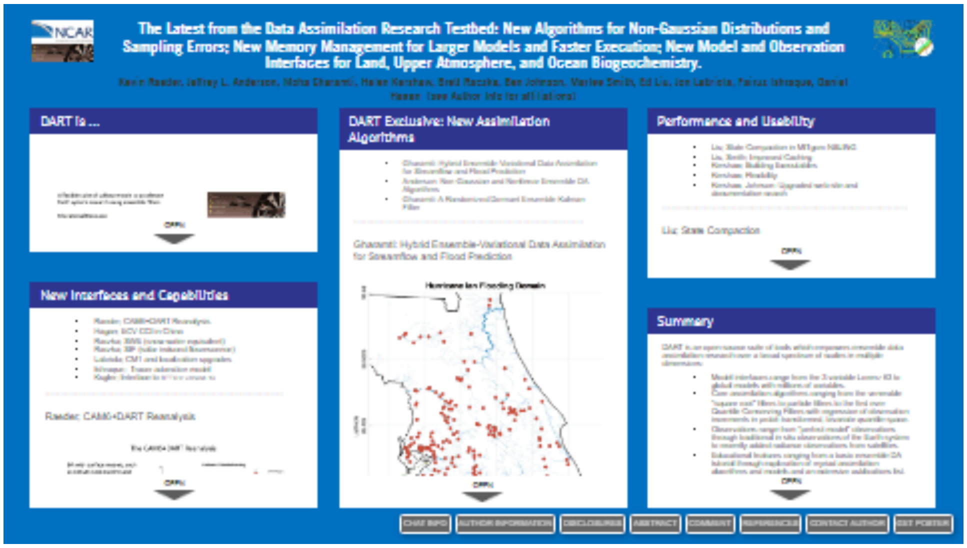
The Latest from the Data Assimilation Research Testbed: New Algorithms for Non-Gaussian Distributions and Sampling Errors; New Memory Management for Larger Models and Faster Execution; New Model and Observation Interfaces for Land, Upper Atmosphere, and Ocean Biogeochemistry.



The Latest from the Data Assimilation Research Testbed: New Algorithms for Non-Gaussian Distributions and Sampling Errors; New Memory Management for Larger Models and Faster Execution; New Model and Observation Interfaces for Land, Upper Atmosphere, and Ocean Biogeochemistry.

Kevin Raeder, Jeffrey L. Anderson, Moha Gharamti, Helen Kershaw, Brett Raczka, Ben Johnson, Marlee Smith, Ed Liu, Jon Labriola, Fairuz Ishraque, Daniel Hagan (see Author Info for affiliations)

DART Is ...

New Interfaces and Capabilities

- Raeder: CAM6+DART Reanalysis
- [illegible]
- Raczka: SWE (snow water equivalent)
- Raczka: SIF (solar induced fluorescence)
- Labriola: CM1 and localization upgrades
- Ishraque: Tracer advection model
- [illegible]

Raeder: CAM6+DART Reanalysis

DART Exclusive: New Assimilation Algorithms

- Gharamti: Hybrid Ensemble-Variational Data Assimilation for Streamflow and Flood Prediction
- Anderson: Non-Gaussian and Nonlinear Ensemble DA Algorithms
- Gharamti: [illegible] Ensemble Kalman Filter

Gharamti: Hybrid Ensemble-Variational Data Assimilation for Streamflow and Flood Prediction

Hurricane Ian Flooding Domain

Performance and Usability

- Liu: State Compaction [illegible]
- Liu, Smith: Improved Caching
- Kershaw: Building Executables
- Kershaw: Flexibility
- Kershaw, Johnson: Upgraded website and documentation search

Liu: State Compaction

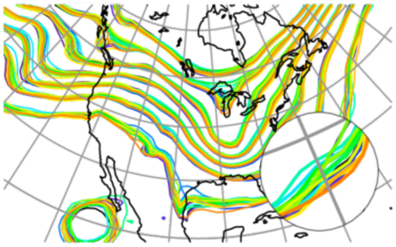
Summary

DART is an open-source suite of tools which empowers ensemble data assimilation research over a broad spectrum of scales in multiple dimensions.

- Model interfaces range from the 3-variable Lorenz 63 to global models with millions of variables.
- [illegible]
- [illegible]
- [illegible]

Kevin Raeder, Jeffrey L. Anderson, Moha Gharamti, Helen Kershaw, Brett Raczka,
Ben Johnson, Marlee Smith, Ed Liu, Jon Labriola, Fairuz Ishraque, Daniel Hagan
(see Author Info for affiliations)

ENTER NAMES OF AFFILIATED INSTITUTIONS



PRESENTED AT:

AGU FALL MEETING
Chicago, IL & Online Everywhere
12–16 December 2022
SCIENCE LEADS THE FUTURE

DART IS …

A flexible suite of software tools to accelerate Earth system research using ensemble filters

Educational Resource

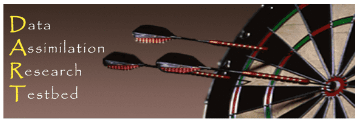
User community:

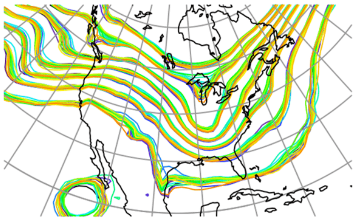
- 50+ Universities
- 100+ other sites
- 1500+ registered users

Open Source. DART team & community members develop:

- Model interfaces (e.g. CESM, WRF-Hydro, MPAS)
- Observation forward operators
- Assimilation algorithms
  (e.g. EnKF, RHF, quantile conserving and many more)

