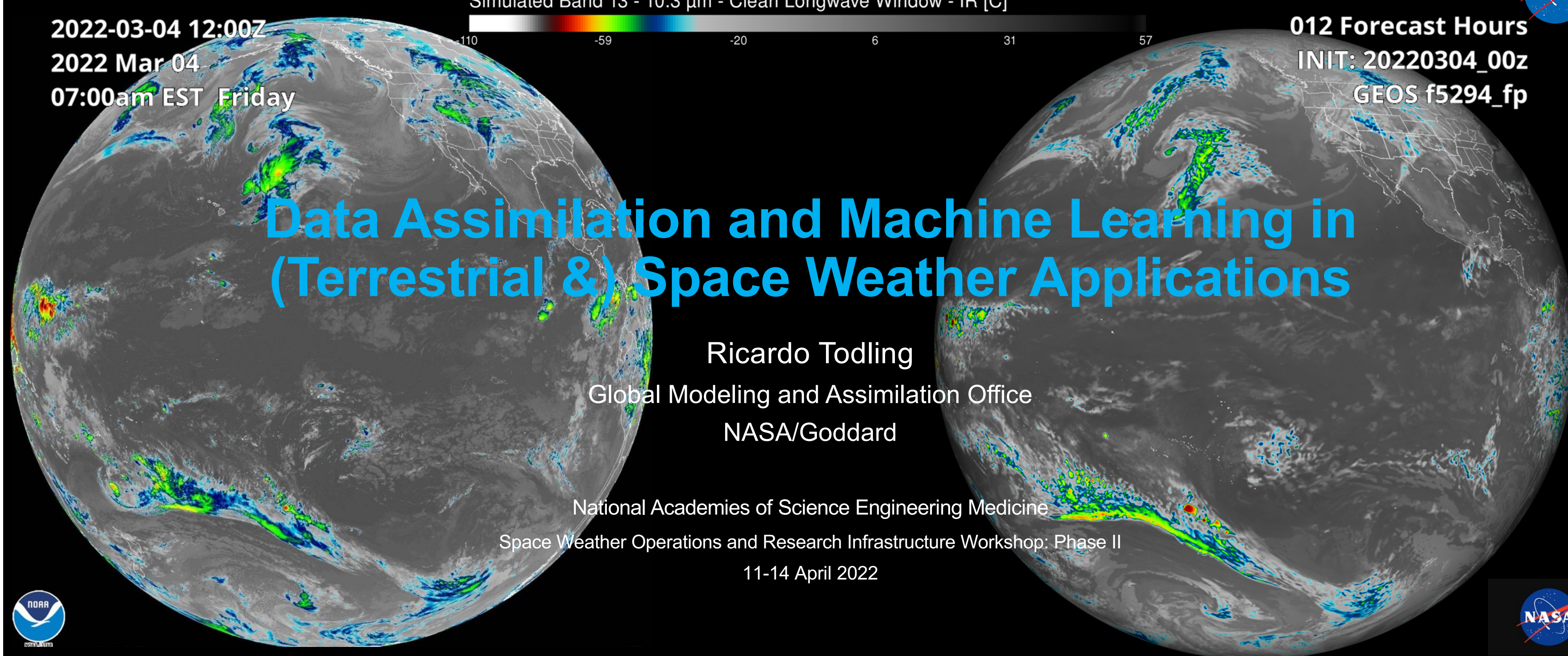
National Aeronautics and Space Administration



Data Assimilation and Machine Learning in (Terrestrial &) Space Weather Applications

Ricardo Todling
Global Modeling and Assimilation Office
NASA/Goddard

National Academies of Science Engineering Medicine
Space Weather Operations and Research Infrastructure Workshop: Phase II
11-14 April 2022

Disclaimer: This presentation has a strong Terrestrial Weather Applications bias; it might need some UQ to adjust it to Space Weather Applications!

OUTLINE

1. The Progress in Terrestrial Weather Prediction through DA

2. The Hierarchy of Models and DA Strategies in Terrestrial and Space Weather

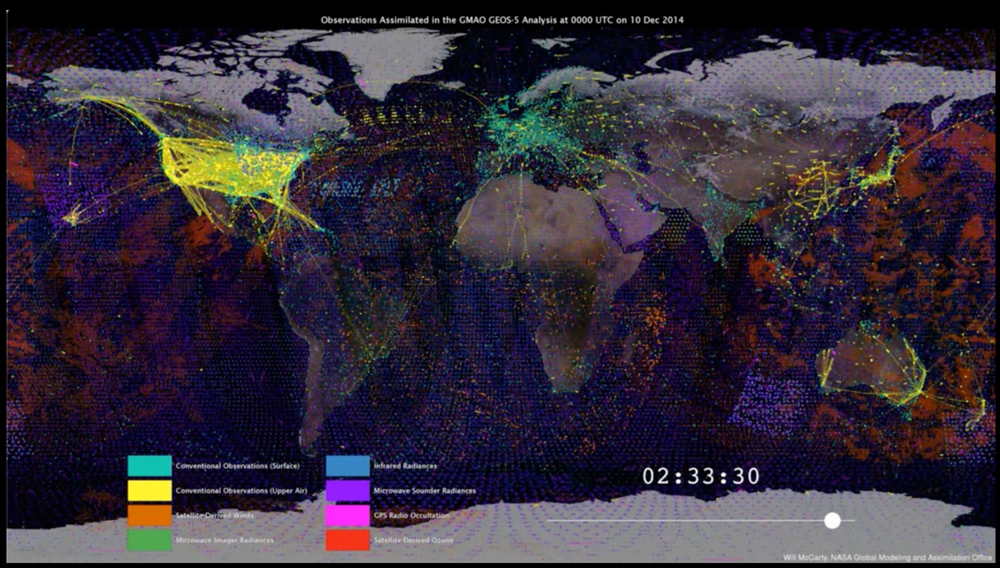
3. Hybrid Concepts

   - The Learning Aspect of DA
   - Machine Learning as Tool to Aid DA

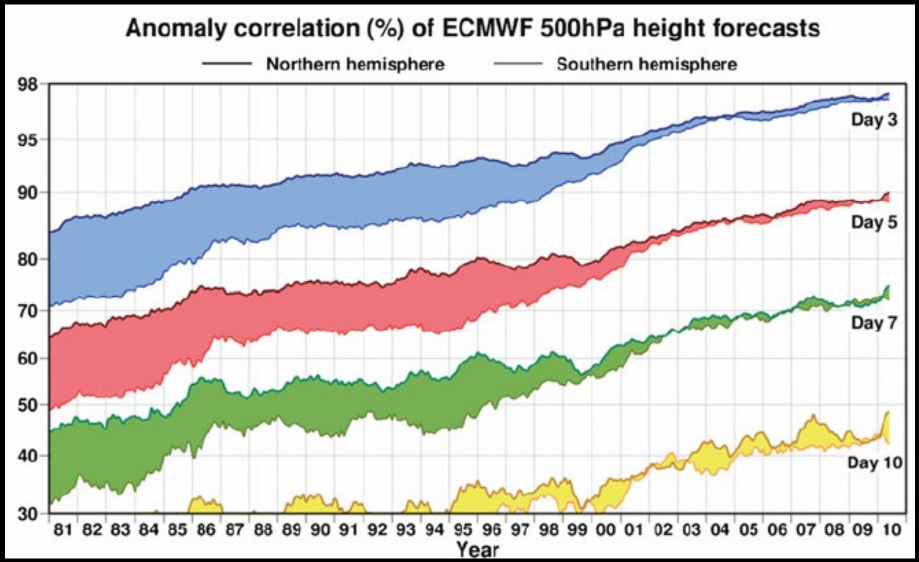
4. A Few Words on Frameworks

5. Closing Remarks

Terrestrial Weather Prediction: Better Data, Models & Techniques



Evolution of ECMWF forecast skill for varying lead times (3 days in blue; 5 days in red ; 7 days in green; 10 days in yellow) as measured by 500-hPa height anomaly correlation. Top line corresponds to the Northern Hemisphere; bottom line corresponds to the Southern hemisphere. Large improvements have been made, including a reduction in the gap in accuracy between the hemispheres (Source: Courtesy of ECMWF. Adapted from Simmons and Hollingsworth (2002).



