost all lead times (days 1–60).

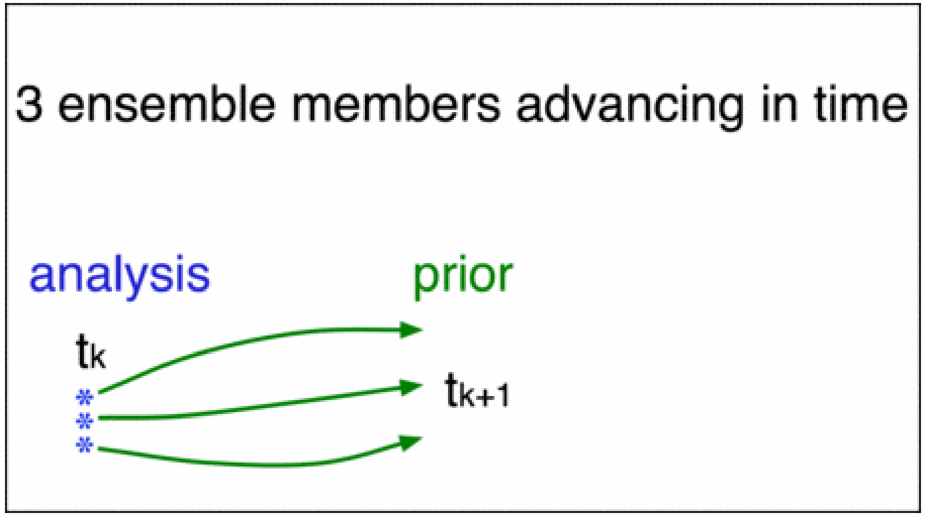
DART-initialized forecasts show the most consistent improvement, especially during days 1–30 for precipitation, soil moisture, surface fluxes, and near-surface temperature.

At longer leads (weeks 4–8), all experiments lose skill—as expected—but **DART40 retains slightly higher correlations** for several thermodynamic and hydrological variables.

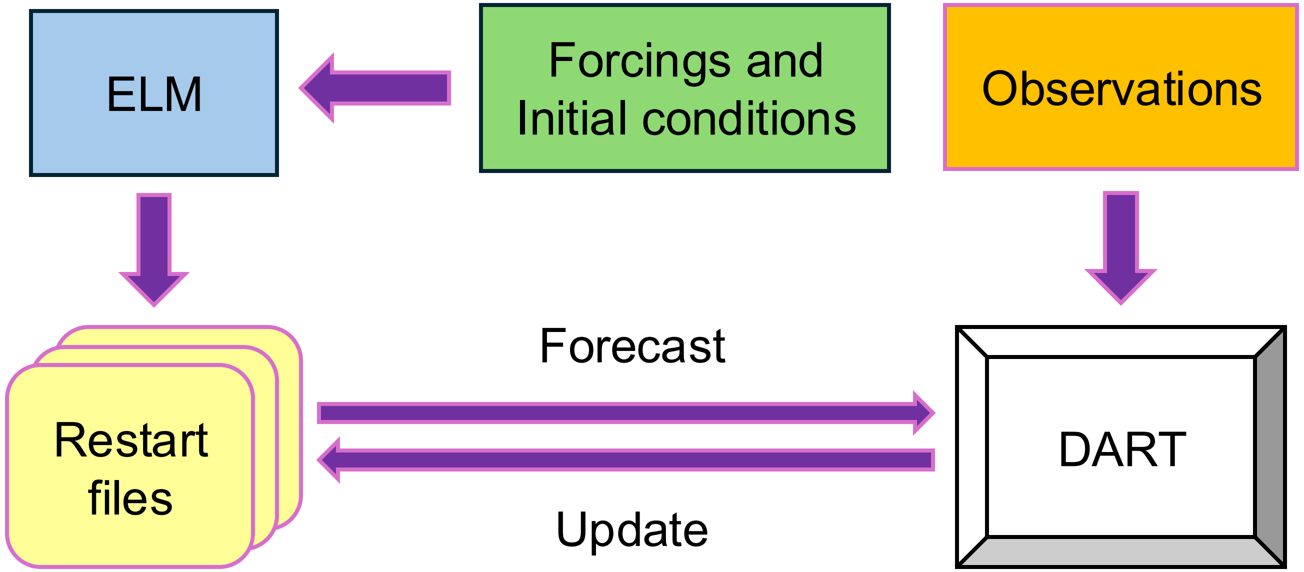
Project progress: ELM-DART Land Data Assimilation System

We developed a land data assimilation system for the E3SM Land Model (ELM) using the Data Assimilation Research Testbed (DART) based on CLM-DART.