Contributions are reviewed, streamlined and tested before merging in public DART





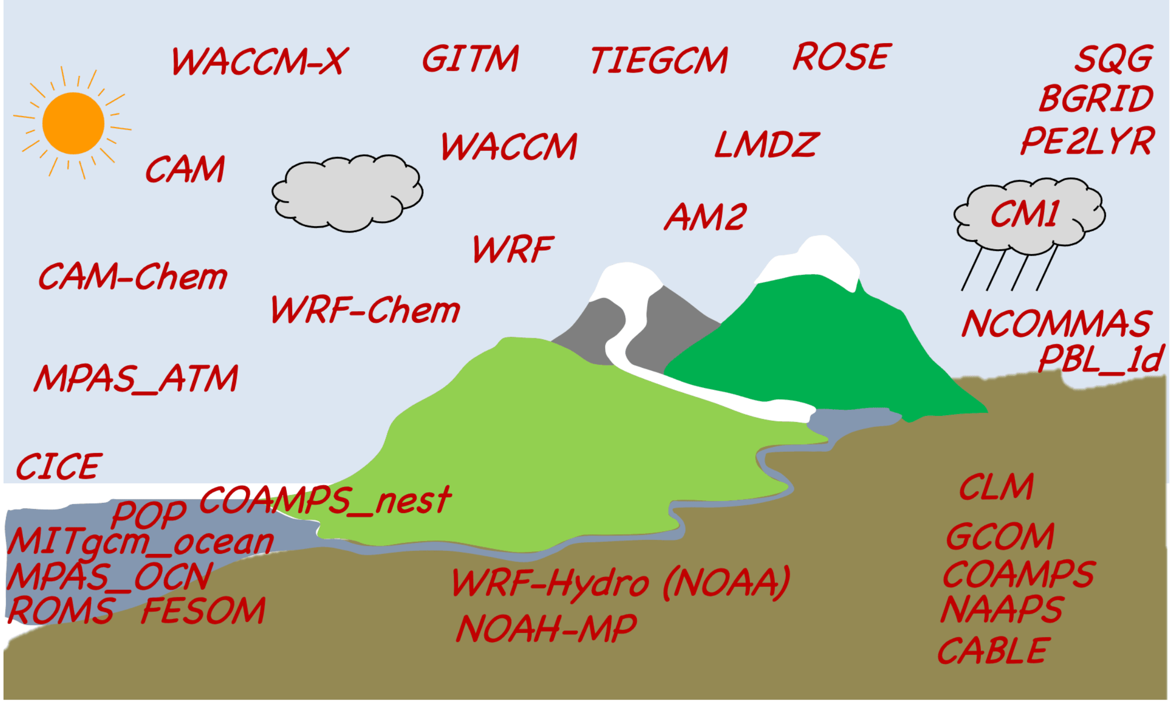
https://dart.ucar.edu

dart@ucar.edu

CISL
Computational & Information Systems Lab

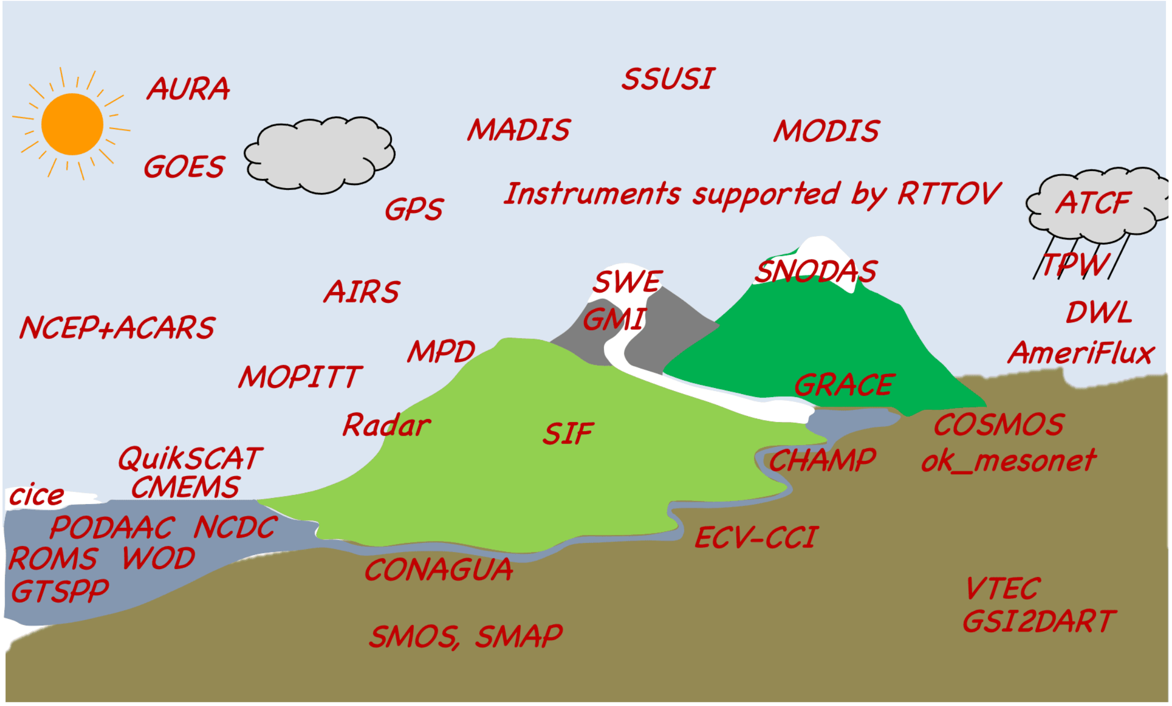
We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

Geophysical Models Interfaced to DART



DART also includes many low order models for rapid DA research and clear tutorials: Lorenz 63, Lorenz 96, simple advection (new; Ishraque 2022), ...

Earth System Observations (others available)



Assimilation algorithms

DART provides numerous core assimilation algorithms, from the traditional "square- root" filters, such as the Ensemble Adjustment Kalman Filter (EAKF), to the Rank Histogram filter, particle filters, and the new Quantile Conserving Filter (QCF, see Jeff Anderson's section in the middle column). There are numerous support algorithms to make ensemble data assimilation work well and efficiently in large (Earth system) models; e.g.,

- localization,
- ensemble inflation,
- sampling error correction,
- highly parallel computation,
- efficient interprocess communication, ...

To analyse the output of assimilations effectively there are postprocessing tools to examine it in state space (on the model grid), in observation space (at the observation locations) and from an ensemble statistical perspective. The ensembles can be used to explore the evolution of patterns of sensitivity of a quantity of interest to all of the model state variables.

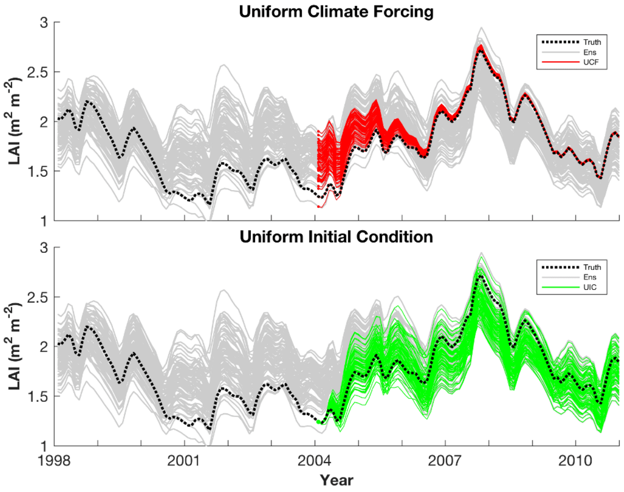
NEW INTERFACES AND CAPABILITIES

- Raeder; CAM6+DART Reanalysis
- Hagan; ECV-CCI in China
- Raczka; SWE (snow water equivalent)
- Raczka; SIF (solar induced flourescence)
- Labriola; CM1 and localization upgrades
- Ishraque; Tracer advection model
- Kugler; Interface to RTTOV version 13

Raeder; CAM6+DART Reanalysis

The CAM6+DART Reanalysis

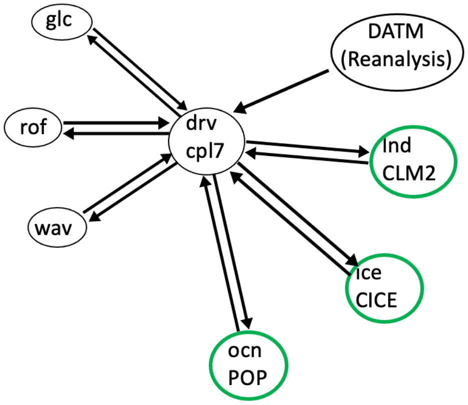
DA with surface models, such as CESM's Community Land Model, requires not only a good model, but good forcing from the atmosphere, both in the mean and ensemble spread.



In the top figure, an ensemble of CLM members having different initial conditions is forced by a single atmospheric evolution. The ensemble collapses (red), making it unusable for ensemble data assimilation. The bottom figure shows an ensemble whose members all have the same initial conditions, but each is forced by a different atmospheric evolution. The members (green) diverge into an ensemble with a useful spread.

Atmospheric forcing of surface components

CESM components



glc

DATM
(Reanalysis)

rof

drv
cpl7

lnd
CLM2

wav

ice
CICE

ocn
POP

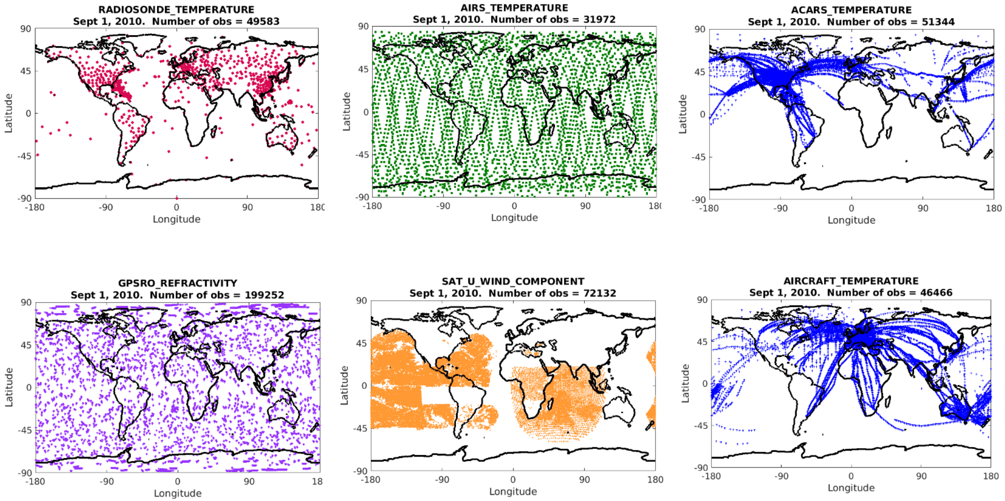
Surface models in CESM2
(CLM, POP, CICE, ...) are forced
by CAM6. DA using any of these
can use an existing CAM6 reanalysis
instead of re-running a CAM6 ensemble
for each new case.
Reanalysis ≅ actual atmosphere.

Forcing is stored in cpl history files:

- frequencies ranging from 1-6 hours
- ready to use in CESM in DATM mode
- 1 year, 1 member per file
- 2011-2020)

These models have DART interfaces for assimilation.