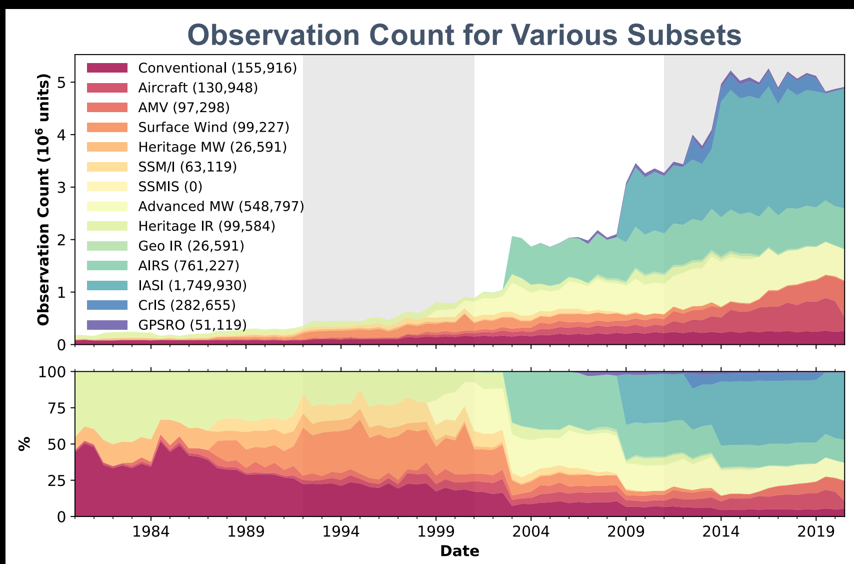
Observation assimilated in GEOS in the 6-hour period between 2100 UTC 9 Dec 2014 and 0300 UTC 10 Dec 2014 (Courtesy of Will McCarty).

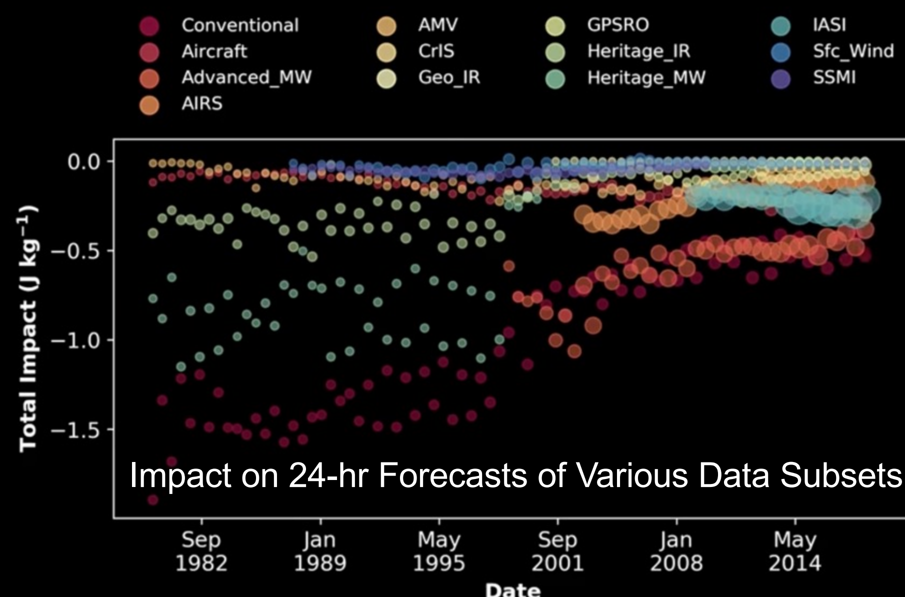
OBS/cycle: 5×10^6

Model: 10^9-10^{10}

Terrestrial DA: Impact of 40 Years of Assimilation

Illustration of the increase in data count in MERRA-2 over the past 40-plus years. The impression of a settling data count toward present day is simply a reflection of limitations in M2 to add newly available sensors; a look at the near-real-time, high resolution, GEOS DA system would reveal a continued rise in data count.





Impact of different types of assimilated observations along the course of MERRA-2. The reduced impact in absolute terms is a consequence of the improved quality in the state of the model due to the assimilation of an increased number of high-quality sensors. (size of dots is obs count (Diniz & Todling 2020)

Terrestrial DA-based Predictions: Range of Scales

The accuracy of weather forecasts is a result of increased model resolution, physical processes representation and the large volume of observations assimilated through advanced DA techniques.

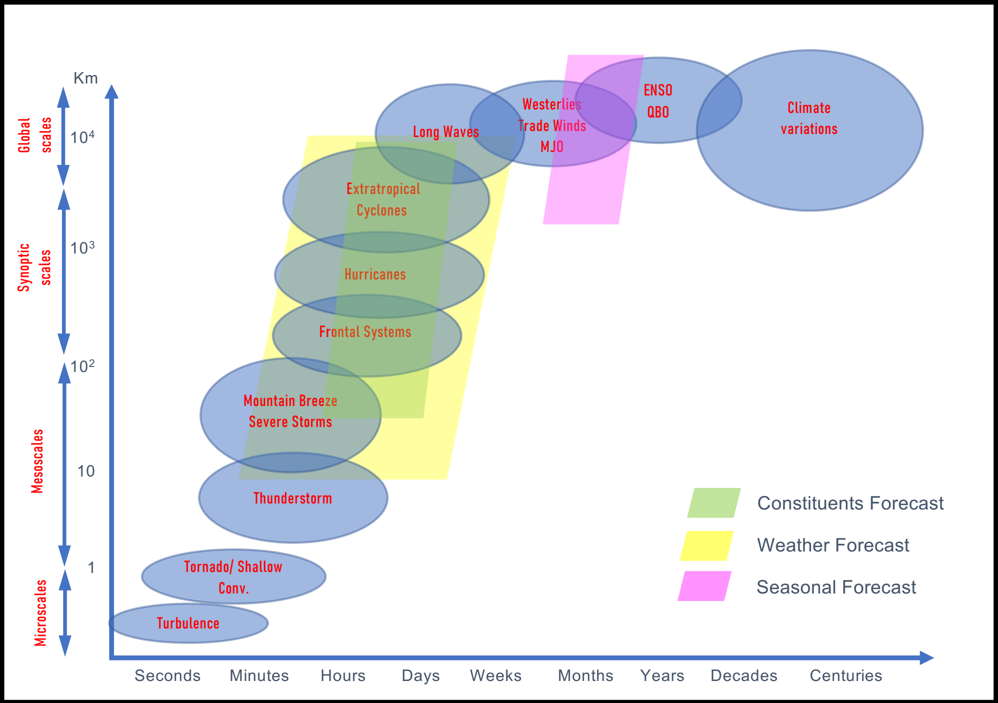
The diagram means to illustrate the range of applicability of DA to Global Terrestrial applications.

Global NWP is now entering the low range of the mesoscales.

Global NWP ranges from hours up to 10 days.

Global Constituent Forecast ranges from hours to 5 days.

Seasonal Prediction extends NWP capabilities in time, with added model complexities, but at the cost of reduced resolution.



Adapted from Tavakolifar et al. (2017; J. Climate)

Terrestrial DA & Prediction: A hierarchy of Components & Strategies

Three Examples from GEOS Forecasting Systems

Different applications invoke different level of model coupling.

Not a one-fits-all approach: Each Forecast System typically includes more than one DA approach.

The Replay strategy roughly nudges one system to results from another.

What's BC today tends to turn into a full modeled component tomorrow.