Dongze Xu and Zhaoxia Pu, Univ. of Utah
Brett Raczka and Jeffrey Anderson, NSF NCAR



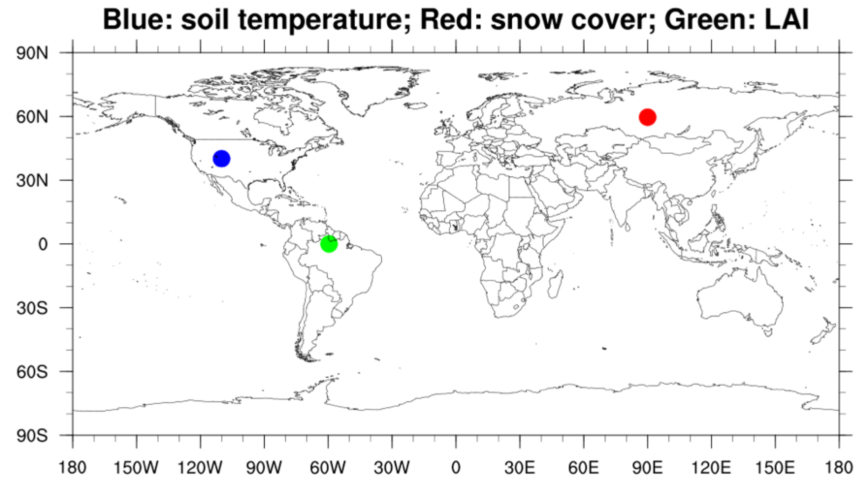
Building ELM-DART Land DA system



- ➢ Modified CLM-DART source codes to align with ELM.
- ➢ Revised DART interface for full compatibility with ELM.
- ➢ Configured and optimized environment settings to ensure consistent operation of ELM and DART.

ELM-DART Single Obs. Experiments

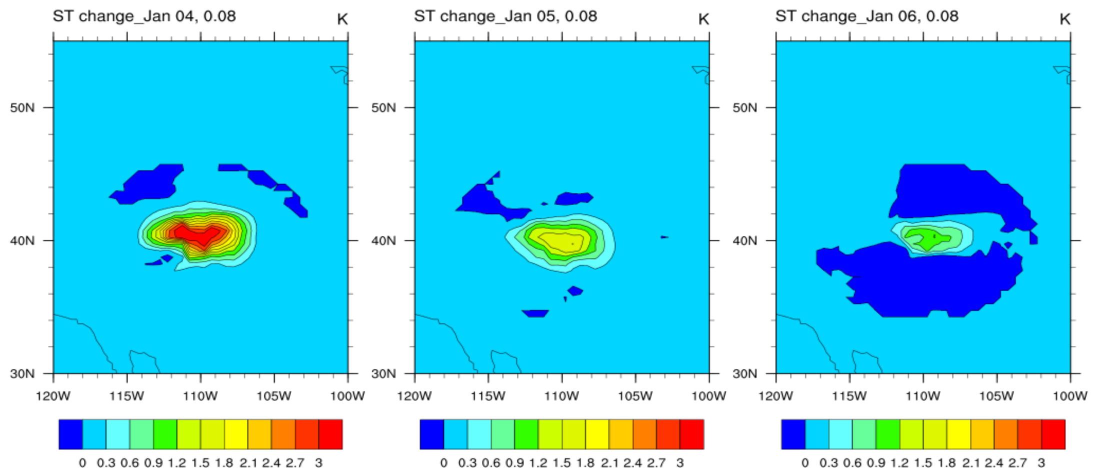
- Time: starts from Jan. 01. 2011 to Jan. 06. 2011, 5 DA cycles
- Land initialization: cold-start
- Atmospheric forcing: CPLHIST Forcing (CAM6, Raeder et al., 2021)
- Ensemble Adjustment Kalman Filter (EAKF)
- Soil temperature (lon: 249.77°, lat: 40.11°, depth: 0.08)
- Snow cover (lon: 89.95°, lat: 59.58°, depth: 0.00)
- Leaf area index (lon: 300.16°, lat: 0.00°)