Assimilated Atmospheric Observations



RADIOSONDE_TEMPERATURE
Sept 1, 2010. Number of obs = 49583

AIRS_TEMPERATURE
Sept 1, 2010. Number of obs = 31972

ACARS_TEMPERATURE
Sept 1, 2010. Number of obs = 51344

GPSRO_REFRACTIVITY
Sept 1, 2010. Number of obs = 199252

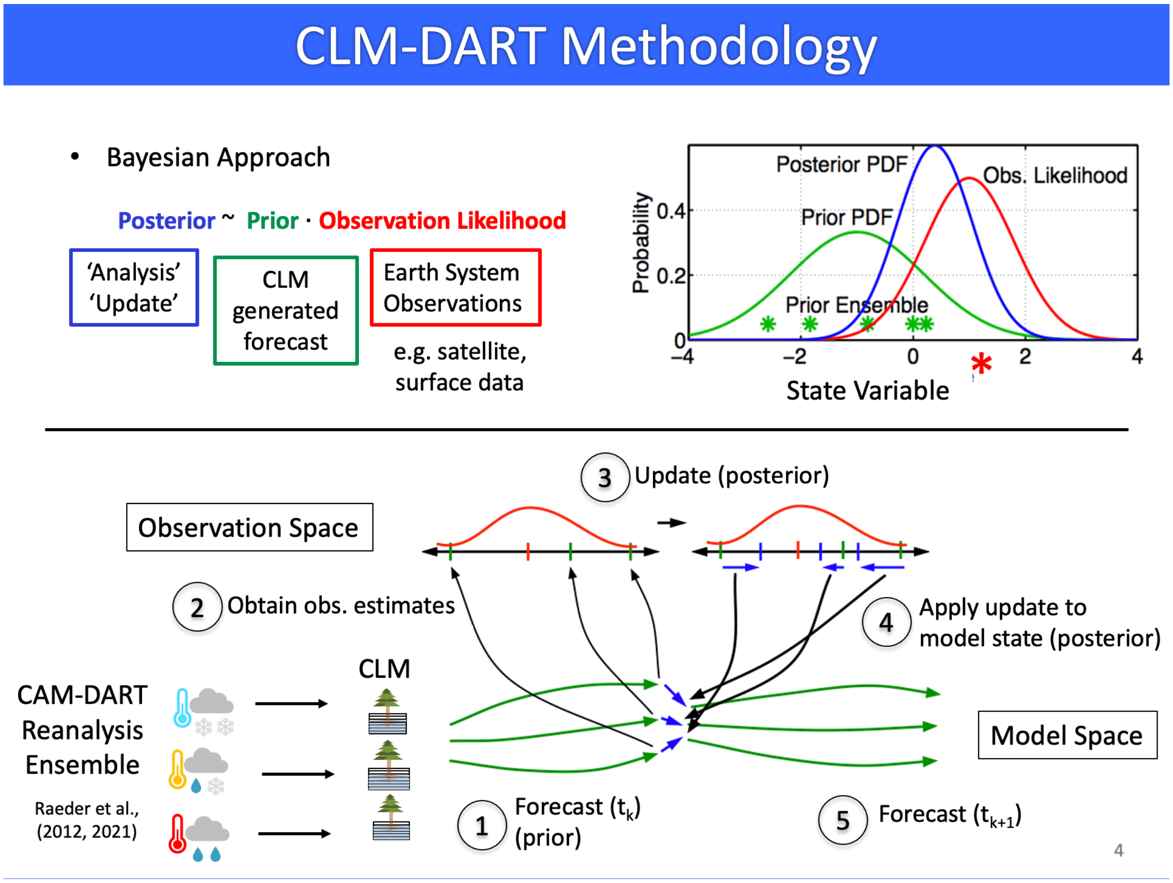
SAT_U_WIND_COMPONENT
Sept 1, 2010. Number of obs = 72132

AIRCRAFT_TEMPERATURE
Sept 1, 2010. Number of obs = 46466

Example of observations used in 1 cycle; > 450,000 in this window.

Reanalysis Data Products

- + The CAM6+DART Reanalysis can accelerate research using non-atmospheric Earth system models at lower cost.
- + It provides objectively derived, realistic variability and uncertainty estimates to surface models.
- + Described in detail in Scientific Reports: https://rdcu.be/ctUVQ
- + Freely available at https://rda.ucar.edu/datasets/ds345.0
- + Organized by CESM component (cpl, atm, esp, ...)
- + Useful units of compressed data for easy download
- + "Observation space" data; ensemble *model estimates* of the observations at the obs locations



CLM-DART Methodology

- Bayesian Approach

Posterior ~ **Prior** · **Observation Likelihood**

'Analysis'
'Update'

CLM generated forecast

Earth System Observations

e.g. satellite, surface data

3 Update (posterior)

Observation Space

2 Obtain obs. estimates

4 Apply update to model state (posterior)

CAM-DART Reanalysis Ensemble

Raeder et al., (2012, 2021)

CLM

Model Space

1 Forecast (t_k) (prior)

5 Forecast (t_{k+1})

4

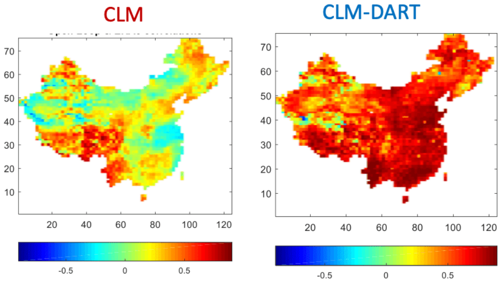
Hagan; ECV-CCI in China

Soil Moisture observations (CLM-DART)

CLM: CLM4.5 free run (no observations)
CLM_DART: CLM4.5 + ECV-CCI observations

- CLM-DART fills in gaps from ECV-CCI retrievals and improves surface correlation with ERA5 benchmark product

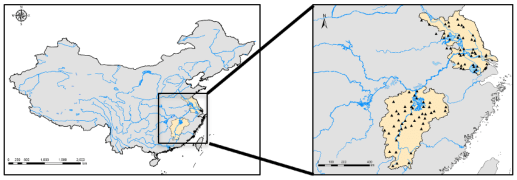
Correlation w/ ERA5 Near Surface Soil Moisture



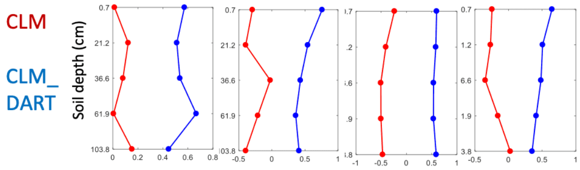
D. Hagan et al, (in prep)

- CLM-DART also improves subsurface soil moisture correlation with in-situ site observations

Site Level Sub-Surface
Soil Moisture Correlation (1-100 cm)

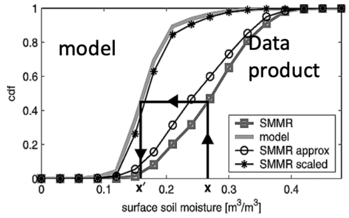


Jiangxi and Jiangsu provinces

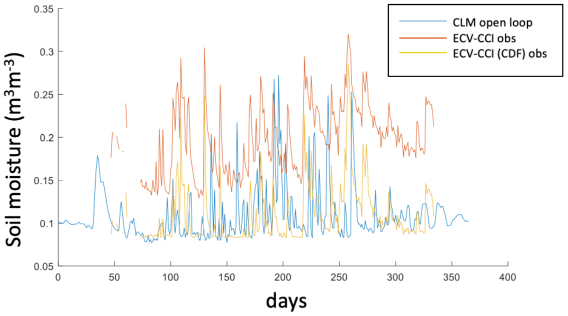


Soil Moisture - CDF matching

- CDF matching re-scales data products to match the bias and variability of the open-loop model



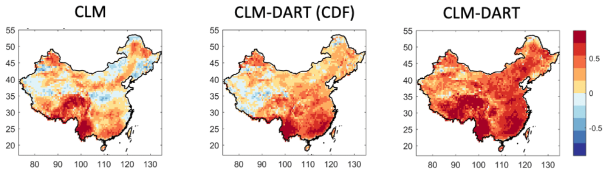
Reichle & Koster 2004 (GRL)



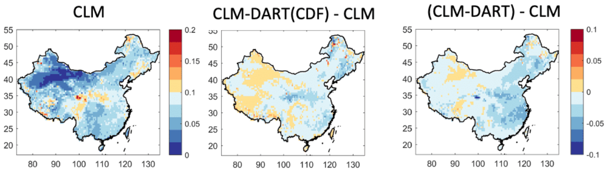
- The CLM-DART soil moisture product using the standard ECV-CCI product shows stronger correlations and reduced RMSD compared to ERA5Land benchmark

- Suggests CDF matched soil moisture product loses information, and inflation helps account for model error & bias