	Forecasting Systems		
Model Coupled Components	**Weather 12.5 Km**	**Seasonal 50 Km**	**Chemical Composition 50 Km**
Meteorology	Hybrid 4DEnVar	3D Replay	3D Replay
Ozone	Hybrid 4DEnVar	3D Replay	3D Replay
Aerosols	3DVar	3D Replay	3D Replay
Land	None (Soon EnKF)	None	None
Sea-Ice	BC	None	BC
Ocean	BC	3D-EnOI	BC
Chemical Constituents			None (Soon 3D-Var)
Emissions	BC	BC	BC

Included Prescribed (BC) Parameterized

Space Weather Prediction: A hierarchy of models
Larger (shorter) range of spatial (temporal) scales

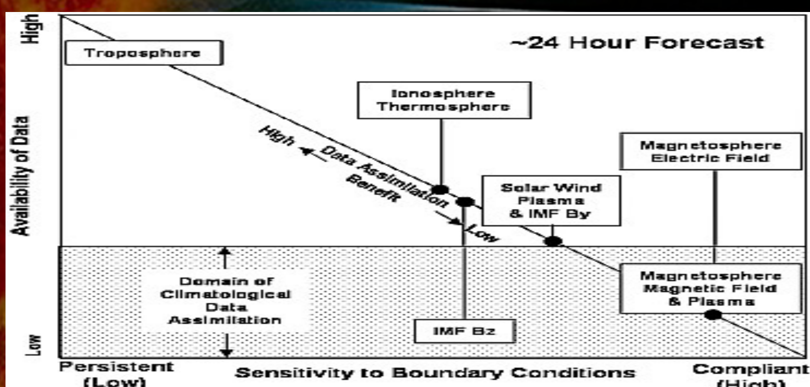
Coronal Model
(e.g., MAS, WSA)

Solar Wind Model
(e.g., Enlil, EUHFORIA)

Magnetosphere Model
(e.g., LFM)

Ionosphere Model
(e.g., Ridley)

Thermosphere/
Atmosphere
(WAM-IPE)

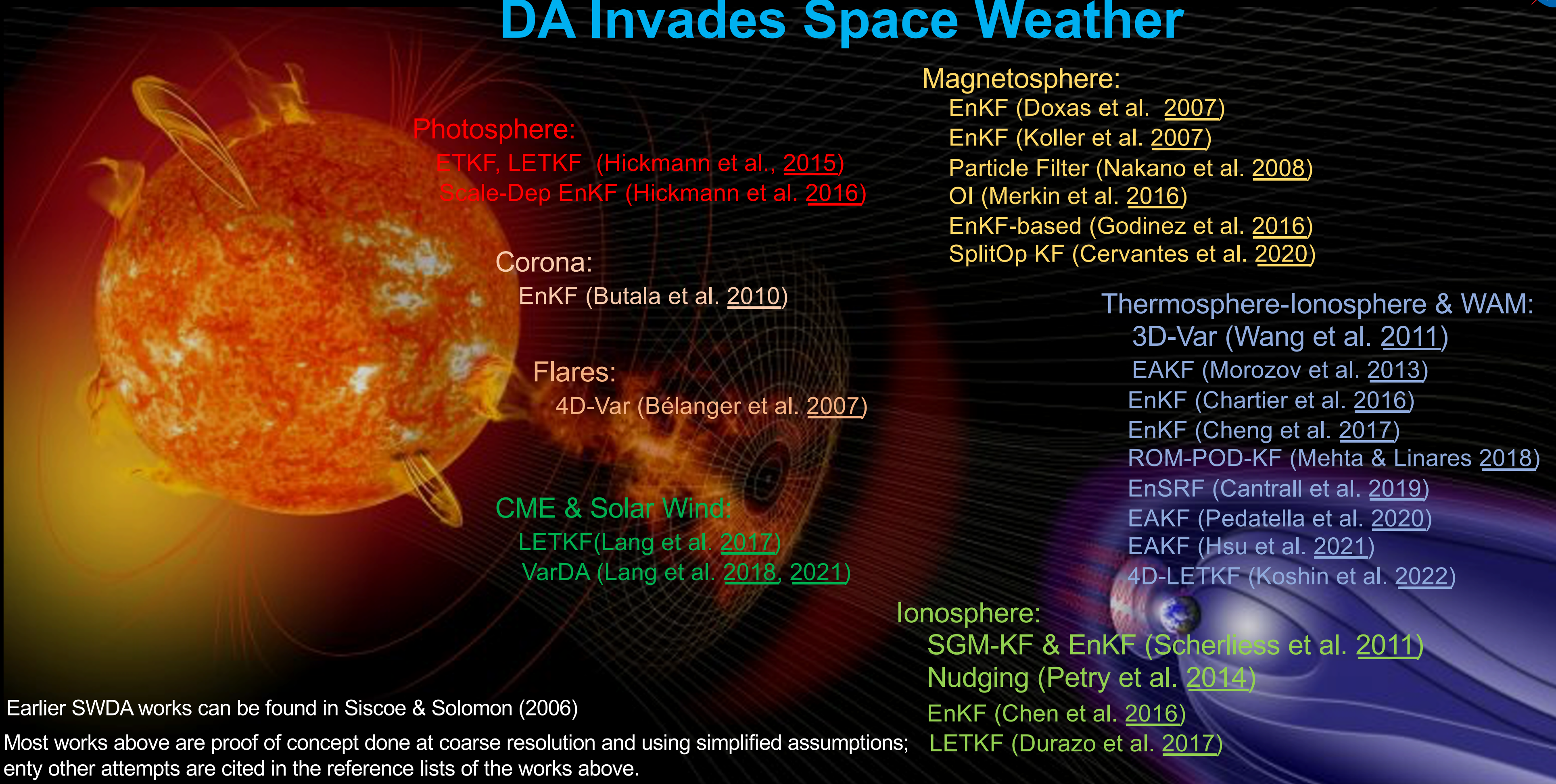


From Siscoe & Solomon (2006)

- Data Sparsity
- Directly vs Instability Driven Dynamics
- Short Timescales
- Range of Forecast Validity
- Intervening Turbulence vs Sensitivity to Initial Conditions

Background pic by: K. Endo

DA Invades Space Weather



Photosphere:
- ETKF, LETKF (Hickmann et al., 2015)
- Scale-Dep EnKF (Hickmann et al. 2016)

Corona:
- EnKF (Butala et al. 2010)

Flares:
- 4D-Var (Bélanger et al. 2007)

CME & Solar Wind:
- LETKF(Lang et al. 2017)
- VarDA (Lang et al. 2018, 2021)

Magnetosphere:
- EnKF (Doxas et al. 2007)
- EnKF (Koller et al. 2007)
- Particle Filter (Nakano et al. 2008)
- OI (Merkin et al. 2016)
- EnKF-based (Godinez et al. 2016)
- SplitOp KF (Cervantes et al. 2020)

Thermosphere-Ionosphere & WAM:
- 3D-Var (Wang et al. 2011)
- EAKF (Morozov et al. 2013)
- EnKF (Chartier et al. 2016)
- EnKF (Cheng et al. 2017)
- ROM-POD-KF (Mehta & Linares 2018)
- EnSRF (Cantrall et al. 2019)
- EAKF (Pedatella et al. 2020)
- EAKF (Hsu et al. 2021)
- 4D-LETKF (Koshin et al. 2022)

Ionosphere:
- SGM-KF & EnKF (Scherliess et al. 2011)
- Nudging (Petry et al. 2014)
- EnKF (Chen et al. 2016)
- LETKF (Durazo et al. 2017)

Earlier SWDA works can be found in Siscoe & Solomon (2006)

Most works above are proof of concept done at coarse resolution and using simplified assumptions; enty other attempts are cited in the reference lists of the works above.