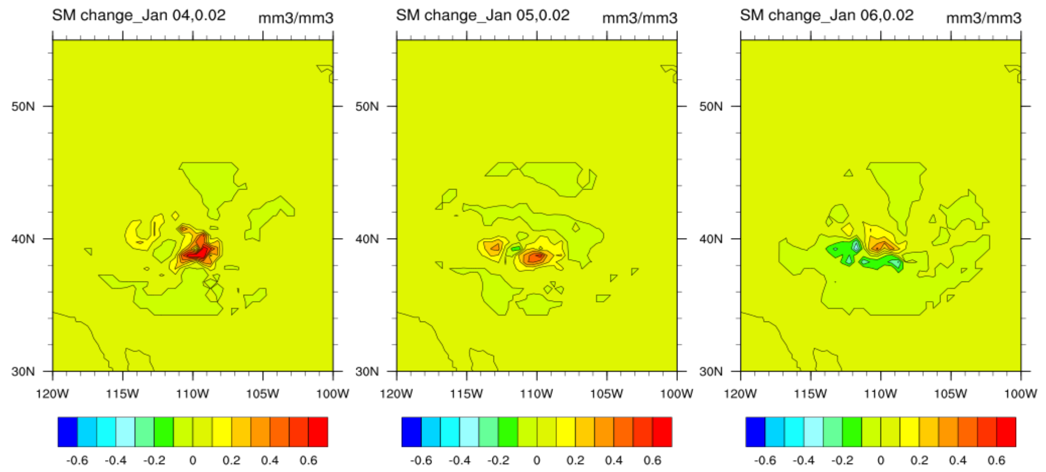
ELM-DART Single Obs. Experiments

The observation was successfully assimilated on January 4, 5, and 6.

