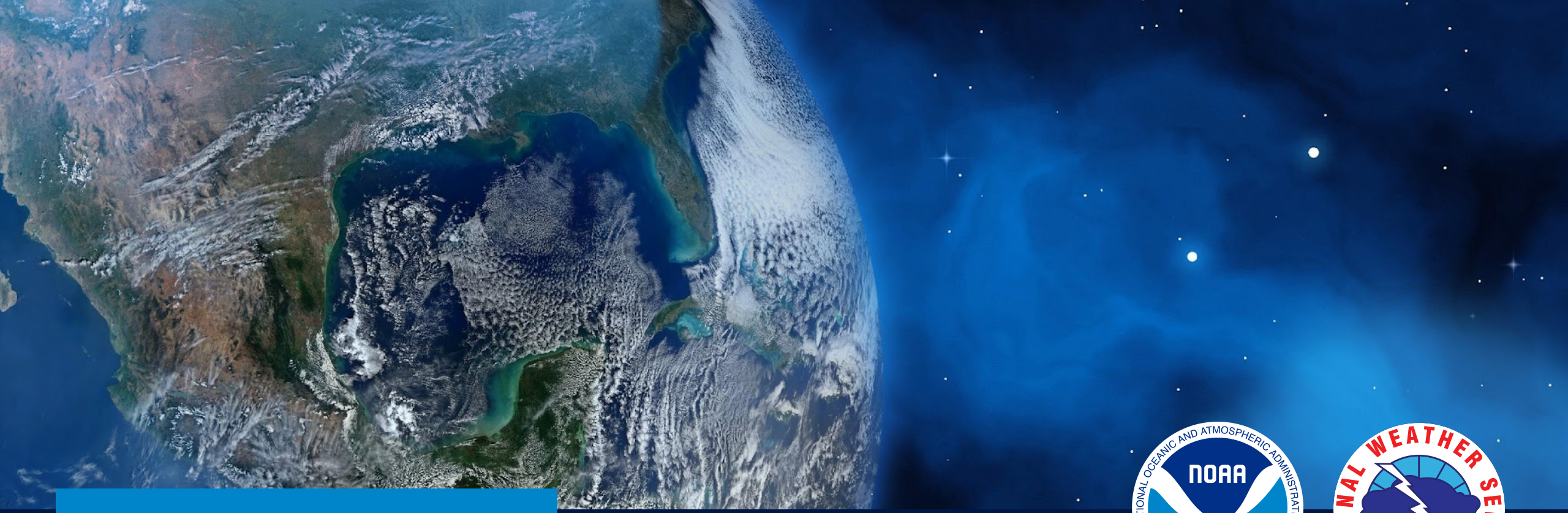


Overview of land data assimilation and land-atmosphere coupling options for NWP

Clara Draper
Mike Barlage, Jiarui Dong, Tseganeh Gichamo, Cory Martin,
Jeffrey Whitaker, Yuan Xue, and many others.
NOAA Office of Modeling and Development

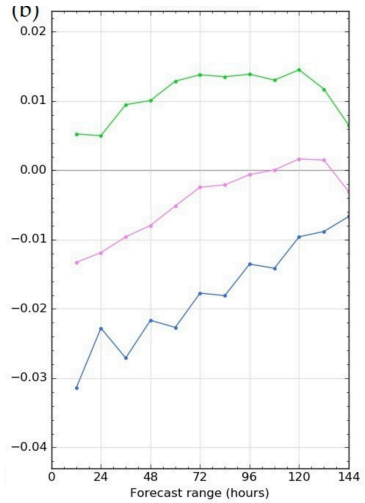
U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service

ECMWF Annual Seminar, 2026.

A (very) brief history of land DA

Land DA in NWP

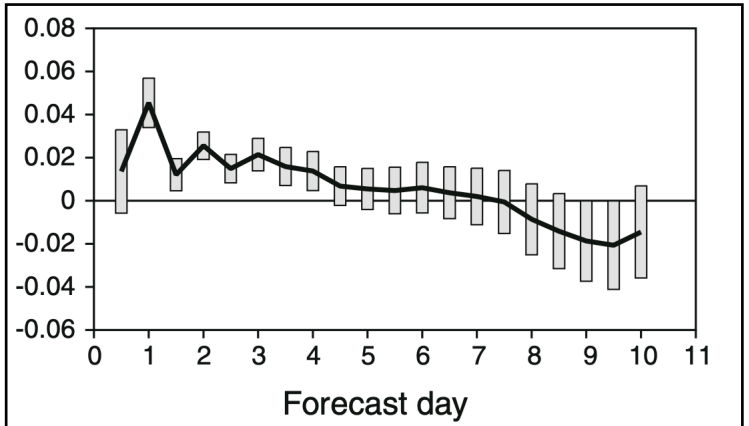
- Météo-France was the first to use land DA in an NWP system in 1985
  - Applied a simple scheme to update soil moisture and soil temperature from screen-level station observations
- Today global NWP systems update soil and snow states from a combination of conventional and satellites observations
- Some centers use observations of land forcing fields (precipitation, radiation)
  - Better suited to non-NRT systems
- Land DA useful for improve initialization of land states, leading to improved atmospheric forecasts; and also for detecting and evaluation model errors



Change in forecast T2m RMSE [K] at UKMO due to the SEKF soil moisture analysis (three different versions shown; Gomez et al, 2020).

Land DA in NWP

- Land DA in NWP typically includes some of:
  - Updating soil moisture, from
    - Screen-level T and q
    - From 2010's, satellite soil moisture (Tb/backscatter/retrievals)
  - Soil (and snow) temperature, from
    - Screen-level T
  - Snow amount, from
    - Satellite snow cover
    - Station snow depth



Improvement in normalized root mean square forecast 1000 hPa geopotential error [-] at ECMWF, from updating the snow depth analysis (de Rosnay et al, 2014).

- The land DA update is done separately from the atmospheric DA (weakly coupled)
- Relatively simple methods (Simplified Extended Kalman Filter, 2DVar, Optimal Interpolation)
- Different methods applied to different land variables
- For soil analysis, the horizontal and vertical components of the update are often separated

Land DA for Land-Only/Hydrology Models

- The land community has traditionally favored EnKF-type methods
  - Account for errors of the day, more robust to non-linearities, more intuitive specification of model errors
  - Var has not traditionally been used, although is gaining traction

- Greater use of land satellite observations than in NWP
  - Soil moisture analysis from satellite observations; screen-level observations less relevant

Land-Only Example: NASA Hydroglobe

- EnKF assimilation of:
  - Satellite soil moisture (ESA CCI)
  - Terrestrial Water Storage (GRACE, GRACE-FO)
  - Snow Water Equivalent (MODIS)
  - Leaf Area Index (MODIS)
- Forced by reanalysis (MERRA-2), bias corrected with:
  - Precipitation (GPM)
  - Radiation (CERES)
- Ensemble generated by perturbing atmospheric forcing, and model state variables
- 1D-EnKF (assimilated obs first projected onto model grid, the assimilated with no horizontal error correlations)