The Data Assimilation Soup

Bayesian View of DA

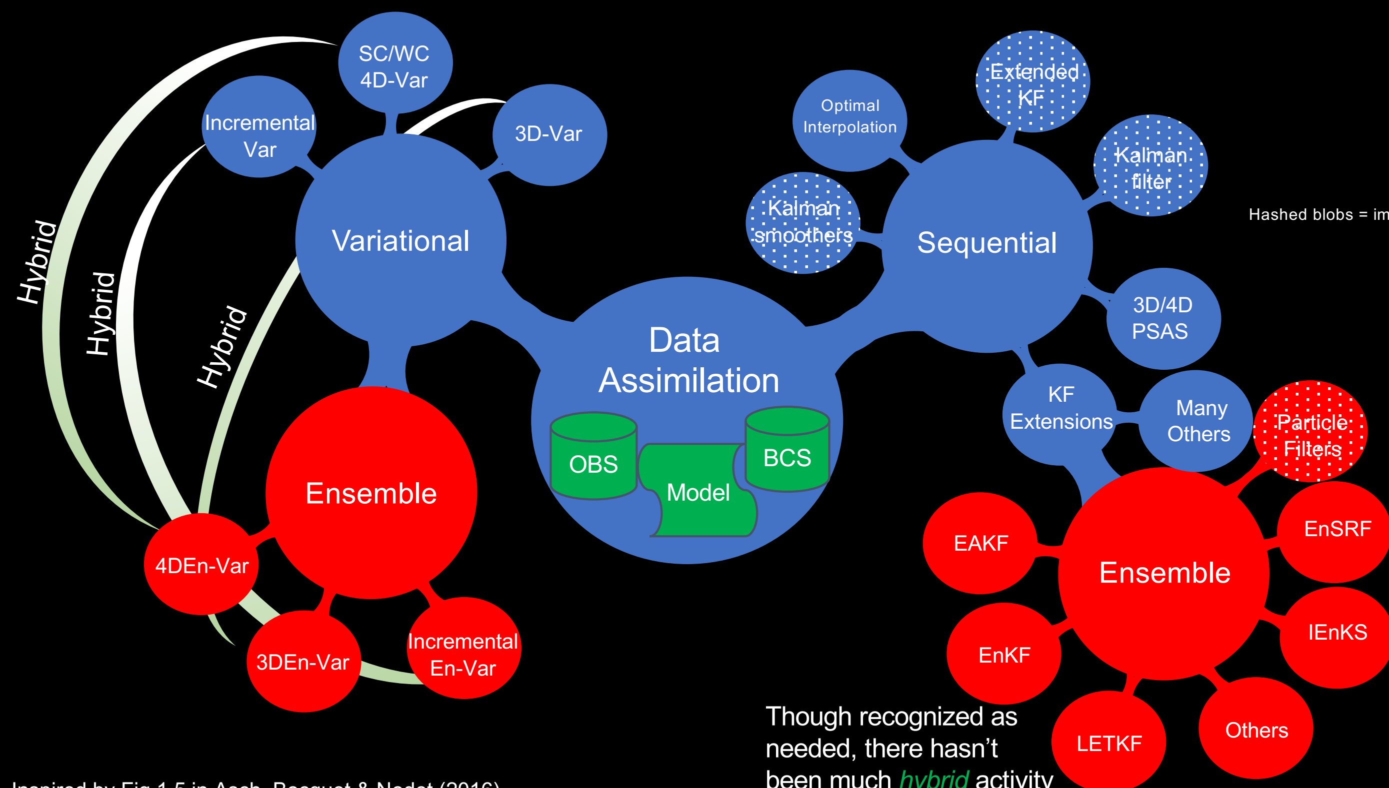
$$p(\mathbf{X}|\mathbf{O}) = \frac{p(\mathbf{O}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{O})}$$

with X and O being a time history of model states and observations over a given time interval.

Provides foundation for both sequential and variational DA frameworks.

Provides insight for hybrid DA.

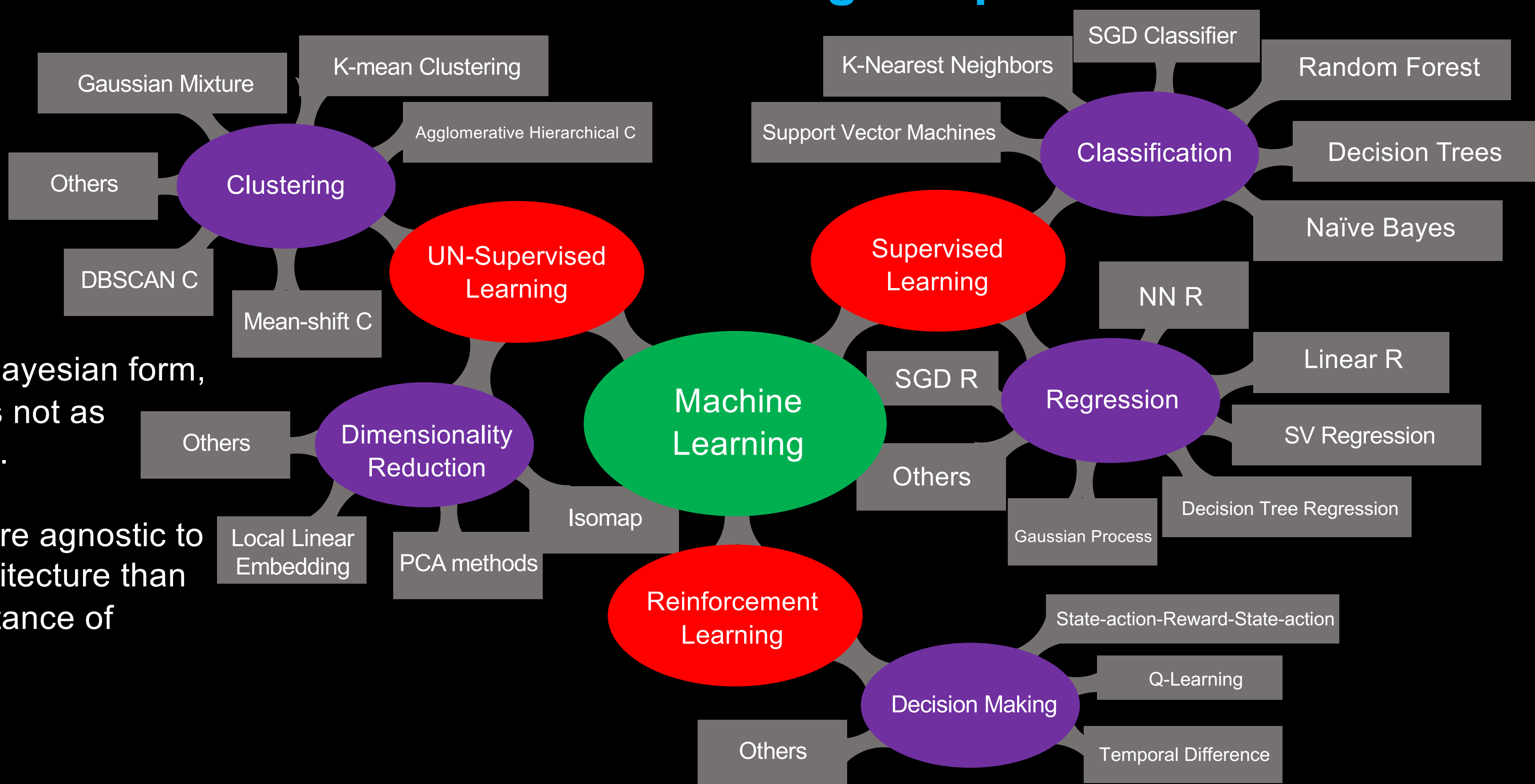
Hybrid DA combines traditional Var (or Seq) with Ensembles.



Though recognized as needed, there hasn't been much *hybrid* activity on this side of the picture.