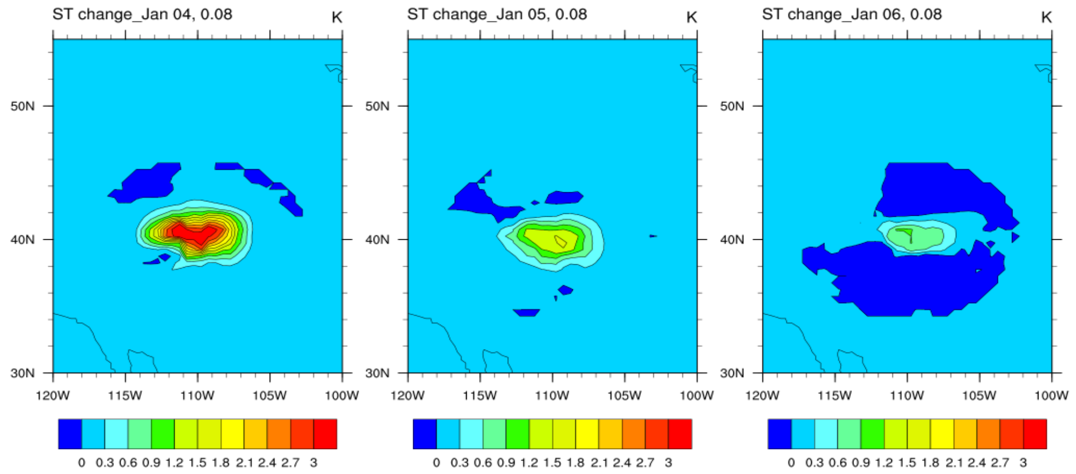
Soil Temperature Increment



Soil moisture Increment

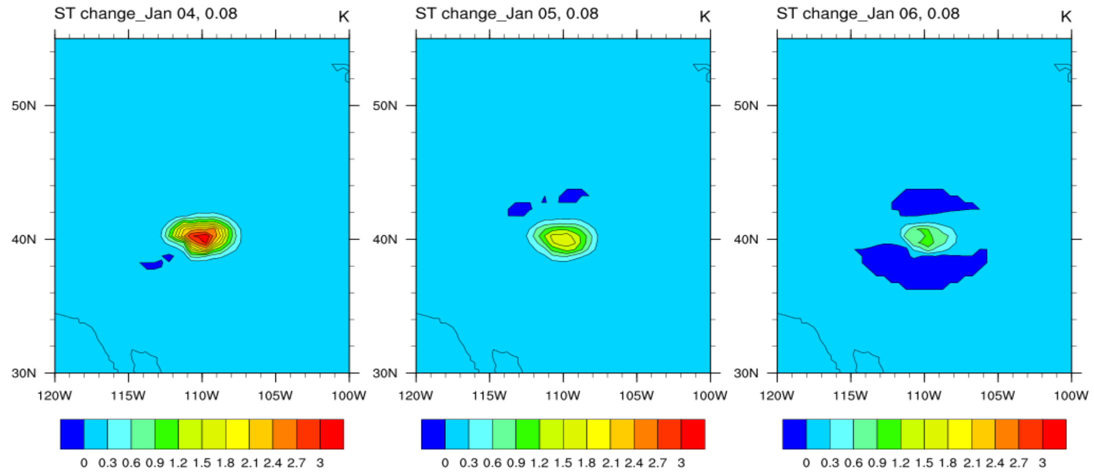


Soil Temperature Increment



Sensitivity to inflation

Soil Temperature Increment

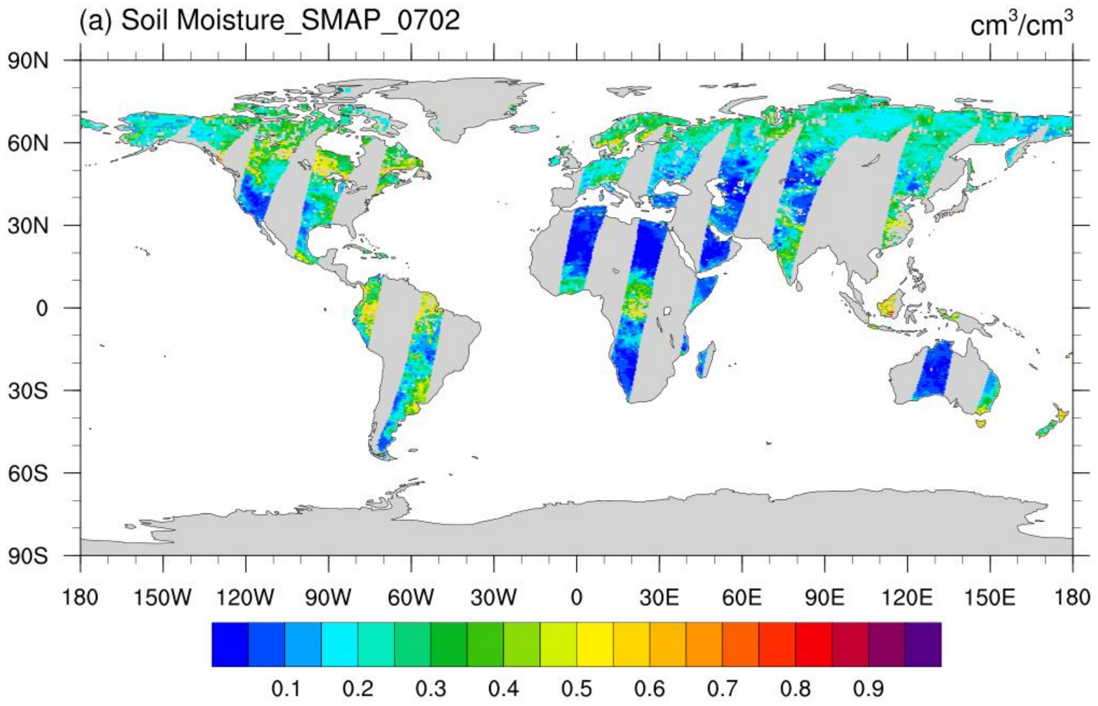


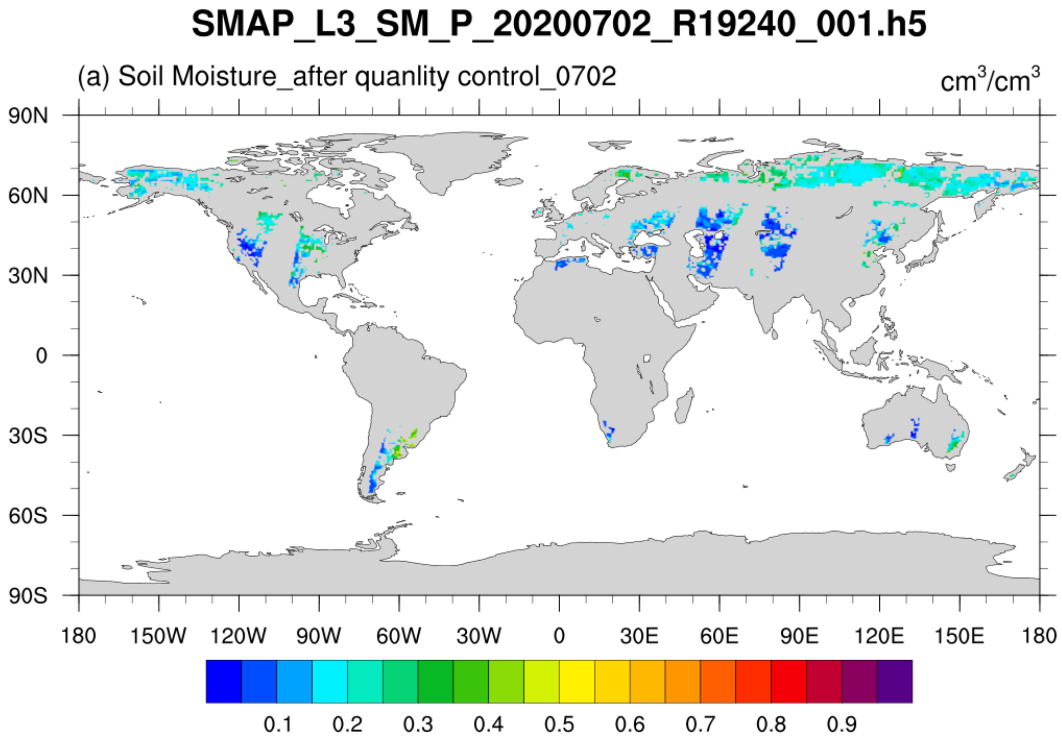
Sensitivity to Localization

ELM-DART Land Data Assimilation with SMAP Satellite Data

SMAP soil moisture data: L3_SMAP: a composite of daily estimates of global land surface conditions retrieved by the Soil Moisture Active Passive (SMAP) passive microwave radiometer. Global, 36 km.