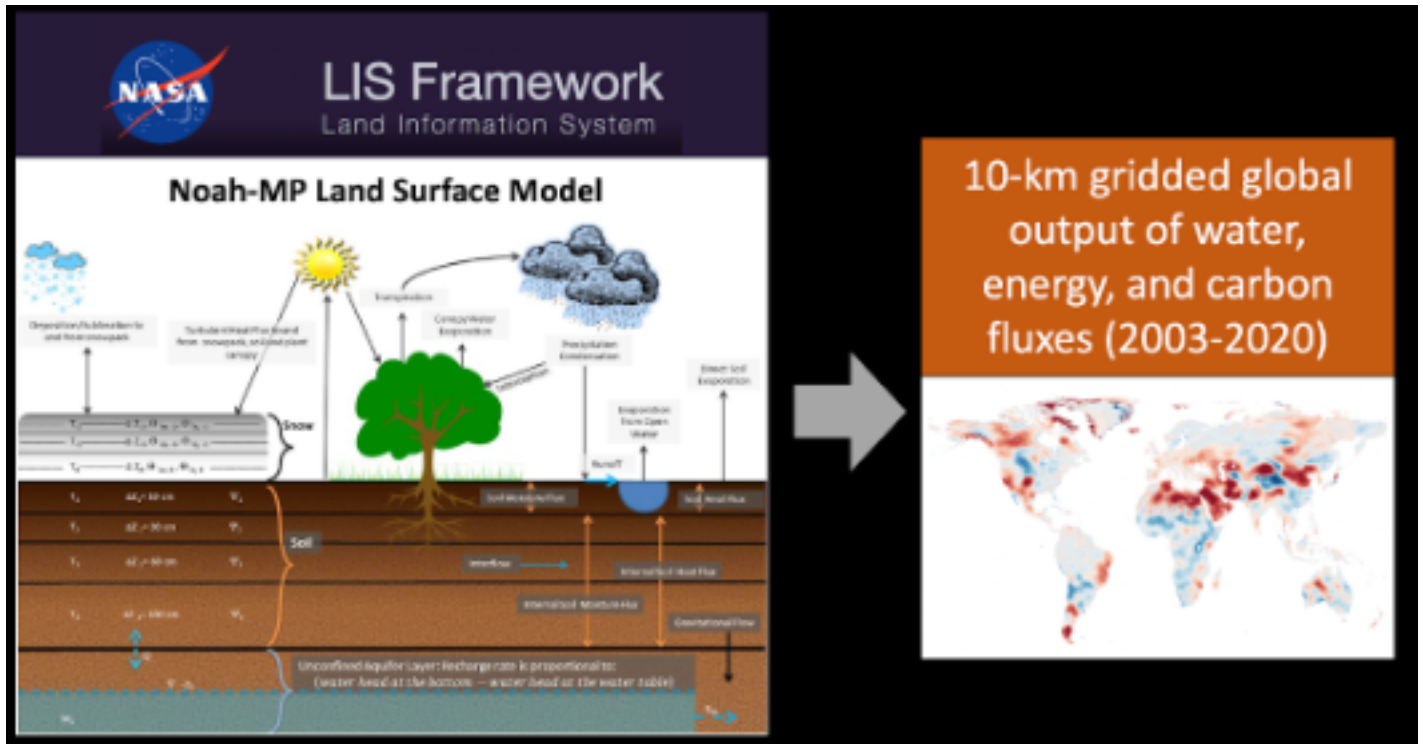
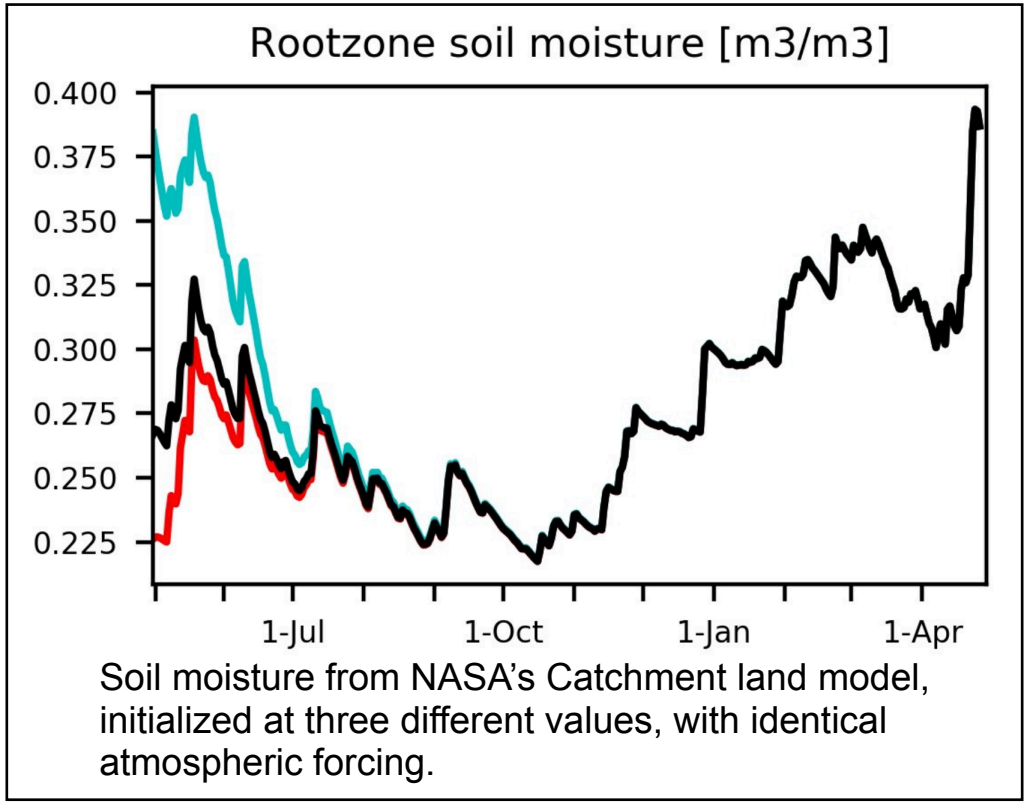


Image from Sujay Kumar

Differences Between Land and Atmosphere Modeling

- The land is strongly-forced (dissipative); over time it will converge to state determined by its forcing
  - No sensitive dependence on initial conditions (not chaotic!)
- Land surface models typically do not simulate horizontal flow between grid cells
  - No direct horizontal flow of errors
- Land physics are highly non-linear, include many switches
  - No model adjoints
- The land is highly heterogenous
  - Comparison of models and observations is difficult, due to differences in spatial support / difficulty extrapolating
  - Re-gridding can introduce systematic errors
- Time scales of land variables can be much longer than the atmosphere



Soil moisture from NASA's Catchment land model, initialized at three different values, with identical atmospheric forcing.