Inspired by Fig.1.5 in Asch, Bocquet & Nodet (2016)

The Machine Learning Soup



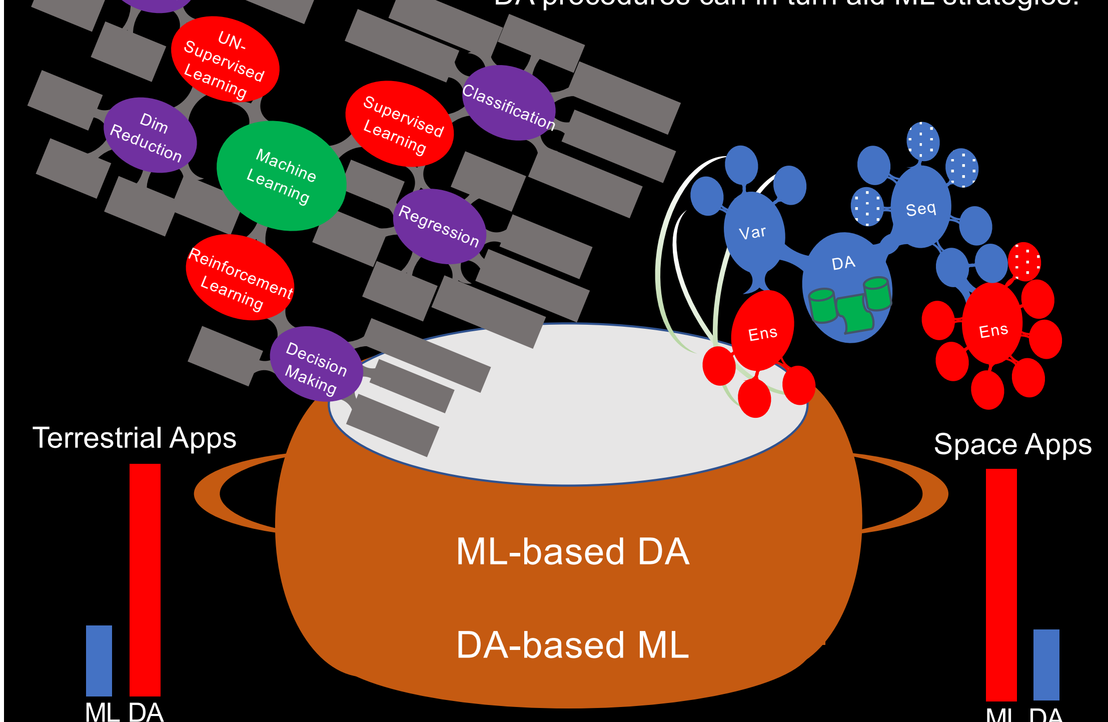
Can be put in Bayesian form, though perhaps not as clear-cut as DA.

Tends to be more agnostic to computing architecture than DA, thus importance of frameworks.

Hybrid Concept: Hybridizing the Hybrid

The past few years has seen a substantial rise in the number of proposals to get ML to aid DA.

In the process, it is also being discovered that DA procedures can in turn aid ML strategies.



See Geer (2021) for broader aspects of this symbiosis.

DA for ML

Handing of space/time sparse incomplete observations; obs operators.

Handling of noisy data.

Inference of processes indirectly related to observations.

Incorporation of prior knowledge, along with Bayesian approach.

Availability of ensembles.

Quantitative representation of uncertainties.

Uncertainty propagation.

Normalization based on physical principles (viz. background errors)

ML for DA*

Process Emulator:

- TL/AD modeling
- Transport/Dynamics
- Chemical integrator
- Obs Error Cov. Construct

Surrogates

Physical Parametrizations

Observations Retrievals

Data Homogenization

Post-processing enhancement

*DA's had its Gray-Box for some time; the Gray-Box of ML (Camporeale) exacerbates it (DA's) further.

The *Learning* Aspect of DA …

Traditional DA is a learning machine …

Adaptive DA is a self-correcting robust machine …

But an inefficient machine in many ways …

The *Learning* Aspect of DA: Terrestrial Applications
Adaptive Estimation

DA procedures have incorporated *learning* mechanisms to correct biases and uncertainties for quite a while: ADAPTIVE schemes.

- Variational Bias Correction (VarBC)

From Dee & Uppala (2009)

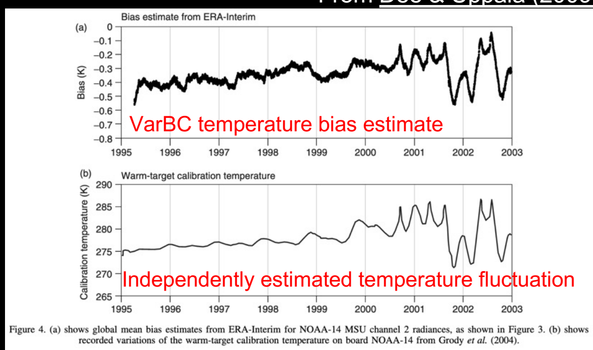
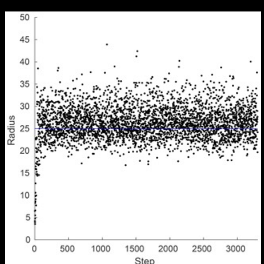


Figure 4. (a) shows global mean bias estimates from ERA-Interim for NOAA-14 MSU channel 2 radiances, as shown in Figure 3. (b) shows recorded variations of the warm-target calibration temperature on board NOAA-14 from Grody *et al.* (2004).

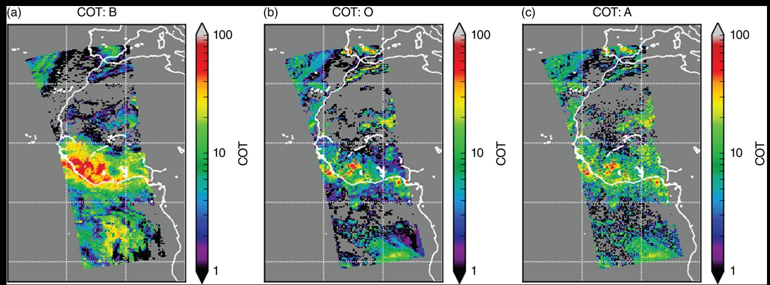
- Adaptive Error Covariance Localization



Left: Localization radii adaptively estimated for a QG model error covariance; Popov & Sandu (2019)

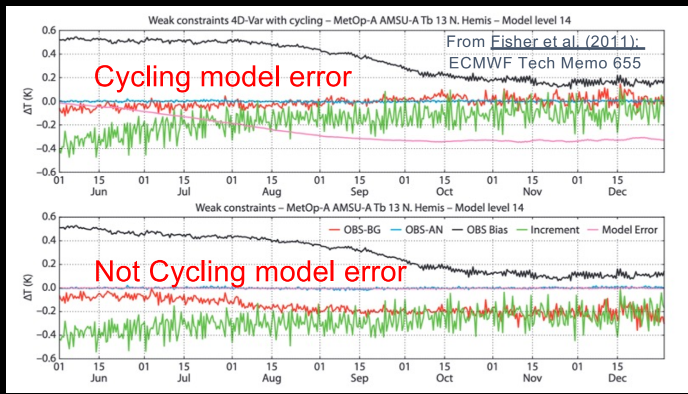
Other works have also explored procedure for adaptive inflation in ensemble DA schemes. Some have made it into SWDA literature, e.g., Godinez & Koller (2012).

- Model parameter estimation



E.g., Cloud Optical Thickness over MODIS pass: background, observations, analysis; Norris & da Silva (2016; Part II)

- Weak Constraint 4D-Var



The *Learning* Aspect of DA: Terrestrial Applications
Adaptive Estimation

DA procedures have incorporated *learning* mechanisms to correct biases and uncertainties for quite a while: ADAPTIVE schemes.

- Variational Bias Correction (VarBC)

From Dee & Uppala (2009)

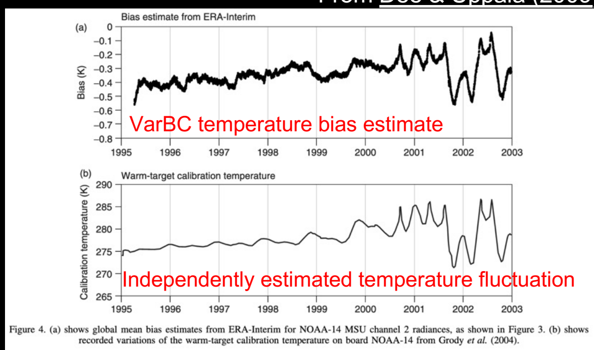
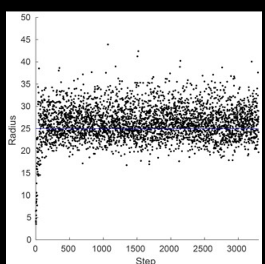


Figure 4. (a) shows global mean bias estimates from ERA-Interim for NOAA-14 MSU channel 2 radiances, as shown in Figure 3. (b) shows recorded variations of the warm-target calibration temperature on board NOAA-14 from Grody et al. (2004).