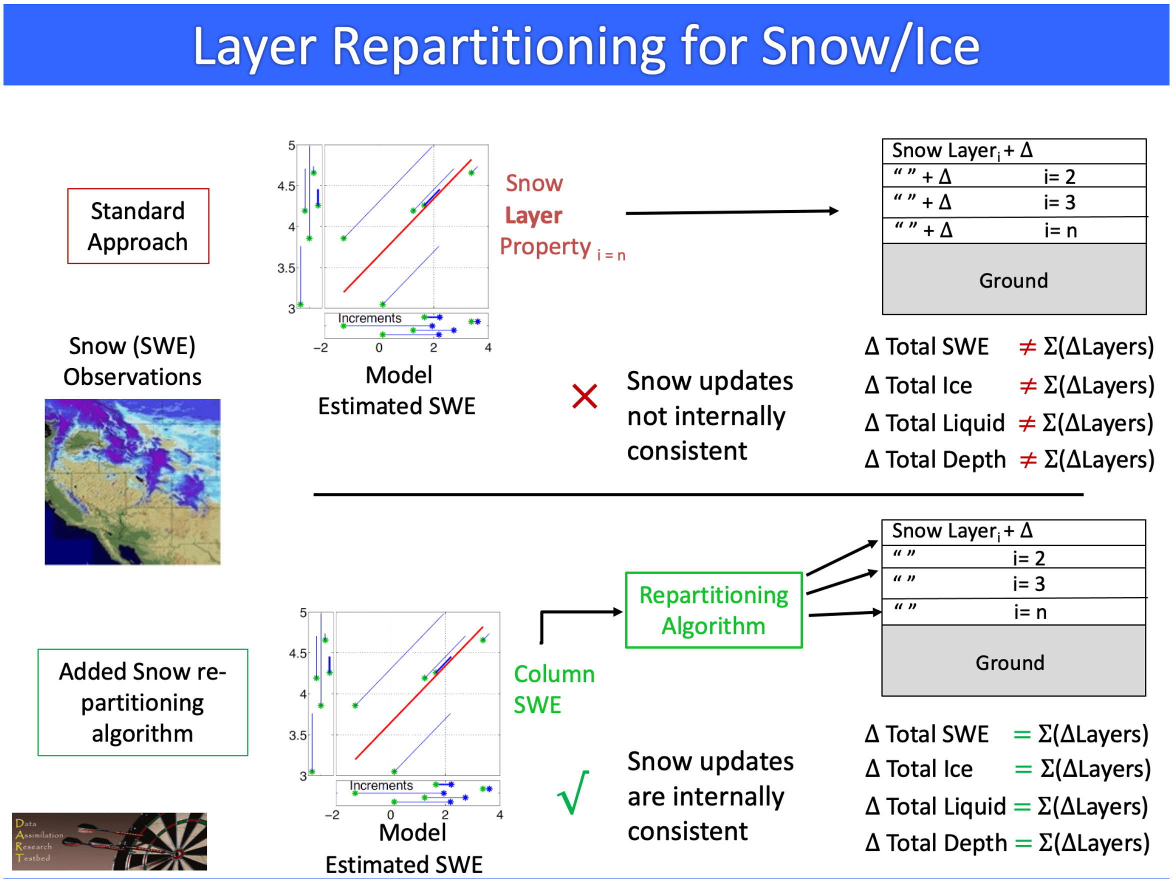
Correlation w/ ERA5



ubRMSD w/ ERA5



Raczka; SWE (snow water equivalent)



Layer Repartitioning for Snow/Ice

Standard Approach

Snow (SWE) Observations

Model Estimated SWE

Snow Layer Property $_{i = n}$

Snow Layer$_i$ + Δ	
“ ” + Δ	i= 2
“ ” + Δ	i= 3
“ ” + Δ	i= n
Ground	

✗ Snow updates not internally consistent

Δ Total SWE ≠ Σ(ΔLayers)

Δ Total Ice ≠ Σ(ΔLayers)

Δ Total Liquid ≠ Σ(ΔLayers)

Δ Total Depth ≠ Σ(ΔLayers)

Added Snow re-partitioning algorithm

Model Estimated SWE

Column SWE

Repartitioning Algorithm

Snow Layer$_i$ + Δ	
“ ”	i= 2
“ ”	i= 3
“ ”	i= n
Ground	

√ Snow updates are internally consistent

Δ Total SWE = Σ(ΔLayers)

Δ Total Ice = Σ(ΔLayers)

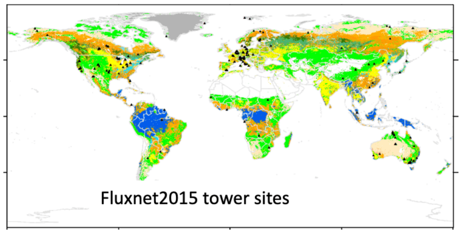
Δ Total Liquid = Σ(ΔLayers)

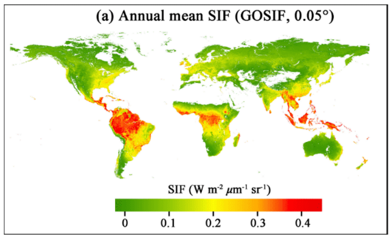
Δ Total Depth = Σ(ΔLayers)

Raczka; SIF (solar induced flourescence)

Solar-Induced Fluorescence (SIF)

- Spatial satellite coverage of SIF far more representative than EC flux tower network

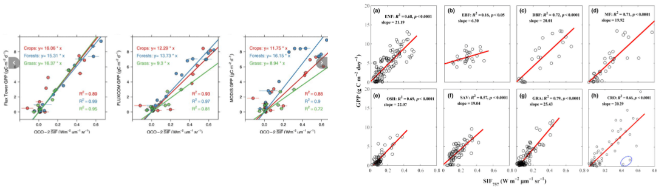



(a) Annual mean SIF (GOSIF, 0.05°)
SIF (W m^{-2} μm^{-1} sr^{-1})
0
0.1
0.2
0.3
0.4

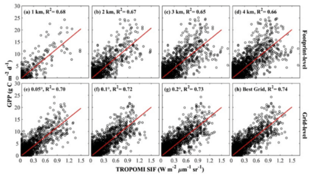
- Strong SIF-GPP (linear) relationship exists regardless of SIF platform, GPP product, biome
- Sub-daily? Slope?

Sun et al., (2017); Li et al., (2018); Xiao et al., (2019, 2022) Pierrat et al., (2022)

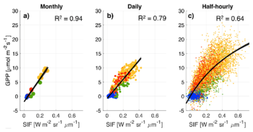
OCO$_2$-SIF vs. Tower GPP and GPP products



TROPOMI-SIF vs. Tower GPP

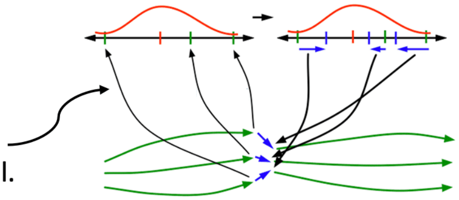


Photospec-SIF vs. Tower GPP



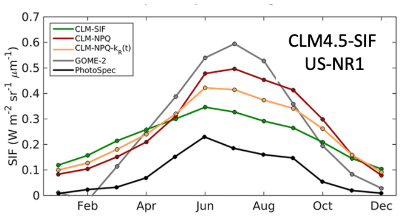
However, do we have a mechanistic understanding to simulate SIF ?

This is important in DA to generate an 'expected' SIF observation from a model.



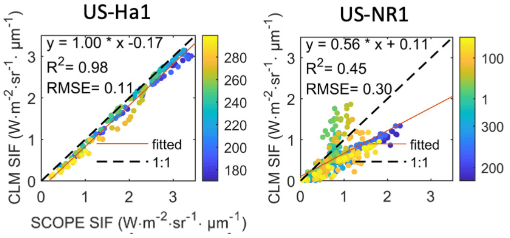
Sustained Nonphotochemical Quenching Shapes the Seasonal Pattern of Solar-Induced Fluorescence at a High-Elevation Evergreen Forest

Brett Raczka, A. Porcar-Castell, T. Magney, J. E. Lee, P. Köhler, C. Frankenberg, K. Grossmann, B. A. Logan, J. Stutz, P. D. Blanken, S. P. Burns, H. Duarte, X. Yang, J. C. Lin, D. R. Bowling



Representation of leaf-to-canopy radiative transfer processes improves simulation of far-red solar-induced chlorophyll fluorescence in the Community Land Model version 5

Rong Li, Danica Lombardozzi, Mingjie Shi, Christian Frankenberg, Nicolas C. Parazoo, Philipp Köhler, Koong Yi, Kaiyu Guan, Xi Yang



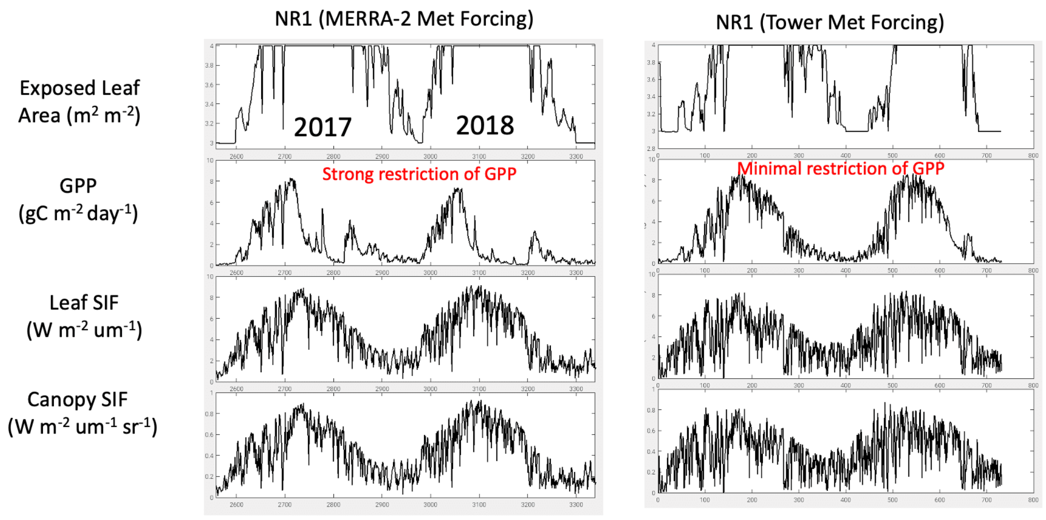
SIF Model Inter-Comparison 2 (SIF-MIP2)