Differences Between Land and Atmospheric Data Assimilation in NWP

- Land DA is weakly coupled, and uses much simpler methods than atmospheric DA
- Done for historical reasons
  - Atmospheric DA methods not necessarily appropriate for land analysis
  - Computationally more affordable
    - (vertical) soil analysis done independently at each grid cell
  - Allows doing the land DA on the model grid (at the time, atmos DA was relatively course)

- The observations used in the land DA (snow depth, snow cover, screen-level T,q, satellite soil moisture) are not necessarily used in the atmospheric DA
  - Screen-level observations have been gradually introduced into atmospheric DA (after first being used for land DA)

Unifying Land and Atmosphere DA

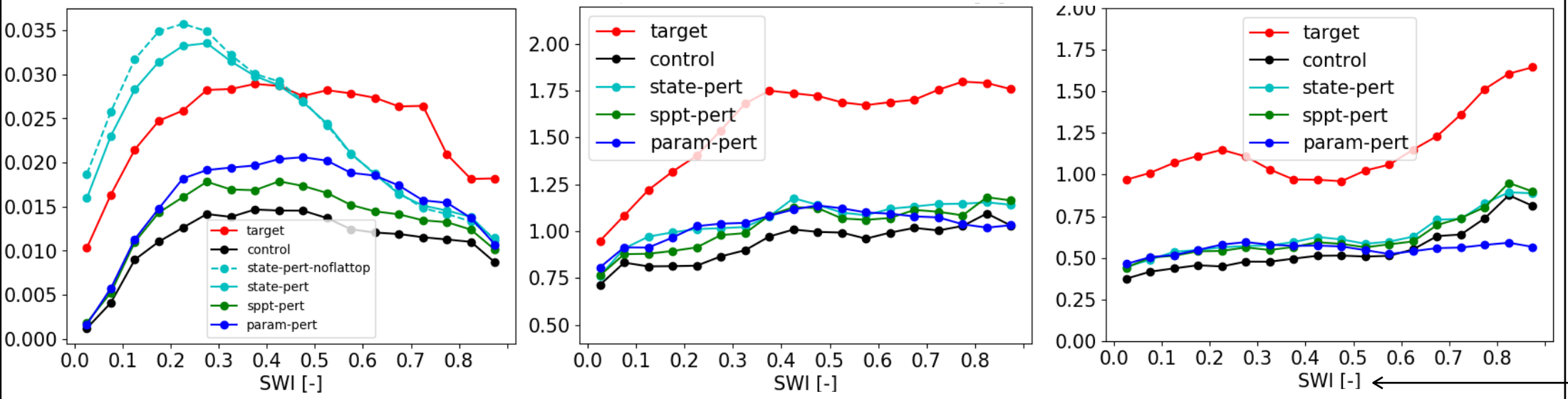
- Ensemble-based DA methods favored by the land community
  - Atmospheric DA has used ensemble-based methods for some time
    - e.g., NOAA's Hybrid 4D-EnVar
  - Not traditionally been used in NWP, but are being introduced
    - ECCC regional NWP system and NASA (LIS/AWFA; GMAO) use weakly coupled land EnKF
      - Each based on land-only ensembles, similar to Hydroglobe

- Current opportunity to better unify land and atmospheric DA systems
  - Use information from the atmospheric DA in the land DA
    - Use coupled error covariances, or h(x) for interface obs
  - Use atmospheric DA system/method to also do land DA
    - Introduces option of strongly coupled DA

Using NWP Ensembles in Land DA

NWP Ensemble Spread for Land DA

Model error stdev, estimated from NOAA's operational ensemble (black), and independent estimates (red)



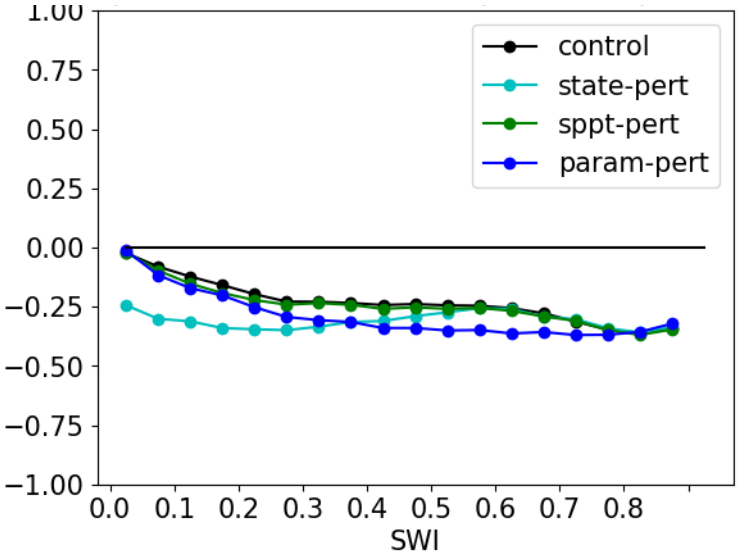
Figures from Draper 2021 (JHM)

Soil Wetness Index = Soil moisture, scaled between dry (0) and wet (1) limits.

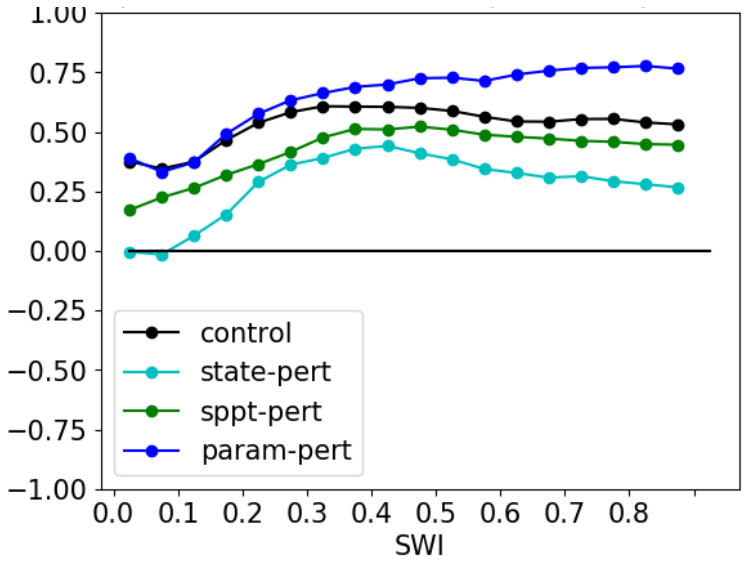
- NWP ensembles are under-dispersed at/near the land surface (known issue), but do have some spread for use in land DA
- ASIDE: Figures show experiments to induce more spread. Aqua line was generated by adding perturbations to soil moisture at each time step; ensemble spread represents model memory of those perturbations (very different to atmosphere)

NWP Ensemble Correlations for Land DA

GFS daytime correlation
(soil moisture layer 1, T2m)



GFS daytime correlation
(soil temperature layer 1, T2m)



Figures from
Draper 2021 (JHM)

Left: Ensemble vertical correlations, estimated from NOAA's operational ensemble (black), and various perturbation experiments.

- NWP ensembles have sufficient cross-correlation to support coupled DA
- ASIDE: Adding perturbations to soil moisture (aqua line) increases correlations where land is driving the coupling, and reduces it where the atmosphere is driving the coupling; consequences for coupled DA?