- Adaptive Error Covariance Localization



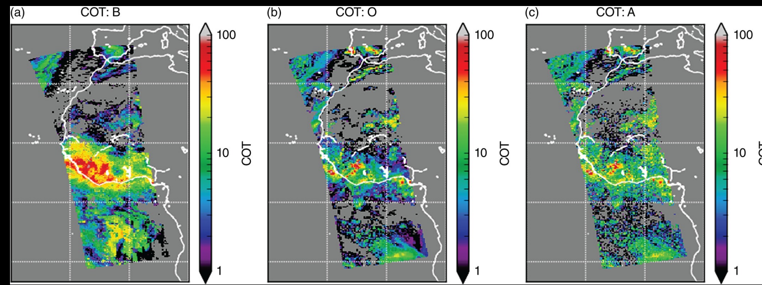
Left: Localization radii adaptively estimated for a QG model error covariance; Popov & Sandu (2019)

Other works have also explored procedure for adaptive inflation in ensemble DA schemes. Some have made it into SWDA literature, e.g., Godinez & Koller (2012).

All these procedure Error Estimates

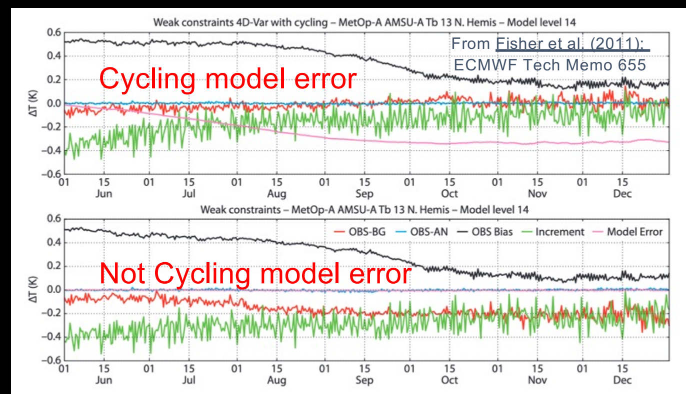
AKA: Uncertainty Quantification

- Model parameter estimation



E.g., Cloud Optical Thickness over MODIS pass: background, observations, analysis; Norris & da Silva (2016; Part II)

- Weak Constraint 4D-Var



Ensemble-derived for DA & UQ …

Introduced to address nonlinearities & efficiency …

Also good for UQ …

Uncertainty Quantification (UQ): Ensemble DA

Observations (O)

Model

Constrained (X)
Unconstrained (Z)

Bayesian Applications

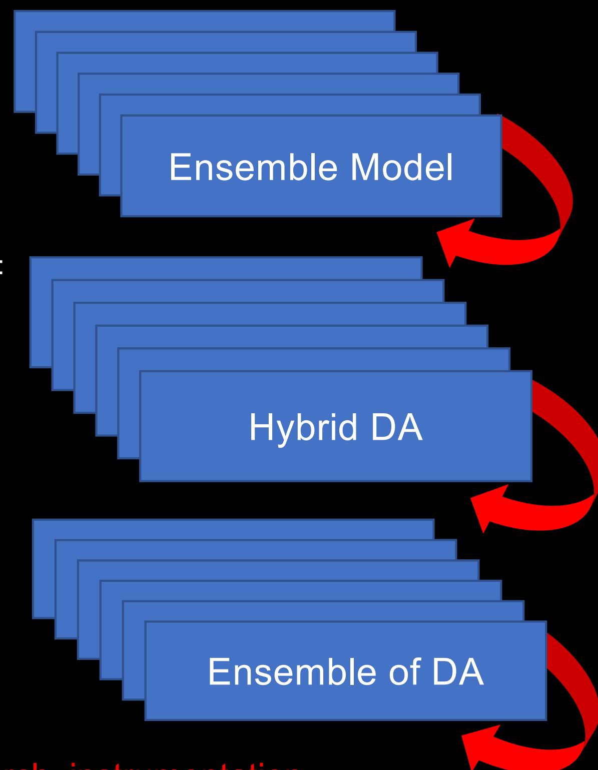
$$p(\mathbf{X}|\mathbf{O}) = \frac{p(\mathbf{O}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{O})}$$

In this perspective there are two levels of uncertainty required to make DA work, those being in the:

Observations: $\mathbf{O} \rightarrow \mathbf{R}$

Model: $\mathbf{X} \rightarrow \mathbf{B}, \mathbf{Q}$

Straightforward DA and ML do not provide UQ on Z, but MC-based do.



Monte Carlo-like DA

- U from Adaptive DA
- U Propagation (X,Z)

- Probabilities (Fair) Scores
- MC / Ensemble Learning
- ML for obs error covariance enhancement (LSTM) RNN.
- ML/WC for model error estimation.

- ML to improve Predictions (offline)
- ML for Downscaling.

Here understood as:

Constrained Vars: variables directly affected by DA/ML;
can include model parameters.
Unconstrained Vars: everything else derived from a model.

UQ of U-Var are desirable in many areas: climate research, instrumentation, risk analysis, improved DA/ML methodologies, validation, etc.

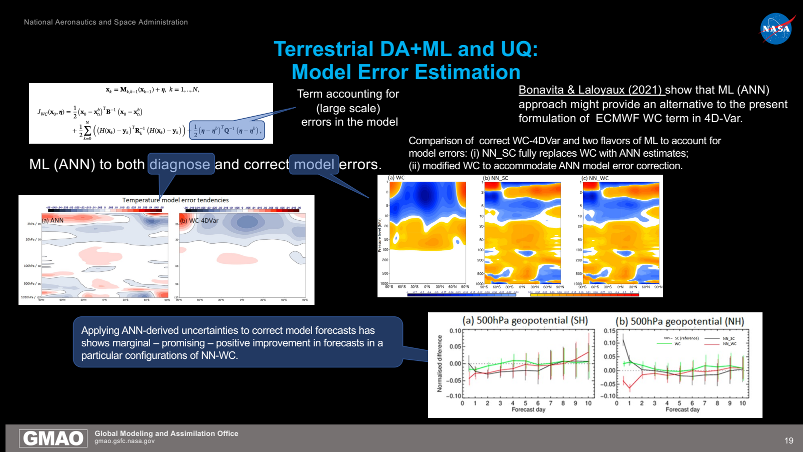
Did anybody say Digital Twin?