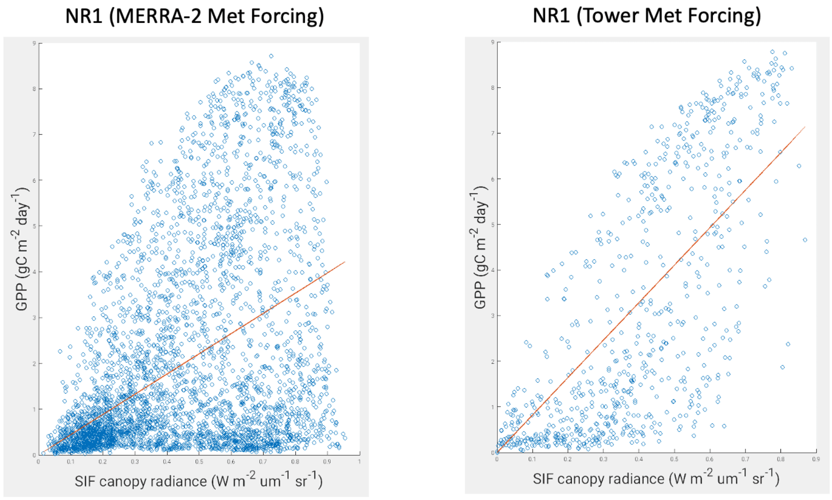
Goal: Apply tower GPP, canopy SIF and leaf SIF observations to CLM-DART

Sites: Evergreen DEJU (Alaska) NR1 (Colorado), OBS (Canada), NR1 (Colorado), OSBS (Florida)

Models: SIB4, ORCHIDEE, BEPS, CLM, JULES, CLiMA, SCOPE, VISIT, CARDAMOM

Groups: Free model simulation, Data Assimilation, Radiative Transfer Models





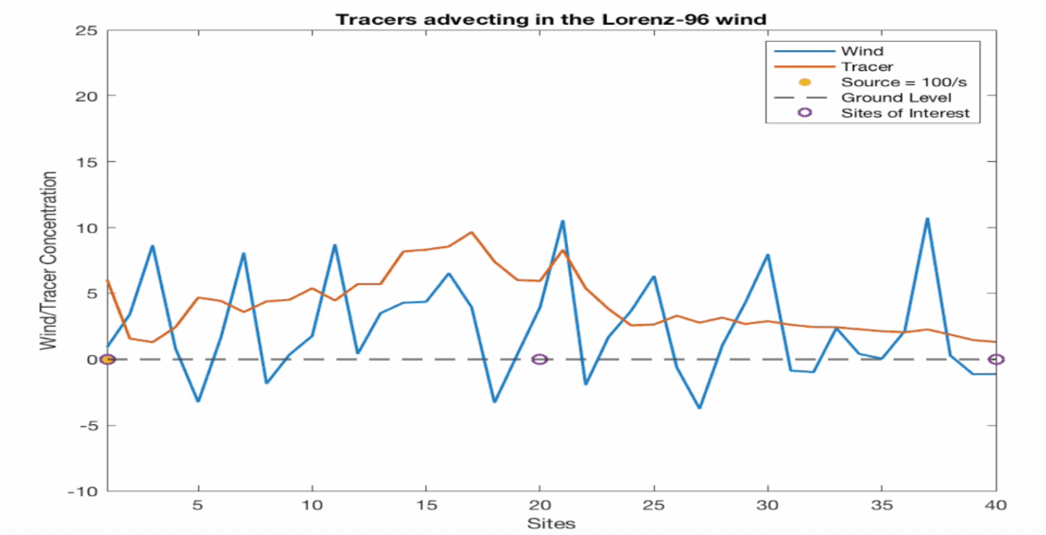
Labriola; CM1 and localization upgrades

The DART interface to Cloud Model v1 (CM1) can now handle mixed periodic boundary conditions; the x andor y dimension can be periodic or not. It also now handles interpolation of 3D fields such as reflectivity.

The threed_cartesian location module can use multiple localization radii, so that each observation type can have a localization radius appropriate for its correlation characteristics.

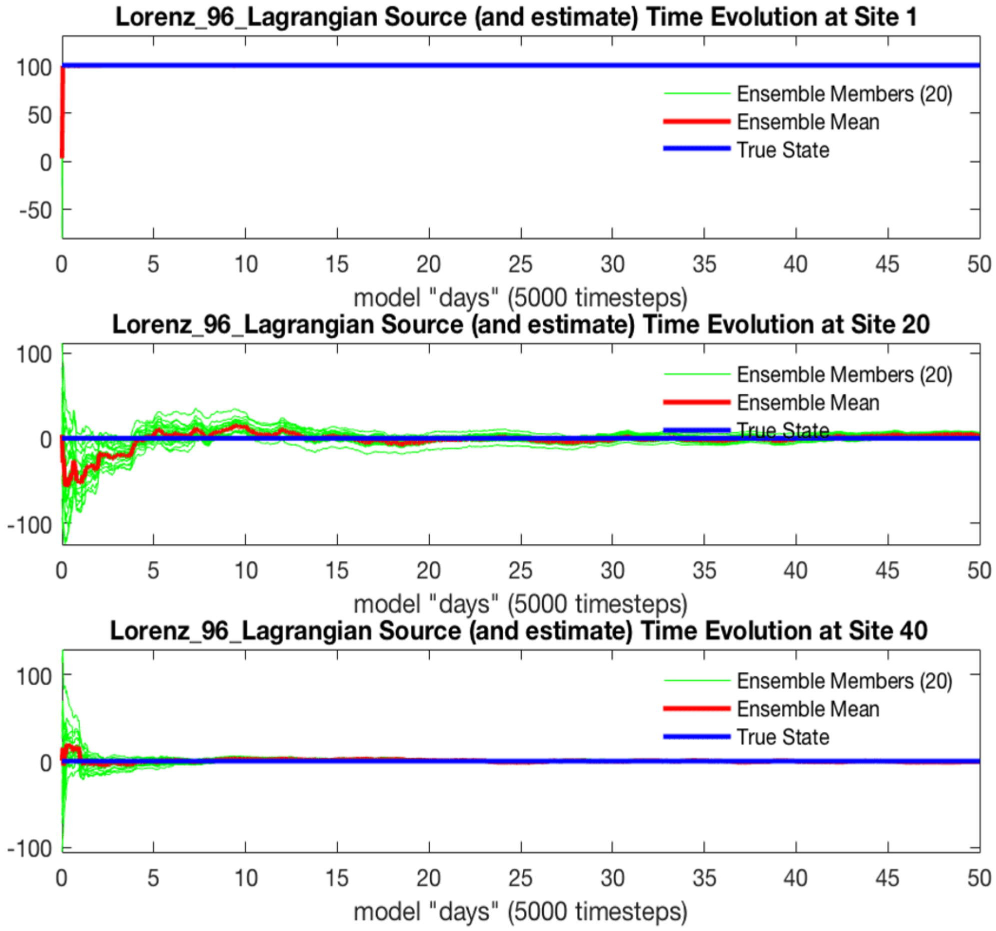
Ishraque; Tracer advection model

This new model interface was implemented by a SIPaRCS summer student to solve a long-standing need for a low order model (Lorenz 96) which uses a semi-Langrangian tracer advection scheme.



Model state: {wind,tracer} at 40 sites on a circle

Tracer source: site 1



Lorenz_96_Lagrangian Source (and estimate) Time Evolution at Site 1

Lorenz_96_Lagrangian Source (and estimate) Time Evolution at Site 20

Lorenz_96_Lagrangian Source (and estimate) Time Evolution at Site 40

The assimilation not only constrains the wind and tracer to be close to the truth (not shown), but can identify the location and strength (left axis) of the tracer source (site 1, 100/s).

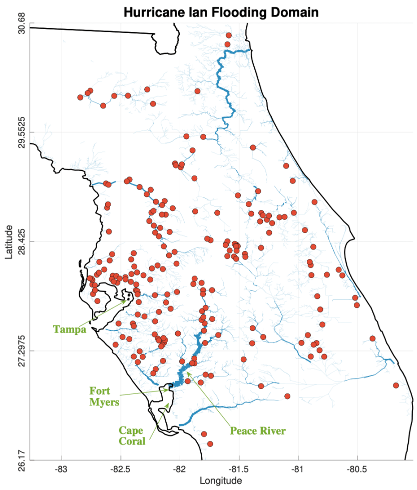
Lukas Kugler; Interface to RTTOV version 13

Forward operators for the RTTOV model (v13) for assimilation of satellite radiances from NOAA-15 ... -18; both RTTOV-direct (visible, infrared, and microwave) as well as RTTOV-scatt (microwave) computations. RTTOV = Radiative Transfer for Advanced TIROS Operational Vertical Sounder; TIROS = Television and Infrared Operational Satellite.