Summary of Ensemble Experiments

- Can we use NWP ensemble to estimate covariances for land DA?
  - Yes, although ensembles are under-dispersed, and this will be difficult/ impractical to address
  - Spatial patterns more realistic than many land-only estimates

- Experiments with ensemble perturbation methods highlight differences in the error propagation between land and atmosphere

Using an Atmospheric EnKF to Update Soil States: Comparison of Weakly and Strongly Coupled Land/Atmosphere DA

The GSI Hybrid 4DEnVar

- NOAA's global NWP uses a Hybrid 4DEnVar
  - Deterministic member updated using Var, with hybrid B
  - Ensemble members updated using EnKF, then re-centered on the deterministic member
- Dual resolution: the ensemble and atmos DA are coarser resolution than the deterministic member
- Experiments here use only the EnKF (no deterministic member)

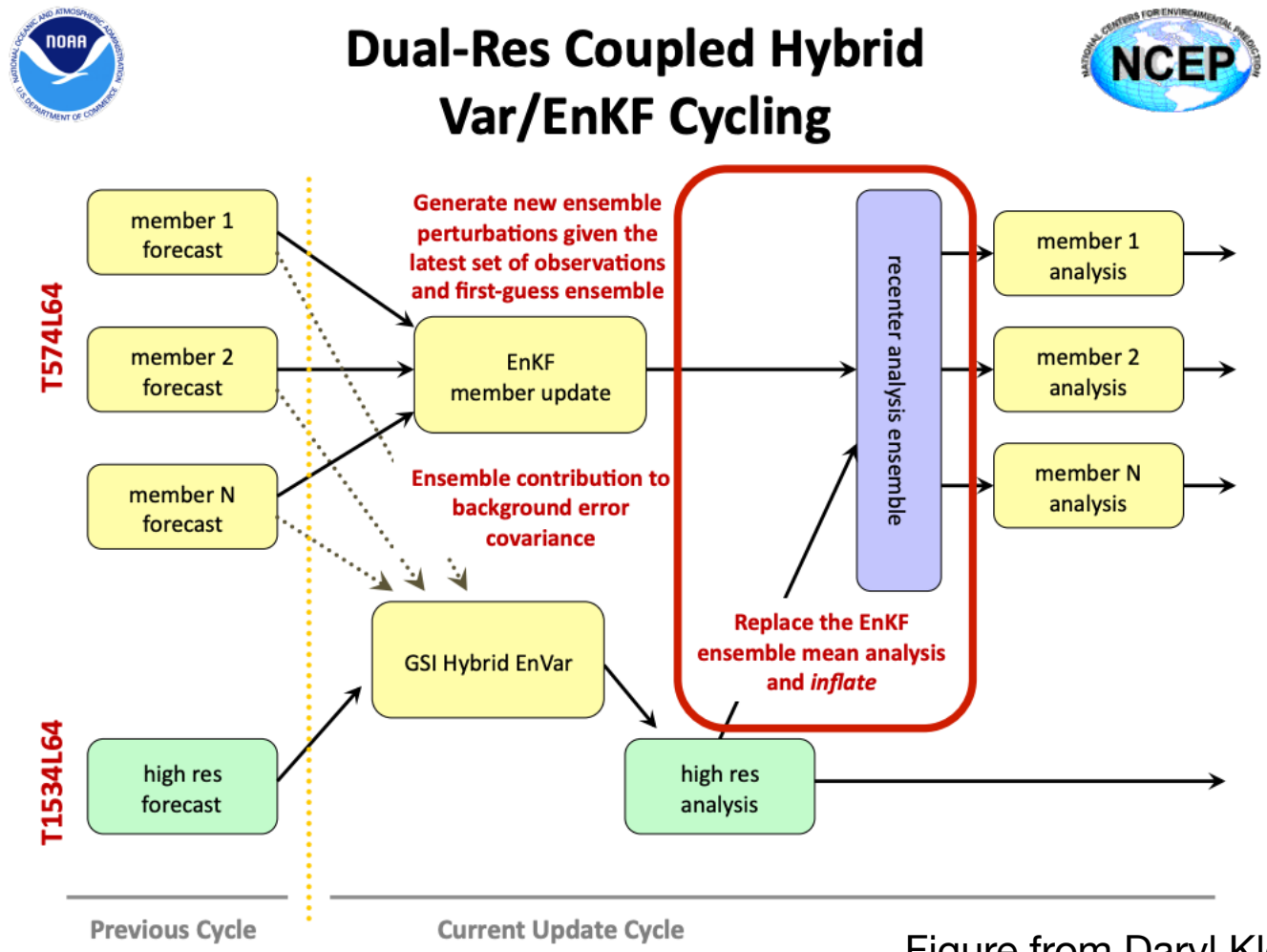


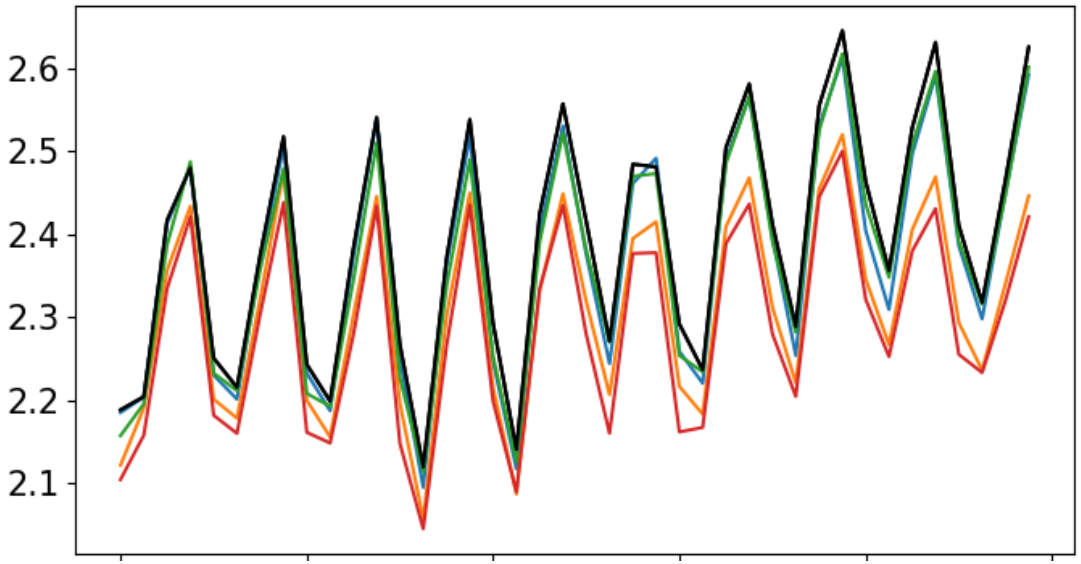
Figure from Daryl Kleist.