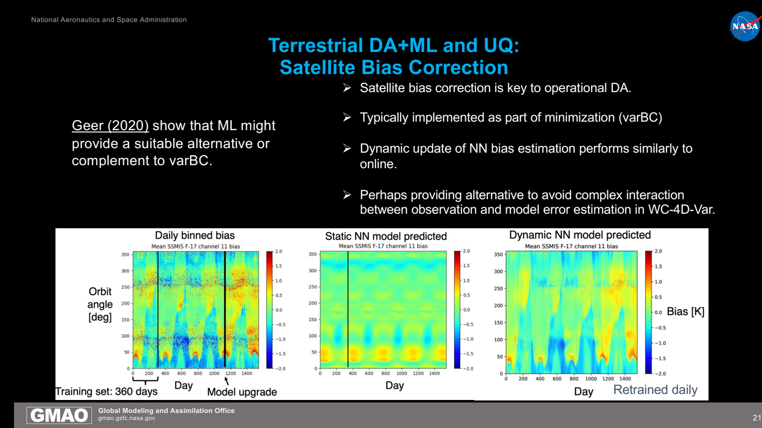
ML as Aiding Device for DA …

Alternative ways to derive uncertainties …

Another tool to address efficiency …

ML to Aid DA: Terrestrial UQ in DA



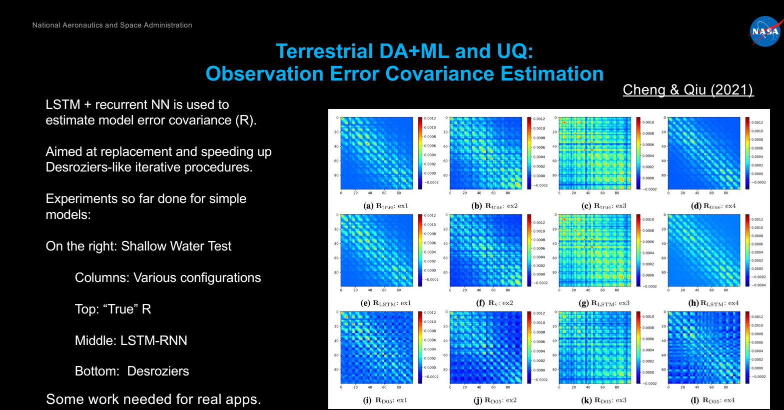


Top Left: ANN for WC-4D-Var model error estimation

Top Right: Dynamic NN for satellite bias estimation

Bottom Right: LSTM-Recurrent NN for obs error cov

Full slides in Appendix



Frameworks …

The way for community collaboration …

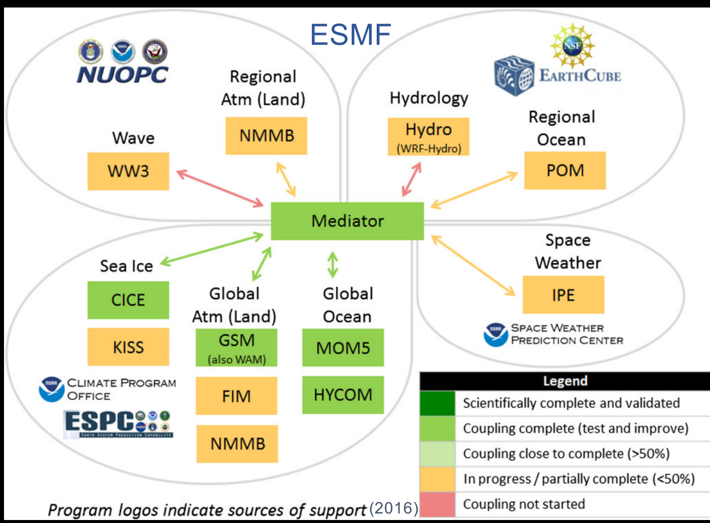
Facilitating R2O & O2R …

Facilitating rapid deployment of science …

Communities Modeling Frameworks

Terrestrial Weather Modeling

ESPS – Earth System Prediction Suite



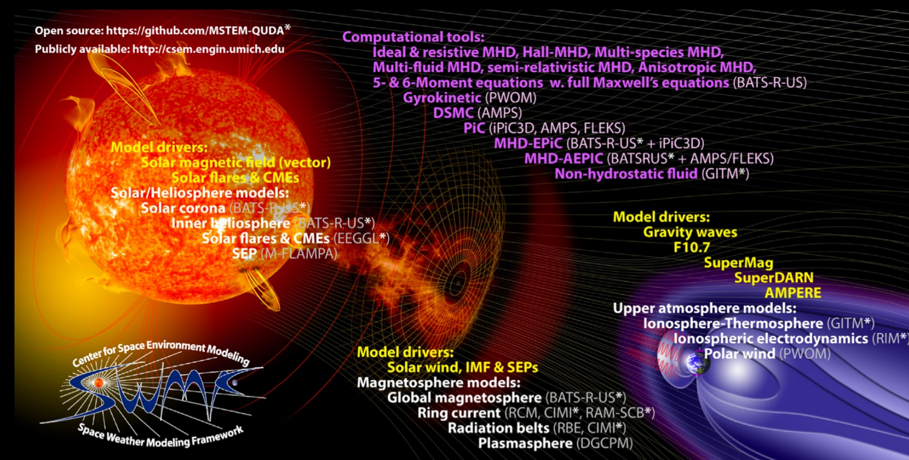
From Theurich et al. (2016)

ESMF – Earth System Modeling Framework

NUOPC - National Unified Operational Prediction Capability

Space Weather Modeling

SWMF - Space Weather Modeling Framework



From Gombosi et al. (2021)

Is a modeling framework adequate for DA?

The Terrestrial community has had this discussion and it has largely decided to answer NO.

Remarks on Frameworks for DA & ML

ML applications have condensed into general, portable, Phyton/C++-based libraries that have been globally-embraced by the user community and has facilitated continual development and enhancement data analytics, e.g.: Keras, SciKit-Learn, TensorFlow, PyTorch.

DA applications have been slow to condense into a "globally-embraced" framework. The past few years has seen the community reach a consensus on the *need* for a framework; competing frameworks exist at present:

- EMPIRE (Reading, UK)
- DART (NCAR, USA)
- PDAF (AWI, DE)
- OpenDA (TU Delft, NL)
- JEDI (JCSDA, USA)

Terrestrial DA Truly Opensource Frameworks

Space DA Opensource? Frameworks?

- ADAPT (Air Force, LANL, USA)
- DREAM (LANL, USA)
- Is there need for a specific SWDAF? Probably NOT!
- NSF/NASA on Frameworks 2020

It has been acknowledged that these frameworks must interface with ML software (e.g., JEDI).

JEDI is the framework for DA development adopted by NOAA, NASA, US Navy, US Air Force, and others. JEDI is not yet operational, but schedules are set on that. The U.K. Met Office is also committed to JEDI.

It might be helpful for the SW-DA-ML community to embrace existing DA frameworks; in the USA, JEDI.

Closing Thoughts

- Data Assimilation:

  - Can be viewed as a *traditional* machine learning device.
  - Adaptive procedures render DA self correcting & robust.

But …traditional DA is hard to implement, maintain, and inefficient:

- Ensemble techniques are fundamental to address part of such issues & provide path to UQ.

- Modern ML techniques allow for further improvement of DA through:
  - Surrogate modeling.
  - Covariance estimation.
  - Characterization of uncertainties.

- DA Frameworks should allow for agile R2O2R & to keep up with Exascale endeavors.

Thank you

Parallels in the Challenges in DA and ML

Possible points for further discussion

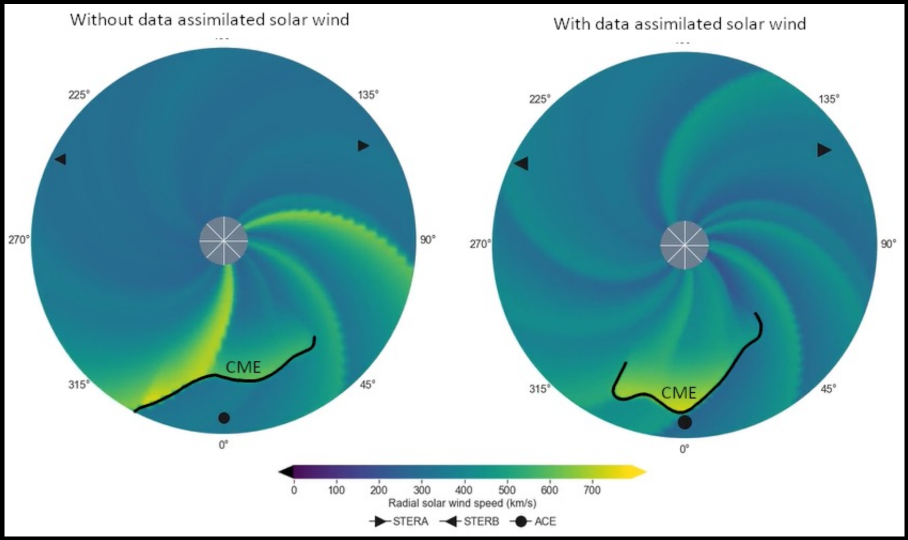
Challenges in DA	Challenges in ML
➢ Construction of proper physical models.	➢ Generation of representative training set.
➢ Adapt to new information (varBC, WC).	➢ Changed scenarios requiring retraining.
➢ Enforcement of conservation properties.	➢ Enforcement of conservation properties.
➢ Big data.	➢ Big data.
➢ UQ for constrained and unconstrained variables.	➢ UQ for constrained and unconstrained variables.
➢ Scalability to Exascale Computing	➢ Scalability to Exascale Computing
	➢ Use of prior knowledge
	➢ Physically-based normalization

- There is considerable similarity between these challenges
- Some might be more mature on one side than the other.