DART EXCLUSIVE: NEW ASSIMILATION ALGORITHMS

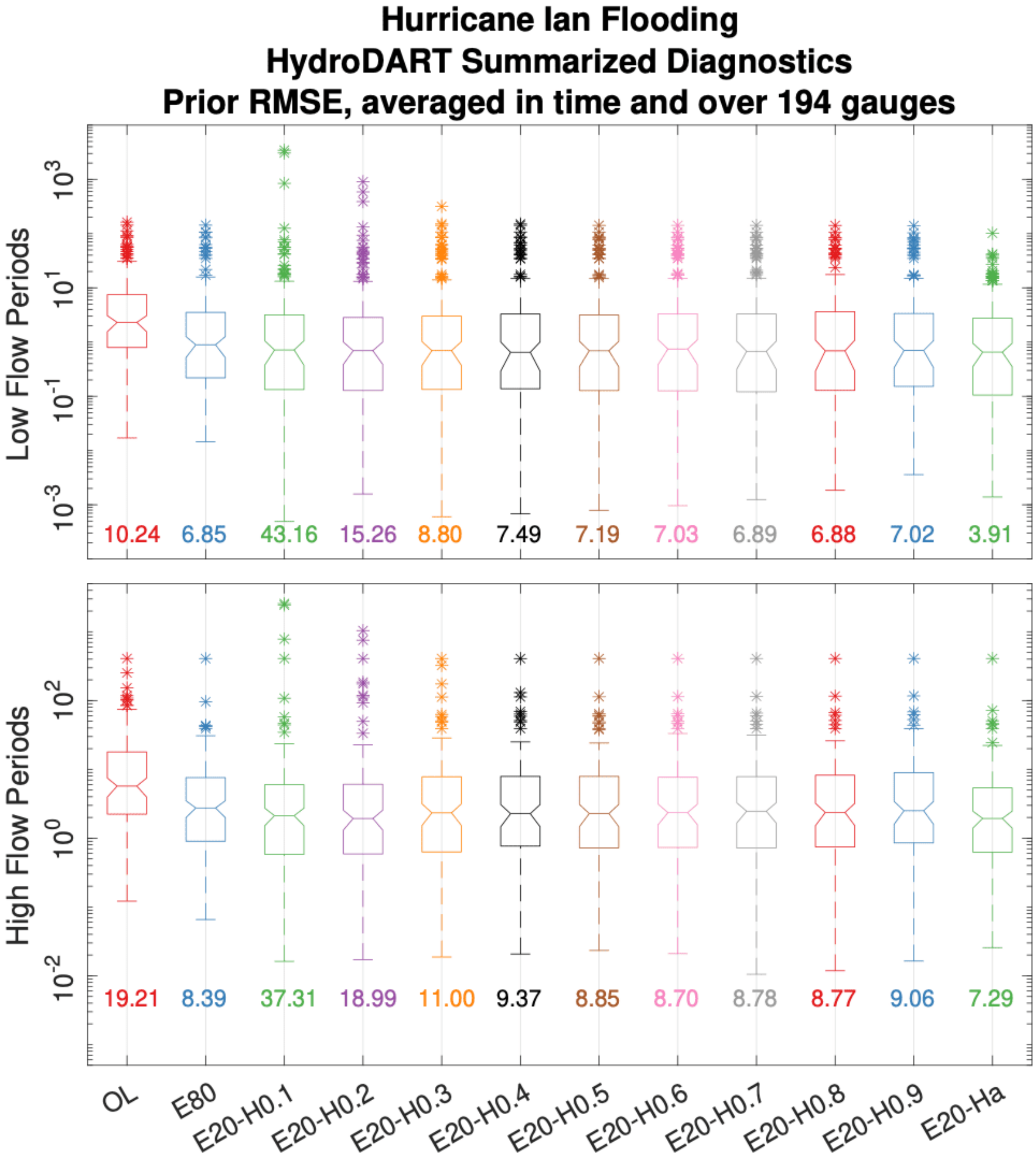
- Gharamti: Hybrid Ensemble-Variational Data Assimilation for Streamflow and Flood Prediction
- Anderson: Non-Gaussian and Nonlinear Ensemble DA Algorithms
- Gharamti: A Randomized Dormant Ensemble Kalman Filter

Gharamti: Hybrid Ensemble-Variational Data Assimilation for Streamflow and Flood Prediction



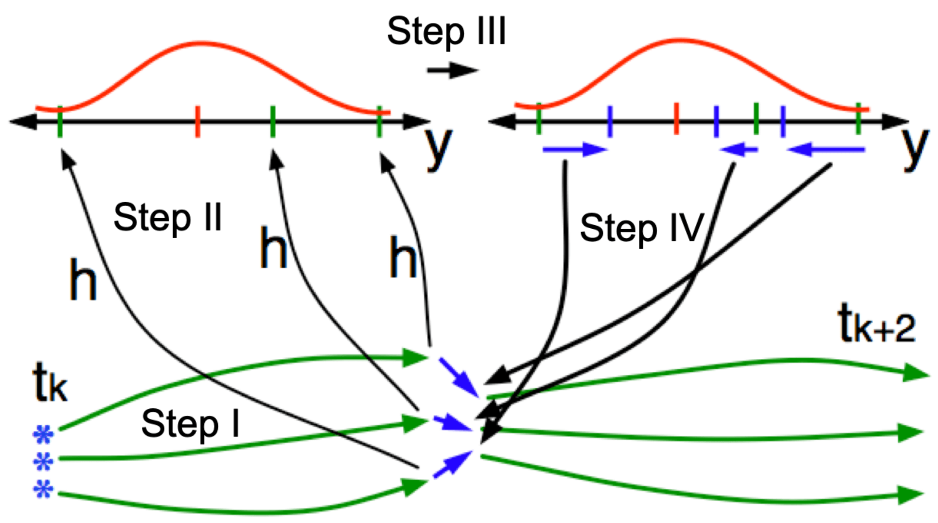
The recently updated version of WRF-Hydro has been coupled to the Data Assimilation Research Testbed (DART). The coupled system, namely HydroDART, is described in detail in El Gharamti et al., 2021. Shown obove is the study area where Hurricane Ian hit Florida causing massive flooding on September 28th, 2022. The stream network (links) is shown on the map in blue color. The thickness of each stream depicts the strength of the streamflow given by a free model run one month prior to the flooding event on August 15th, 2022. USGS gauges (red circles) provided hourly streamflow observations for constraining the model's prediction in DART over a month-period (Sep 15th - Oct 15th, 2022).

The boxplots (below) show the prior root-mean-squared-errors (RMSE) resulting from 12 different experiments where OL denotes open loop (no DA) and E80 is a typical DA run using 80 members (similar to those in El Gharamti et al., 2021). The remaining 10 experiments utilize a hybrid DA approach where the prior sample covariance, , obtained using only 20 members is linearly combined with a static background covariance matrix, , as follows: . 1000 realizations of the state are randomly selected from a 40-year model run (starting on 1970) and then used to construct the static (climatological) background covariance. As shown, the first 9 hybrid experiments assign constant weighting, , throughout the experiment. The last run uses an adaptive hybrid variant (E20-Ha) where the weighting factor remains spatially constant however it changes over time following El Gharamti 2021. The overall RMSE averages for each experiment are reported underneath the boxplots. On top of being computationally more efficient, E20-Ha clearly outperforms E80 especially for low-flow periods where the standard EnKF suffers from low ensemble variability. On average, the estimates suggested by the E20-Ha scheme are 43% and 13% more accurate than those obtained using E80 for the low flow and high flow periods, respectively.


Hurricane Ian Flooding
HydroDART Summarized Diagnostics
Prior RMSE, averaged in time and over 194 gauges

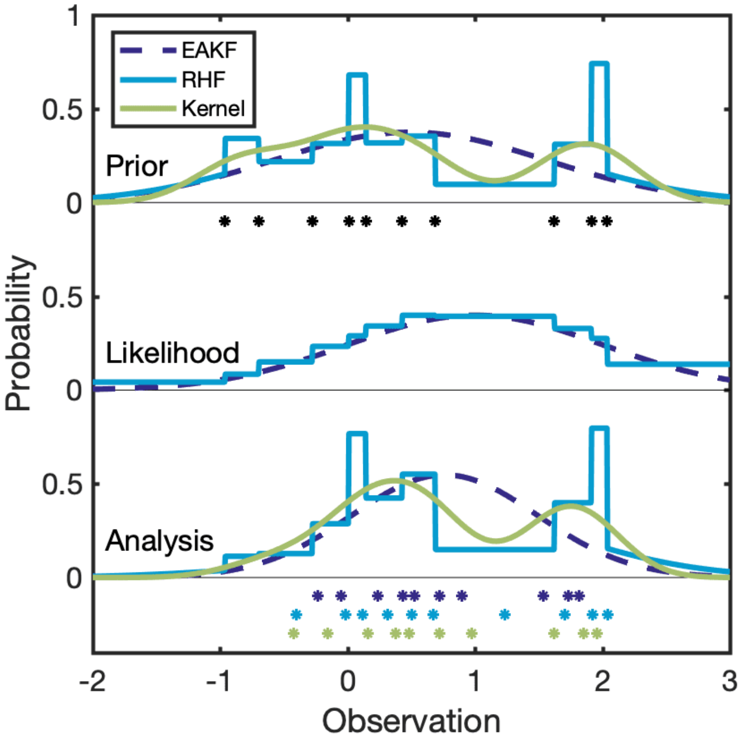
Anderson: Non-Gaussian and Nonlinear Ensemble DA Algorithms

The schematic in Figure JLA.1 shows how DART implements ensemble DA. Computing ensemble increments for an observed quantity (step 3) and regressing those observation increments onto each state variable (step 4) can be done independently. The EAKF that has been a DART mainstay assumes a Gaussian distribution for step 3 and does standard linear regression for step 4.