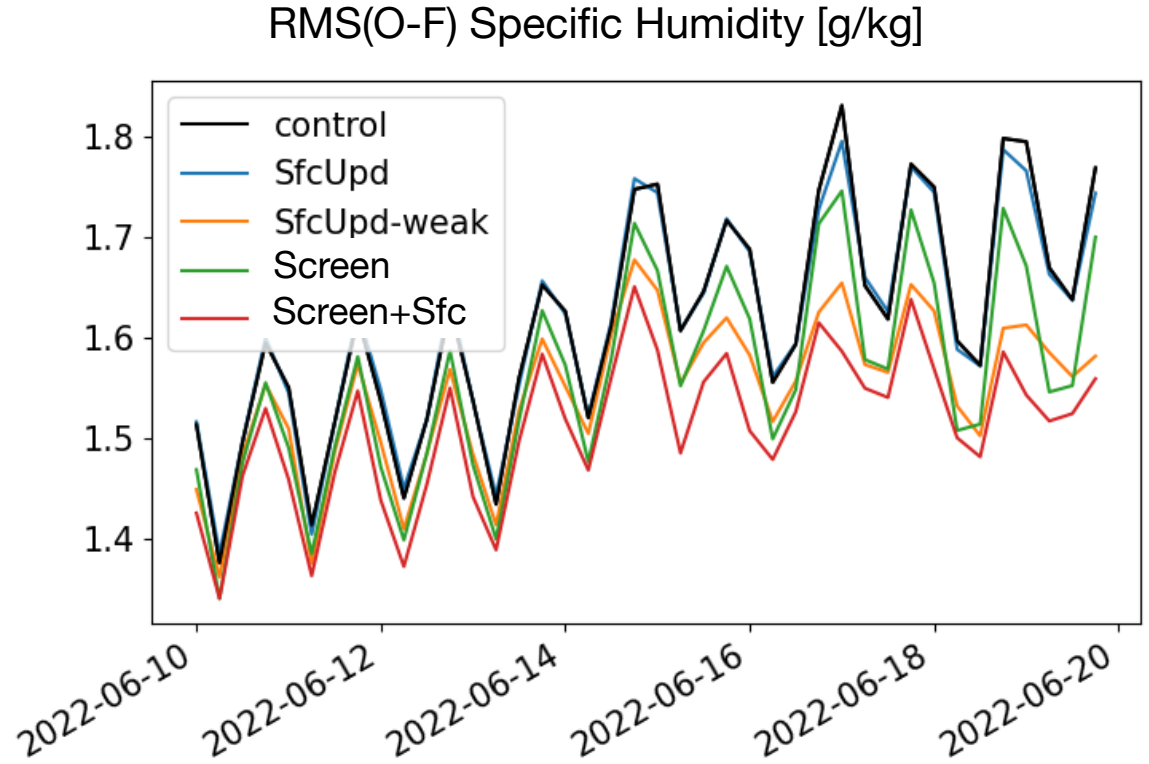
Coupled Soil/Atmosphere EnKF Experiments

- Conducted series of experiments to test coupled soil / atmosphere EnKF
  - Using the EnKF component of the GSI (for land and atmosphere)
  - 30 day experiment from June 2022, at 0.5° using GFSv17 HR-1 tag (with Noah-MP)
  - Atmospheric observations: standard observations assimilated in ops
  - Land observations: screen-level T and q (not presently assimilated at NOAA)

	Update atmosphere		Update soil moisture+temperature	
	from standard atmos obs	from screen-level obs	from standard atmos obs	from screen-level obs
Control	x			
Screen	x	x		
SfcUpd	x		x	
Screen+SfcUpd	x	x	x	x
SfcUpd-Weak	x			x

Screen-Level O-F





Mean RMSE

From Draper et al 2026 (QJRMS)

	T [K]	q [g/kg]
Control	2.39	1.62
Screen	2.37	1.56
SfcUpd	2.37	1.61
Screen+SfcUpd	2.28	1.51
SfcUpd - weak*	2.31	1.55

shaded = sig. difference from Control

- All experiments improve the O-F
- Best results from Screen+SfcUpd (3-4% reduction)
- Followed by SfcUpd-weak (screen-level forecasts constrained more by updated surface than atmosphere)

Upper-Air O-F

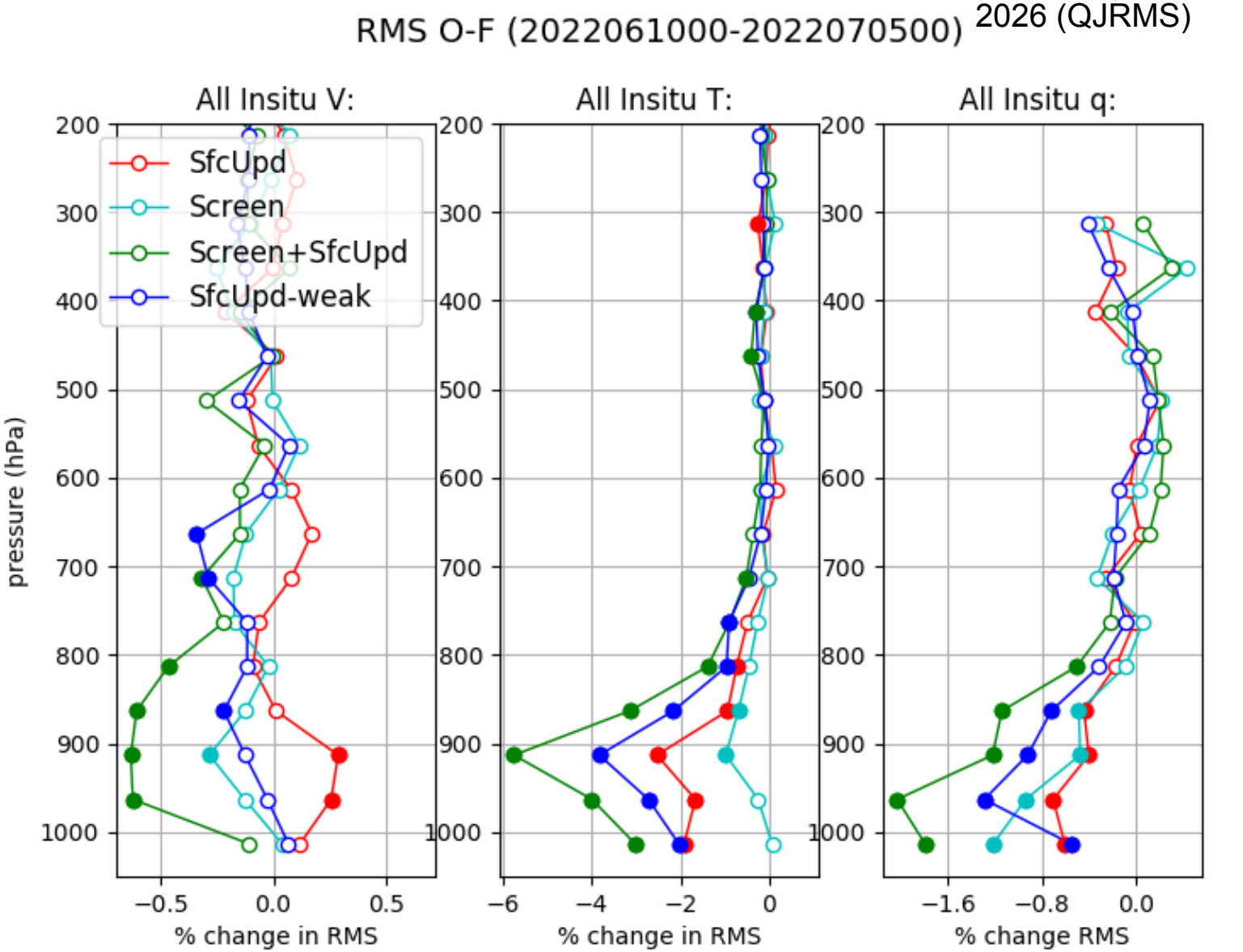
- Comparison against sondes and aircraft

Mean RMSE at ~900 hPA

	T [K]	q [g/kg]
Control	1.29	1.67
Screen	1.28	1.66
SfcUpd	1.27	1.65
Screen+SfcUpd	1.24	1.65
SfcUpd - weak*	1.25	1.66

- Best results from Screen+SfcUpd: Small, but consistent improvement (2-5%)
- SfcUpd-weak again next best

From Draper et al 2026 (QJRMS)