Backup Slides

Data Assimilation for CME Prediction

- U.K. Met Office prediction arrives 10 hours before*.
- Without , DA CME arrives on Earth 41 hours late.
- With DA, CME arrives 19 min before event.
- DA settings:
  - Simplified Solar Wind
  - Obs: STEREO @ 5-hr
  - 576-member VarDA
  - Update Inner BC
  - ACE for verification



See: M. Lang Space Weather Forecast Blog, May 2021
Also: Lang & Owens (2019) and Barnard et al. (2020)

*right at error margin for 3D-MHD models, per Riley et al. (2018)

Authors

- Move from LETKF to Var to allow for Update of Inner BC
- Hope that localization issue associated with wind corrections in LETKF can be solved and EnKS can be used to avoid Adjoint in Var.

Comments (here)

Alternatively, ML can be used to train ANN as Adjoint replacement

Terrestrial DA+ML and UQ: Model Error Estimation

$$\mathbf{x}_k = \mathbf{M}_{k,k-1}(\mathbf{x}_{k-1}) + \boldsymbol{\eta},\ k = 1,..,N,$$

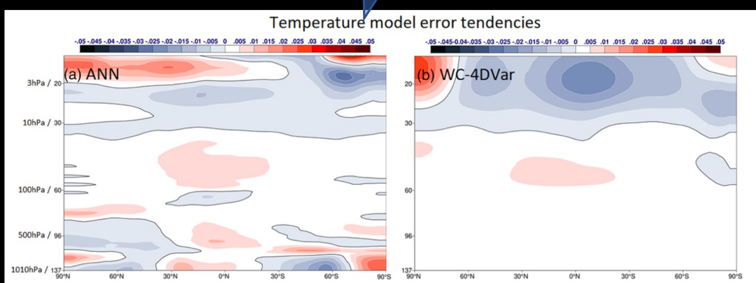
$$J_{WC}(\mathbf{x}_0, \boldsymbol{\eta}) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{k=0}^{N}\left((H(\mathbf{x}_k) - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(H(\mathbf{x}_k) - \mathbf{y}_k)\right) + \boxed{\frac{1}{2}(\eta - \eta^b)^T \mathbf{Q}^{-1}(\eta - \eta^b)},$$

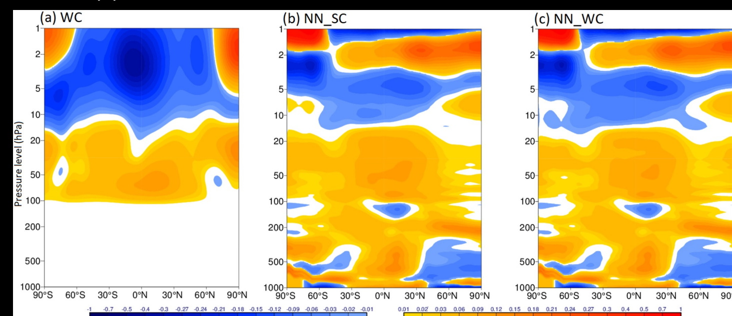
Term accounting for (large scale) errors in the model

Bonavita & Laloyaux (2021) show that ML (ANN) approach might provide an alternative to the present formulation of ECMWF WC term in 4D-Var.

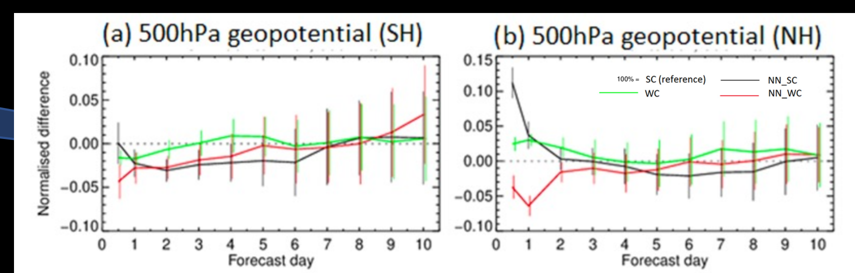
ML (ANN) to both diagnose and correct model errors.

Comparison of correct WC-4DVar and two flavors of ML to account for model errors: (i) NN_SC fully replaces WC with ANN estimates; (ii) modified WC to accommodate ANN model error correction.





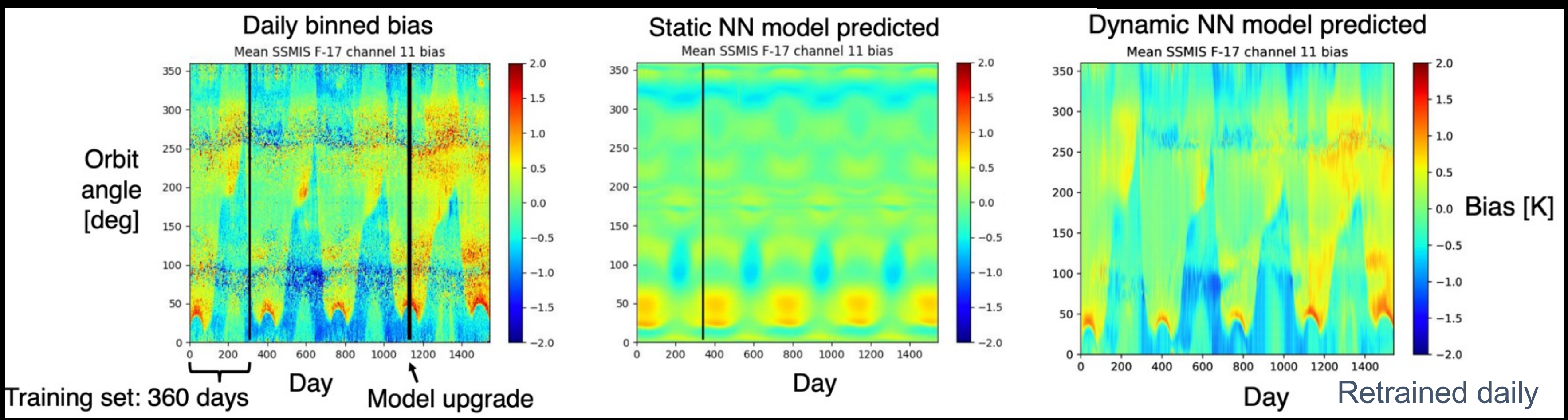
Applying ANN-derived uncertainties to correct model forecasts has shows marginal – promising – positive improvement in forecasts in a particular configurations of NN-WC.



Terrestrial DA+ML and UQ: Satellite Bias Correction

Geer (2020) show that ML might provide a suitable alternative or complement to varBC.

- Satellite bias correction is key to operational DA.
- Typically implemented as part of minimization (varBC)
- Dynamic update of NN bias estimation performs similarly to online.
- Perhaps providing alternative to avoid complex interaction between observation and model error estimation in WC-4D-Var.



Terrestrial DA+ML and UQ: Observation Error Covariance Estimation

Cheng & Qiu (2021)

LSTM + recurrent NN is used to estimate model error covariance (R).

Aimed at replacement and speeding up Desroziers-like iterative procedures.

Experiments so far done for simple models:

On the right: Shallow Water Test

- Columns: Various configurations
- Top: “True” R
- Middle: LSTM-RNN
- Bottom: Desroziers

Some work needed for real apps.