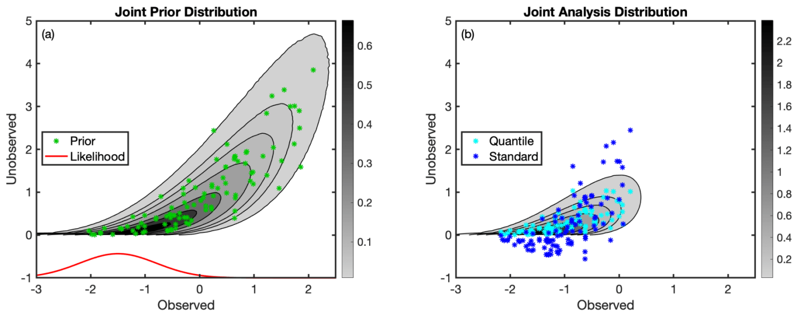
JLA.1: A schematic of the DART assimilation framework. A model produces ensemble forecasts in step I. Observations are assimilated sequentially. The forward operator is applied to each ensemble state vector to obtain a prior observation ensemble in step II. Step III combines the prior observation ensemble with an observation likelihood to produce an updated ensemble and corresponding increments. Step IV regresses the observation increments onto each state variable ensemble independently.

DART now implements a novel efficient algorithm that allows use of arbitrary continuous observation priors and likelihoods for step 3. The key innovation is selecting posterior ensemble members with the same quantiles with respect to the continuous posterior distribution as the prior ensemble had with respect to the prior continuous distribution. This is a generalization of previously documented square root ensemble Kalman filters for normal distributions. It also generalizes non-parametric ensemble filters such as the rank histogram filter. Examples of new continuous priors that can be implemented include gamma, inverse gamma, beta, a sum of normal kernels, and a bounded rank histogram which is a general non-parametric technique that works well for any application. Figure JLA.2 shows some examples of applying different observation distributions to the same prior ensemble.



JLA.2: An example of applying a quantile conserving ensemble filter with three different continuous prior distributions fit to the same prior ensemble (top panel). The priors include a normal (same as the EAKF), a rank histogram, and a sum of individual Gaussian kernels centered around each ensemble member. The continuous posterior distributions and the associated posterior ensembles are shown in the lower panel. Figure is reproduced from Anderson, 2022, MWR, 150, 1061-1074.

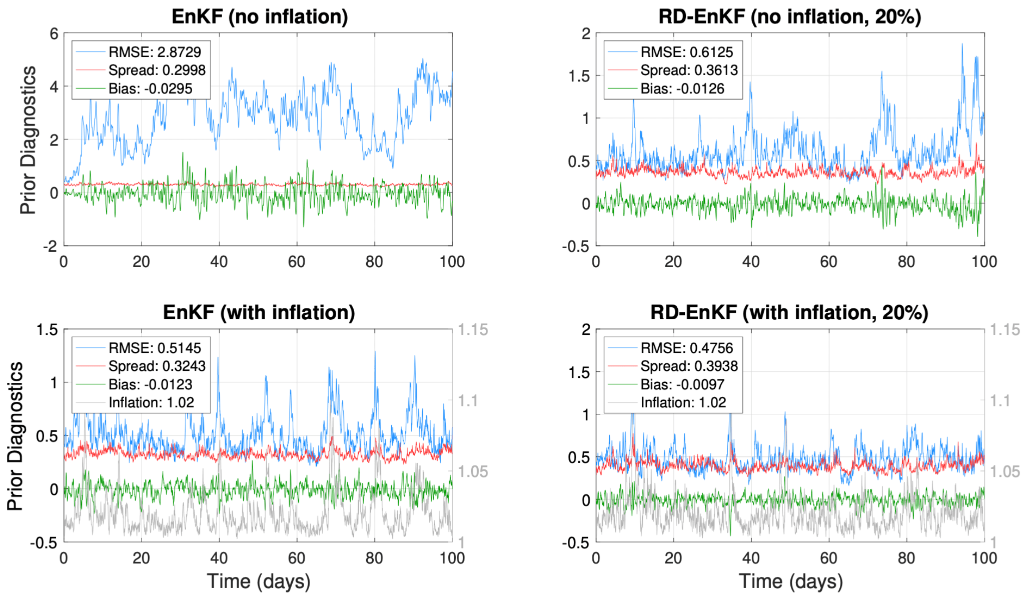
While quantile conserving algorithms for step 3 lead to significant improvements in analysis estimates for observed variables, those improvements can be lost when using standard linear regression of observation increments to update other state variables in step 4. However, doing the regression of observation quantile increments in a probit-transformed bivariate quantile space guarantees that the posterior ensembles for state variables also have all the advantages of the observation space quantile conserving posteriors. For example, if state variables are bounded then posterior ensembles will respect the bounds. The posterior ensembles also respect other aspects of the continuous prior distributions. Figure JLA.3 shows an example of the new regression method. The method can significantly improve data assimilation for non-Gaussian quantities like tracers, snow depth, or sea ice concentration, in Earth system models. The method is also effective for estimating the value of bounded model parameters. A beta release of DART scheduled for the second week of January 2023 will include the new regression algorithms.



JLA.3: Panel a shows a 100-member prior ensemble drawn from a bivariate distribution with the marginal of the observed variable being normal and the marginal of the unobserved state variable being a gamma distribution (appropriate for a tracer). An accurate approximation of the continuous prior PDF is shown by the shading in the figure. The observation increments are computed using a quantile conserving ensemble filter with a normal continuous prior distribution. The blue asterisks in panel b are the result of using linear regression to compute increments for the unobserved variable while the cyan asterisks use a probit-transformed quantile regression. This regression respects the bounds on the bounded quantity and more accurately represents the nonlinear relation between the observed and unobserved variable.

Gharamti: A Randomized Dormant Ensemble Kalman Filter

This work introduces a new variant of the Ensemble Kalman Filter that aims to improve the estimate of the background ensemble perturbations and mitigate variance underestimation. The new filter is called randomized dormant ensemble Kalman filter (RD-EnKF) and uses prior ensembles constructed from active and dormant state realizations. At each assimilation cycle, a set of the ensemble is randomly selected to go through the analysis scheme of the EnKF. This set consists of the active ensemble members. The remaining set consists of the dormant members which do not take part in the analysis. After the update, both active and dormant members are used to perform a forecast to get to the next data assimilation cycle. The number of dormant ensemble members is predefined and can be changed every cycle. When assimilating the observations serially, as in the Data Assimilation Research Testbed (DART), it's also possible to change the dormant set of members for each observation. This makes it possible for all the members to be updated but with different number of observations. Compared to the standard EnKF, the background ensemble spread given by the RD-EnKF is often larger as shown below. To the left, the EnKF with 10 members quickly diverges after few assimilation cycles. The RD-EnKF (right panel), on the other hand, remains robust even with 10 members producing a more reliable forecast. Introducing inflation stabilizes the estimates of the EnKF yet the RD-EnKF is still able to produce better consistency between the prior RMSE and ensemble spread.



Sensitivity experiments of the RD-EnKF with respect to the dormancy rate, localization radius and observation network.

Additionally, the RD-EnKF sample covariances have better statistical properties (e.g., rank) than their EnKF counterparts which makes the algorithm computationally more stable, requiring less tuning of localization. As can be seen (top panel), increasing the dormancy rate improves the estimates of the RD-EnKF even with wider localization radii. The algorithm is further robust to changing observation networks (bottom panel) suggesting accurate prior estimates even in sparsely observed regions.

The algorithm is highly effective under extreme sampling error scenarios. Future work will focus on testing the performance of the RD-EnKF in regional/global ocean DA applications.