Full circle = significant difference from Control

Summary of Coupled DA Experiments

- Demonstrated potential benefit of:
  - Adding soil states to control vector (SfcUpd>Control; Screen+SfcUpd>Screen)
  - Adding assimilation of screen-level observations (Screen>Control; Screen+SfcUpd>SfcUpd)
  - More benefit from assimilating screen-level obs into surface (SfcUpd weak>Screen)
- In these experiments, strongly coupled DA out-performed weakly coupled DA
  - Weakly coupled experiment had much of the benefit of the strongly coupled

- Strongly coupled DA requires that the land and atmospheric DA have the exact same configuration
  - Error propagation for land and atmosphere are fundamentally different, and the best DA configuration for each may well be different
  - The performance gained by individually tuning each DA component in a weakly coupled system may outweigh the small performance advantage of strongly coupled DA seen here

Coupled Land/Atmosphere DA Progress at NOAA

Land DA in Global NWP at NOAA

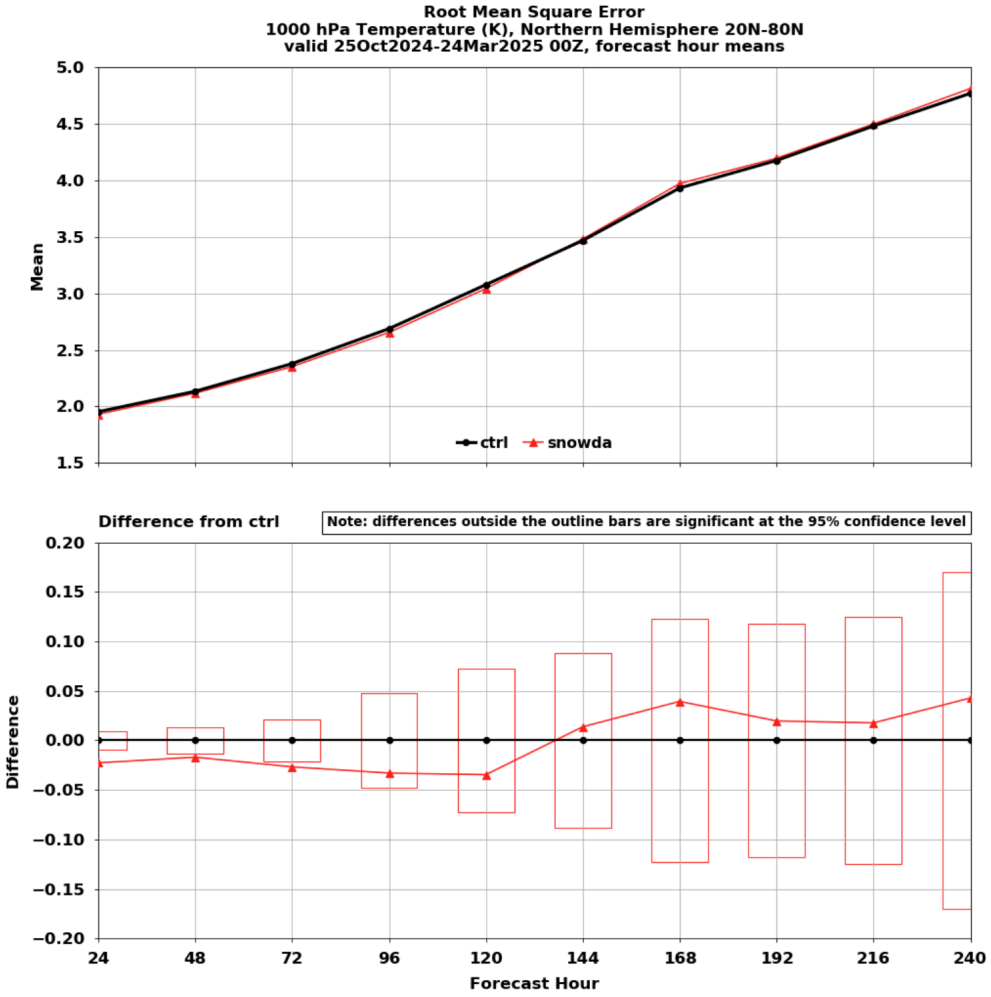
- NOAA's current global NWP system, GFSv16, includes:
  - Snow depth: Simple merging of a gridded snow analysis that has been QC'ed with IMS snow cover
  - Soil moisture and soil temperature: retrospectively corrected with observed precipitation

- For our pending update, GFSv17, we are more directly assimilating the observations, and moving towards unifying the land and atmospheric DA
  - Snow depth: 2DVar assimilation of station snow depth and IMS snow cover
  - Soil moisture and soil temperature: strongly coupled EnKF and addition of screen-level T and q

GFSv17 Snow Depth Analysis

- 2D-Variational assimilation of station snow depth and IMS satellite snow cover
  - Implemented in JEDI
  - DA algorithm based on Optimal Interpolation used elsewhere (Canada, ECMWF); model error covariances account for horizontal and vertical distance

- Evaluation shows largest improvements occur later in the season, associated with improved melt-off.