07/02/2020



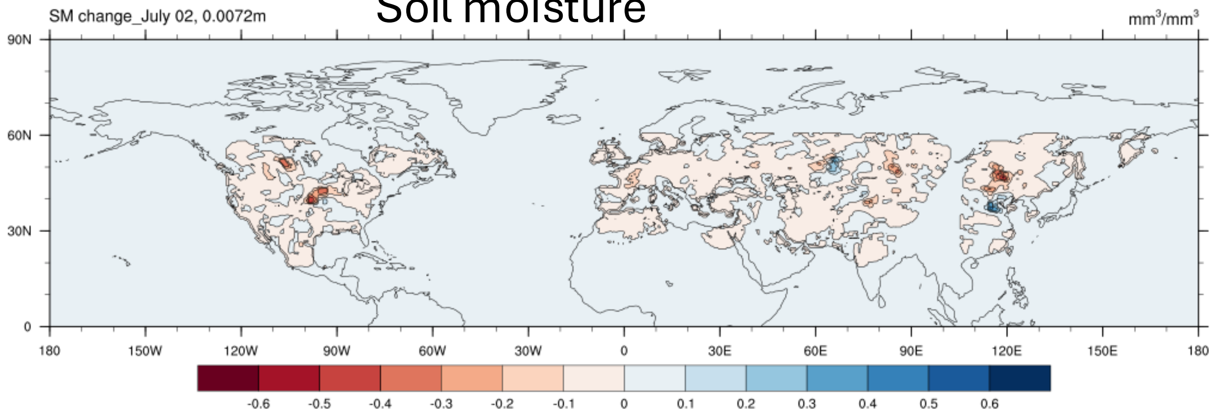


ELM-DART Land Data Assimilation with SMAP Satellite Data

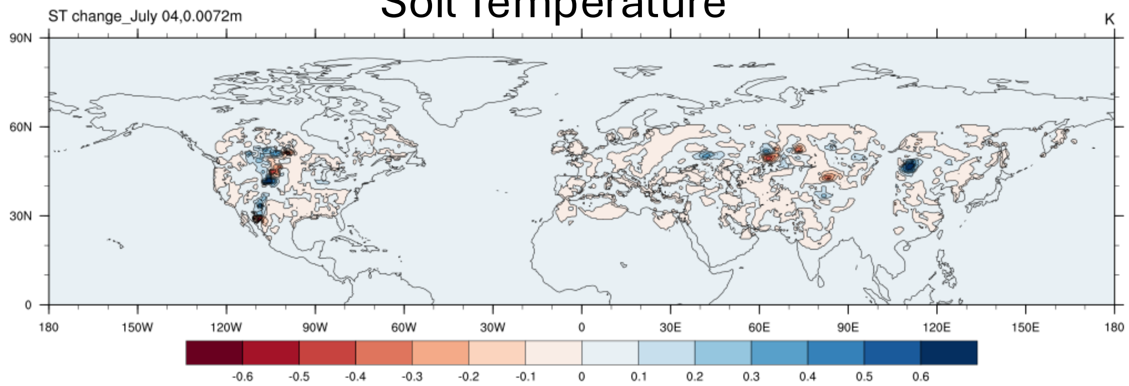
Analysis Increments 07/02/2020

Top soil layer

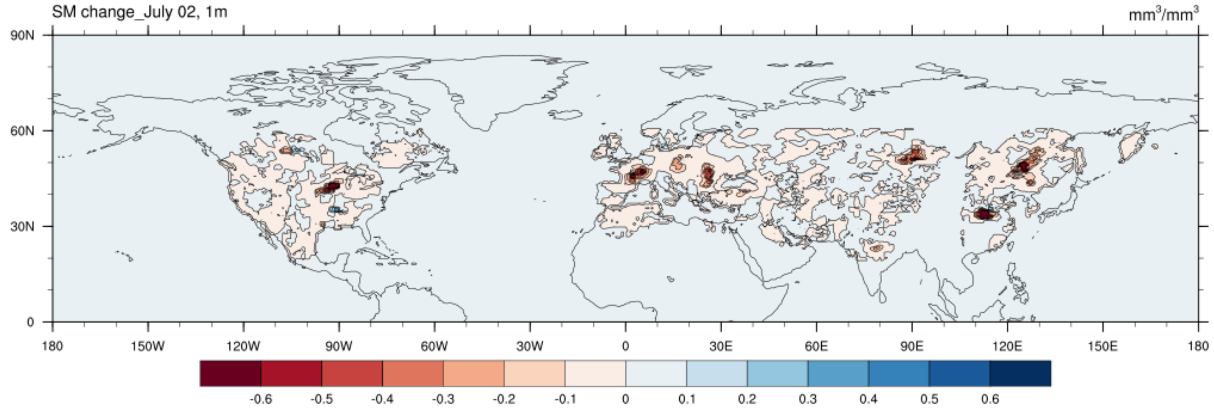
Soil moisture

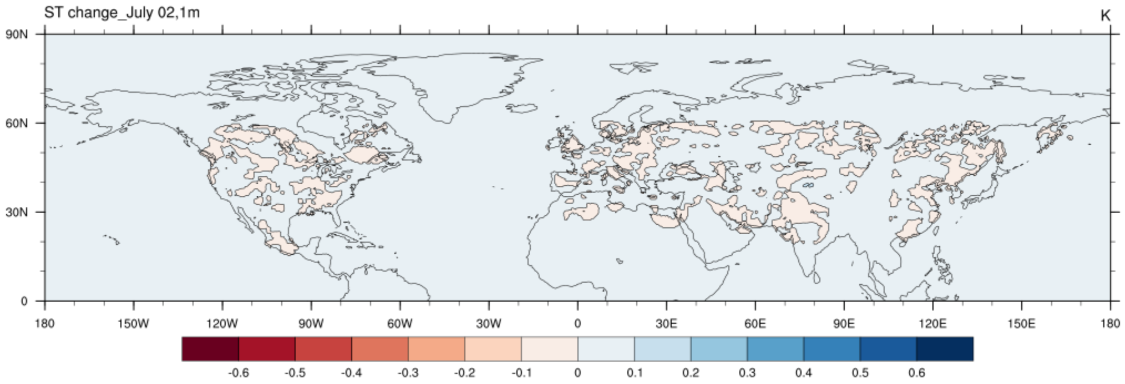


Soil Temperature



at 1 meter





Summary and Ongoing Work

Coupled Data Assimilation for E3SM with DART

- Coupled data assimilation is essential for accurate Earth System Model (ESM) prediction.
- **Team effort**: Developing a strongly coupled land–atmosphere data assimilation system for E3SM using DART.
- **Progress so far**:
  - **Atmospheric components of E3SM/DART DA** system (EAMv3-DART) shows benefits for S2S prediction, with positive feedbacks for land.
  - **Land DA system for ELM** (**ELM-DART**) has been built — assimilation experiments show reasonable results.
- **Ongoing and next steps**:
  - Advancing toward a **strongly coupled land–atmosphere DA system** to improve E3SM initialization and prediction.
  - Working toward the successful of the **project's overarching objectives**.

Thank you!

Zhaoxia.Pu@utah.edu