PERFORMANCE AND USABILITY

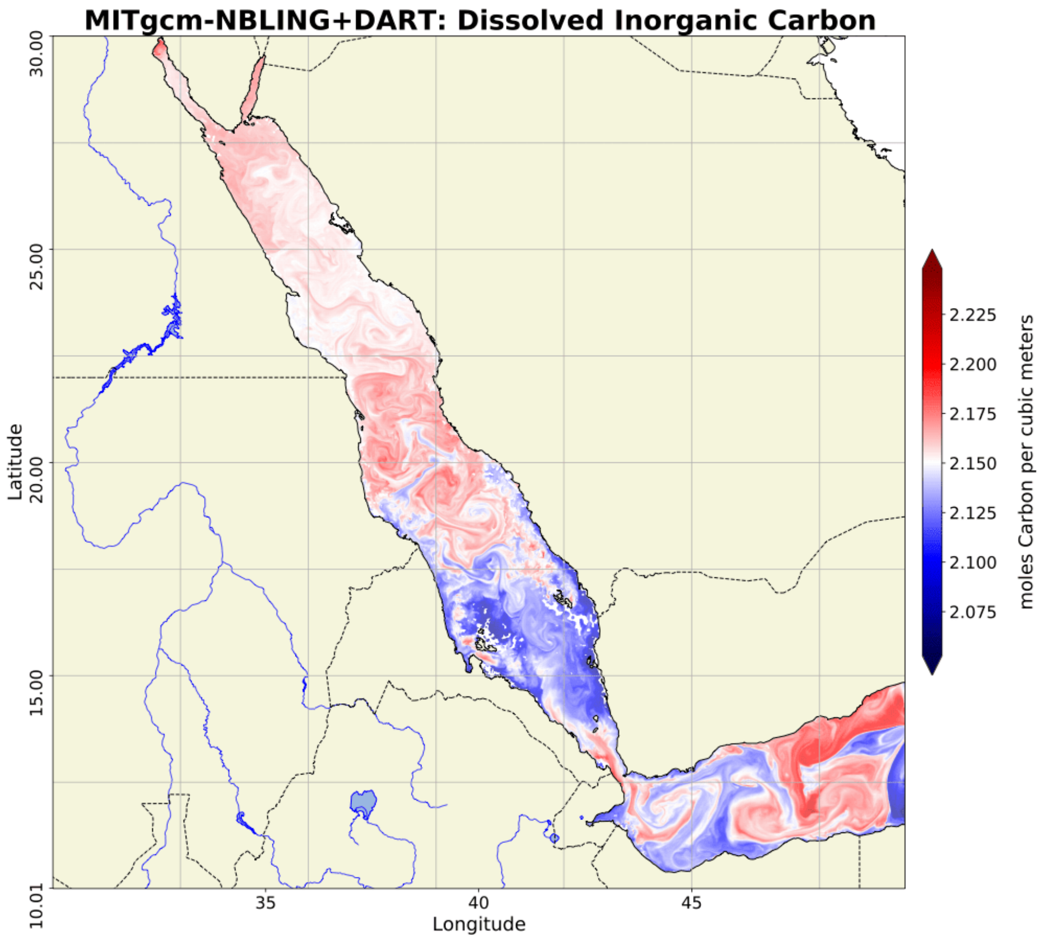
- Liu; State Compaction in MITgcm-NBLING
- Liu, Smith; Improved Caching
- Kershaw; Building Executables
- Kershaw; Flexibility
- Kershaw, Johnson: Upgraded web site and documentation search

----

Liu; State Compaction

Model grid points with no active variables are excluded from the state passed to filter:

- Enables DA with very large, but sparse, state vectors; ocean models where much of the domain is land, and vice versa.
- Only for MIT_gcm at the moment.



This Red Sea model domain (resolution=0.01 degrees) requires model state compaction in order run in DART. It yields ~90% reduction in state vector size. This dissolved organic carbon distribution was generated by an atmosphere+ocean+biogeochemistry model and DART.

Liu, Smith; Improved Caching

- Ed Liu (summer student) discovered+profiled, Marlee Smith fixed
- Redundant caching in the get_close_obs_cached and get_close_state_cached subroutines has been removed.
- Savings of up to 20% run time.

Kershaw; Building Executables

New system:

- Reduced the number of files in DART by ~30%
- Simplified the building of a single executable (default = build all)
- Enabled modifying a local ("work" directory) copy of a source file instead of the copy in its usual home.

Nancy Collins, Kershaw; Flexibility

New mechanism for defining obs quantities, localization, and model interface building:

- Removes the need to have a hard coded list of integers for DART quantities.
- Users can add new quantities by defining a QTY_NAME

Kershaw, Johnson: Upgraded DART web site and documentation search

https://docs.dart.ucar.edu/en/latest/README.html (https://docs.dart.ucar.edu/en/latest/README.html)

SUMMARY

DART is an open source suite of tools which empowers ensemble data assimilation research over a broad spectrum of scales in multiple dimensions:

- Model interfaces range from the 3 variable Lorenz 63 to global models with millions of variables.
- Core assimilation algorithms ranging from the venerable "square root" filters to particle filters to the first ever Quantile Conserving Filters with regression of observation increments in probit-transformed, bivariate quantile space.
- Observations range from "perfect model" observations through traditional in situ observations of the Earth system to recently added radiance observations from satellites.
- Educational features ranging from a basic ensemble DA tutorial through exploration of myriad assimilation algorithms and models and an extensive publications list.
- Ancillary tools range from simple algorithms for minimizing the required ensemble size to scripts for evaluating assimilation performance to large reanalysis data sets for use in surface model assimilation experiments.
- Contributions to the software from senior researchers all the way through high school students.

Join us at AMS 2023 for a celebration of 20 years of DART development!

DISCLOSURES

The National Center for Atmospheric Research is sponsored by the National Science Foundation. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

AUTHOR INFORMATION

Kevin Raeder[1], Jeffrey L. Anderson[1], Moha Gharamti[1], Helen Kershaw[1], Brett Raczka[1], Ben Johnson[1], Marlee Smith[1], Ed Liu[2], Jon Labriola[3], Fairuz Ishraque[4], Daniel Hagan[5]

[1]National Center for Atmospheric Research, Computational and Information Systems Laboratorr, Data Assimilation Research Section

[2]Drexel University

[3]University of Oklahoma

[4]Princeton University

[5]Nanjing University of Information Science & Technology, Nanjing, China

ABSTRACT

The Data Assimilation Research Testbed (DART) is a community facility for ensemble data assimilation developed and maintained by the National Center for Atmospheric Research (NCAR). This poster highlights new capabilities that have been added to DART in the last year that are useful to researchers doing data assimilation for Earth system applications:

- Completion of a decade of CAM6+DART reanalysis, which provides ensemble atmospheric forcing for surface models, and more.
- New or improved model interfaces include: TIEGCM, MITgcm_ocean NBLING (biogeochemistry), a low order tracer advection model (student contribution), and CAM Spectral Element.
- New observation interfaces include: snow water equivalent, soil moisture, with Solar Induced Fluorescence (SIF) and snow cover in development.
- Efficiency gains and handling higher resolution models by compaction of sparse state vectors in ocean and land models, and improved caching (student contributions).
- Quantile conserving and quantile regressing filters support a wide range of non-Gaussian and nonlinear ensemble updates.
- A randomized dormant ensemble filter deals with sampling errors.
- More flexibility in defining observation quantities, localization, and model interface building.
- Upgraded web site documentation search and a CLM5+DART tutorial.
- Improved soil moisture estimates over China using CLM+DART and soil moisture observations (ESA-CCI).

REFERENCES

An extensive, but incomplete, list of DART papers can be found in https://dart.ucar.edu/publications/ (https://dart.ucar.edu/publications/)

For the most recent advancements, see presentations at AGU2022 by the authors and others mentioning DART:

Jeffrey Anderson:

- A33D; Removing the Kalman from the Ensemble Kalman Filter
- NG32A-06: A Quantile Conserving Ensemble Filtering Framework: Regressing Quantile Increments to Update Unobserved Variables

Nick Pedatella:

- SM25C-1992: Assessing the Impact of Assimilating Ionosphere-Thermosphere Observations on Analysis and Short-term Forecasts (invited)
- SA32B-01: Validation of Ionosphere Data Assimilation Systems using GNSS Positioning Algorithms (invited)

Xeuli Huo:

- B46B-02: Assimilating Leaf Area Index and Aboveground Biomass into the Community Land Model to Constrain Carbon Dynamics in the Arctic and Boreal Region

Chih-Chi Hu:

- NG33A-04: An observing system simulation experiment (OSSE) using the Particle Flow Filter (PFF) in a high-dimensional geophysical system in the Data Assimilation Research Testbed (DART)

Tomoko Matsuo:

- SA25C-1935: Community Data Assimilation Tools for the GDC mission and Beyond

Nicholas Dietrich:

- SA42B-06: Specifying Neutral Density and Satellite Drag Through Coupled Thermosphere-Ionosphere Data Assimilation of Radio Occultation Electron Density Profiles

Hristo Chipilski

- NG32A-01: A Conjugate Transform Filter for Geophysical Applications with Varying Degrees of Nonlinearity (invited)

Ben Gaubert:

- A42O-1900: Assimilation of CO from MOPITT, CrIS and TROPOMI to evaluate small fires, biogenic and anthropogenic sources in Sub-Saharan Africa