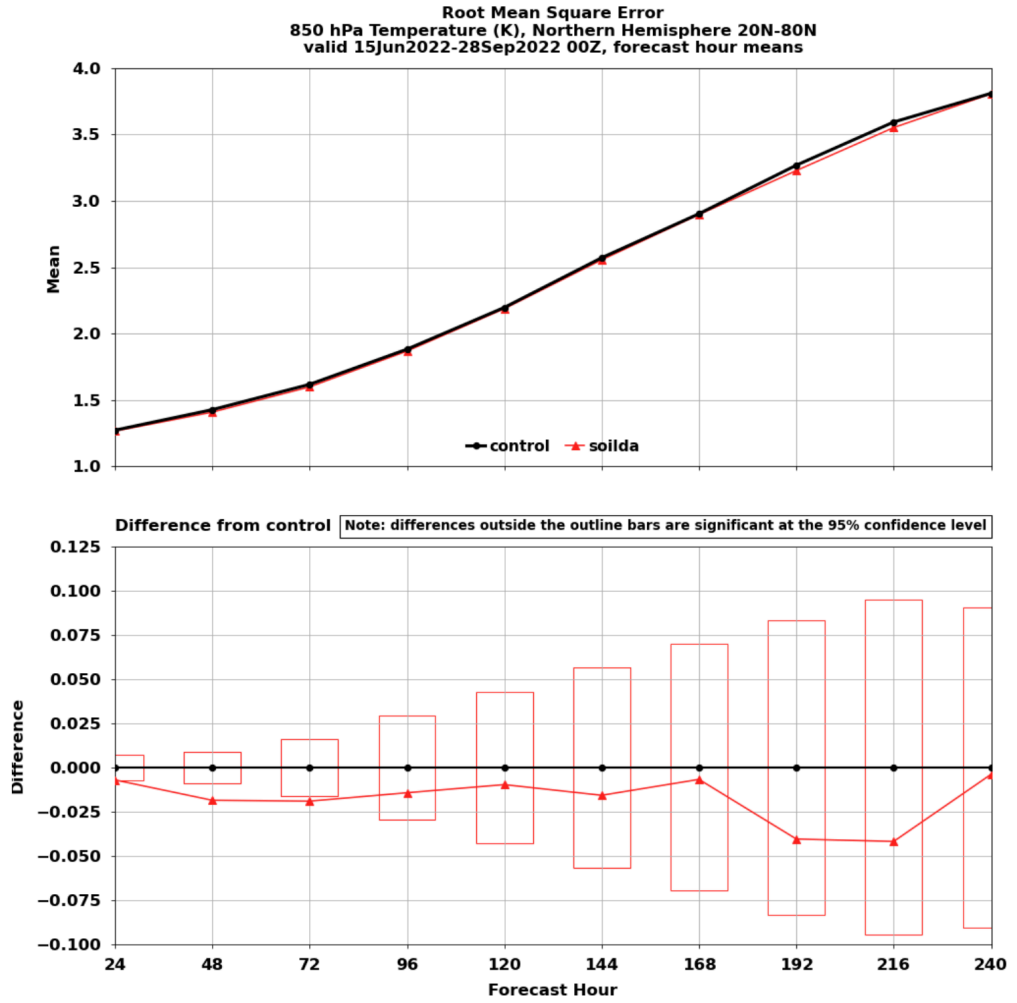


Change in 1000 hPA Temperature RMSE over Boreal winter, from the new snow DA scheme.

GFSv17 Soils Analysis

- Strongly coupled EnKF assimilation, and addition of screen-level T and q
  - Expand the EnKF control vector to include soil moisture and soil temperature in the uppermost 50 cm (SfcUpd+Screen experiment)
  - Do not (yet?) have a variational update for soil states, instead add mean ensemble increment to the deterministic member

- Recall, our atmospheric DA is dual-resolution
  - Resolution of atmospheric DA is coarser than the model.
  - Use of strongly coupled DA requires doing the land DA at a coarsened resolution, likely problematic due to the heterogeneity of the land surface

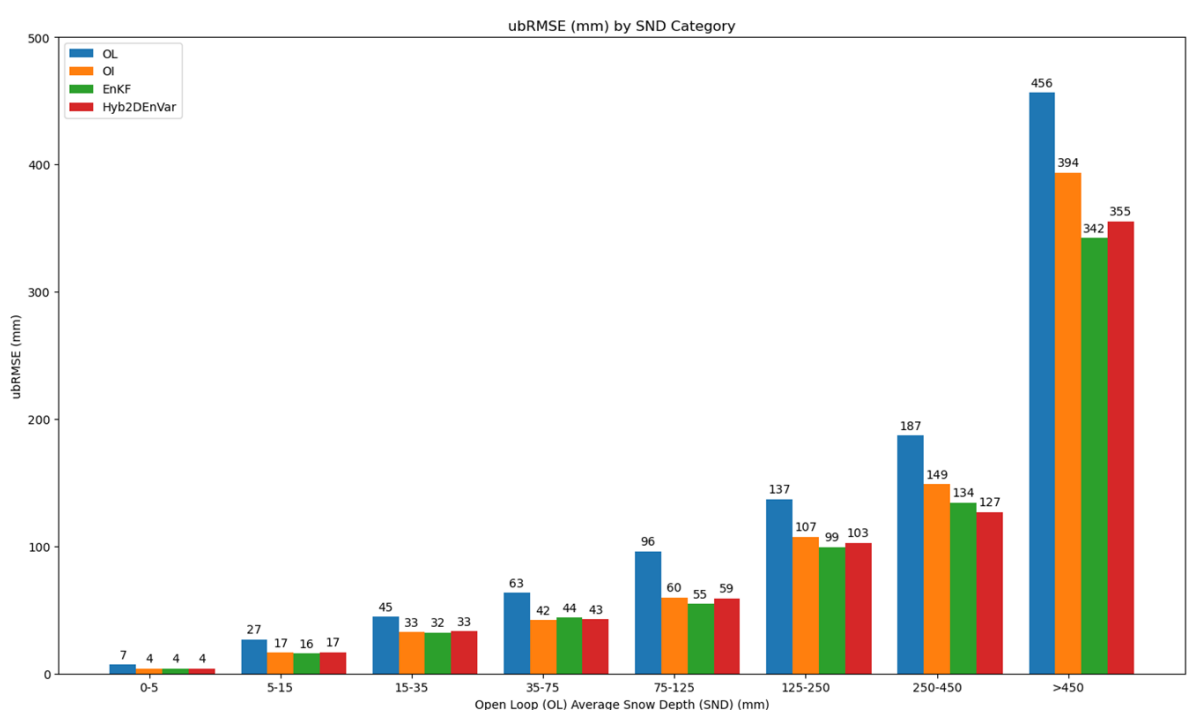


Change in 850 hPA Temperature RMSE over Boreal summer, from the new soil DA scheme.

Beyond GFSv17

Snow depth analysis (with Tseganeh Gichamo):

- Upgrade 2DVar to a Hybrid 2DEnVar
  - Ensemble covariances can better represent large variation in model errors (errors increase as snow depth itself increases)
  - Early results for assimilation of snow depth show Hybrid (with no tuning) matches EnKF performance, both outperform OI/2DVar
  - [to be tested] expect Hybrid to be better for assimilation of snow cover
    - EnKF problem with adding snow (all members no snow)



Snow depth ubRMSE [mm] for open-loop (blue) and different DA methods, binned by snow depth. Note: ensemble method (green; orange) outperform OI for large snow depth.

MAE (mm)

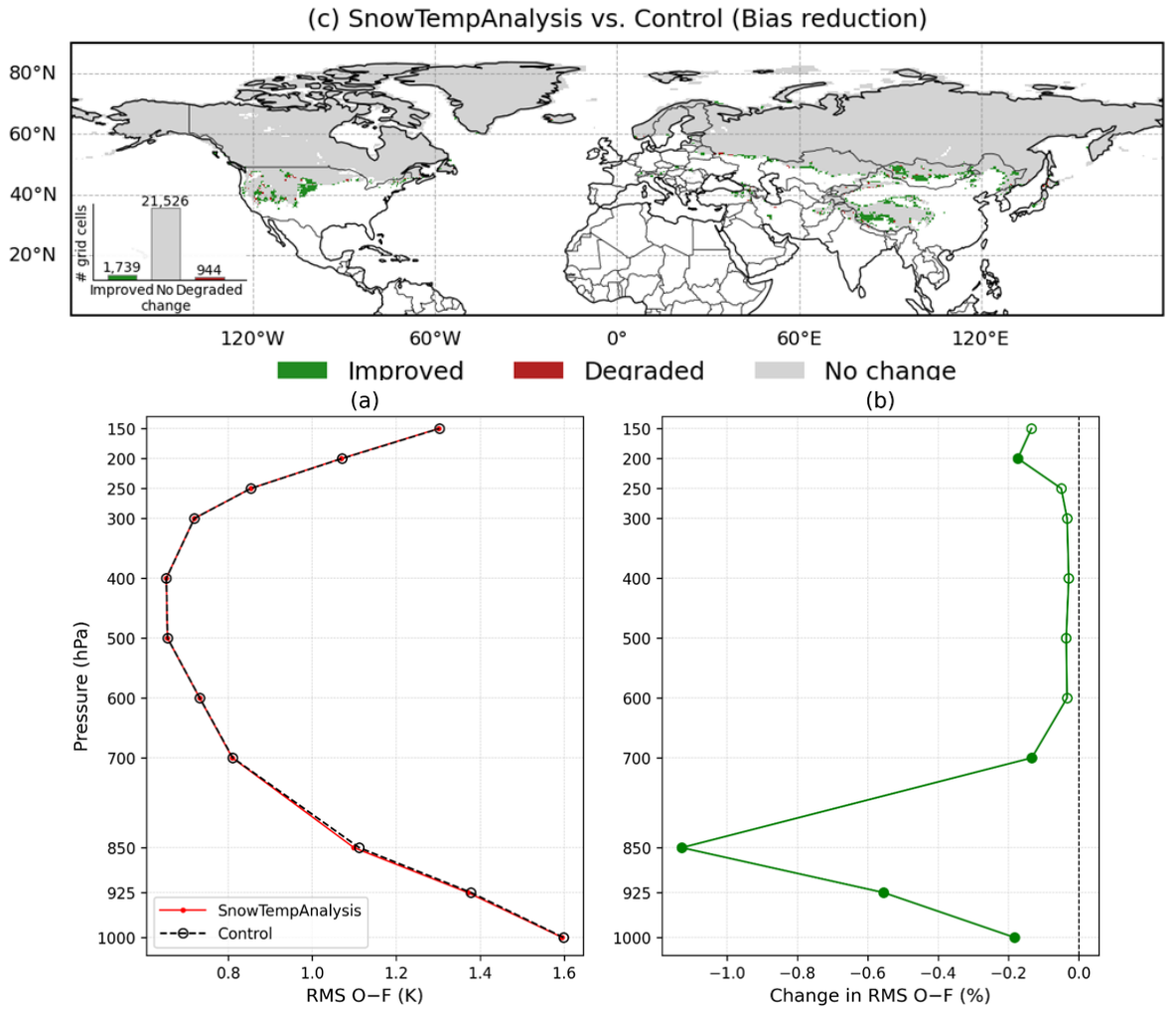
	OL	OI	EnKF	2DVar	JLETKF	Hyb2DEnVar
Mean	105.7	82.1	70.4	80.6	74.0	73.8

Mean Absolute Error over full experiment from each method.

Snow Temperature Analysis

Snow temperature analysis (with Yanjun Gan, NOAA PSL/CIRES):

- The GFSv17 soil analysis does not update soil states when snow is present
- Expand the strongly coupled soil / atmosphere EnKF control vector to also include snow temperature
- Preliminary work shows improvements in snow cover on the edge of the snow pack
  - Part of effort to improve snow cover and snow melt-off in GFS

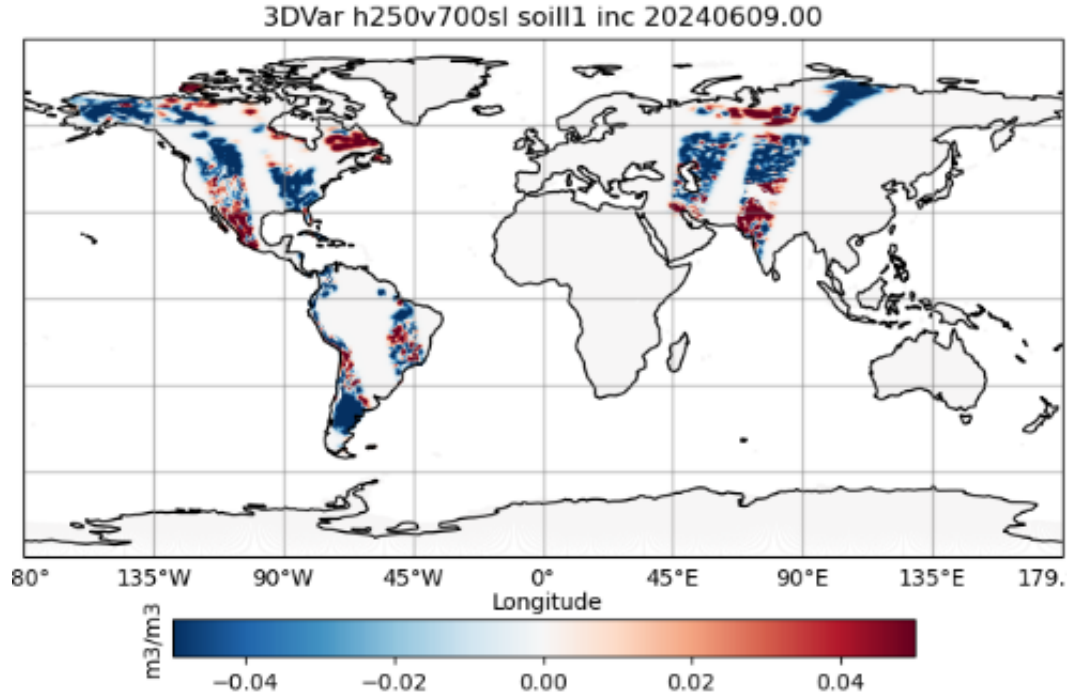


RMSE [K] in T against sondes and aircraft, for NH Boreal winter.

From Gan et al submitted (QJRMS)

Soil Analysis

- Transition to JEDI software package
- Parallel: developing a JEDI 3DVar soil analysis for use in deterministic spin-up/ experiments
- Better solution for calculating soil increment at the resolution of the deterministic member
- Addition of ensemble-recentering for soil (and snow?) states



Example increment to layer 1 soil moisture [m3/m3] from assimilation of SMAP soil moisture: test only, B matrix needs development.

Summary / Conclusions

Summary

- Land and atmosphere and fundamentally different, in ways that affect growth and propagation of errors
  - DA systems should be tailored to each
- Current land DA methods in NWP tend to be relatively simple
  - Partly for historical reasons (some simplifications are still appropriate!)
- Hydrology / land community tends to use more advanced methods; ensemble-DA methods with more use of satellite obs
- Atmospheric DA now uses ensemble-based methods that are more appropriate for land DA
- Opportunity to better unify land and atmosphere DA
  - Upgrade methods used in land DA to be more advanced / closer to those used in hydrology community
  - Make use of ensemble information from atmospheric DA
  - Stronger coupling between land and atmospheric DA
- Experiments here suggest strongly coupled land/atmos DA slightly out-performs weaker coupling
  - Strongly coupled DA requires applying same method/configuration to both components
  - Weakly coupled DA, with both components tuned to suit their characteristics may yield best results
- From a workflow perspective, strongly coupled is easier to implement and maintain

Summary

- At NOAA, moving towards more direct use of observations and unifying the land and atmospheric DA
- For GFSv17:
  - Soil depth: weakly coupled 2DVar
  - Soils: Strongly coupled EnKF

Some(what random) Concluding Thoughts

- Land is not chaotic, has no horizontal flow; ensemble response to applied perturbations (errors) depends on pattern of applied perturbations (errors) and how the model persists these perturbations (errors)
- Latter depends on local conditions, can be reasonably estimated from climatology. Do we even need ensemble-based data assimilation?
  - Yes! Need errors of the day to include atmospheric errors of the day (precip!). Atmospheric events like precip can dramatically shift the land/atmosphere coupling

- Are there any advantages / disadvantages to de-coupling the horizontal and vertical components of the land DA update?
- Discussed only ensemble DA here; are the conclusions the same/similar for ensembles of DA [Var,OI]?
- Still very few land satellite observations used in NWP, particularly compared to land-only systems

Thank you for listening

clara.draper@noaa.gov