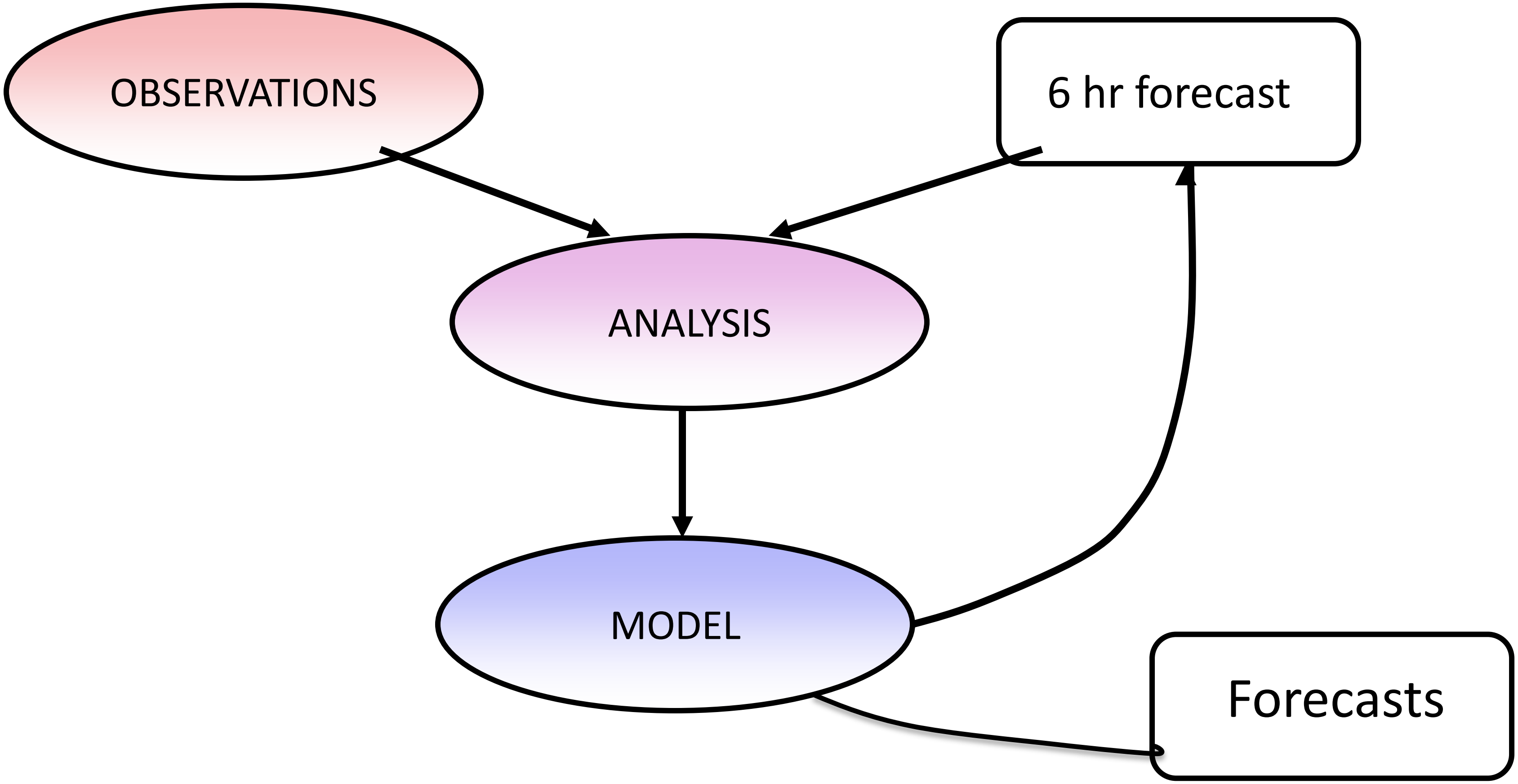
New Applications of Advanced Data Assimilation to Improve Models and Observations

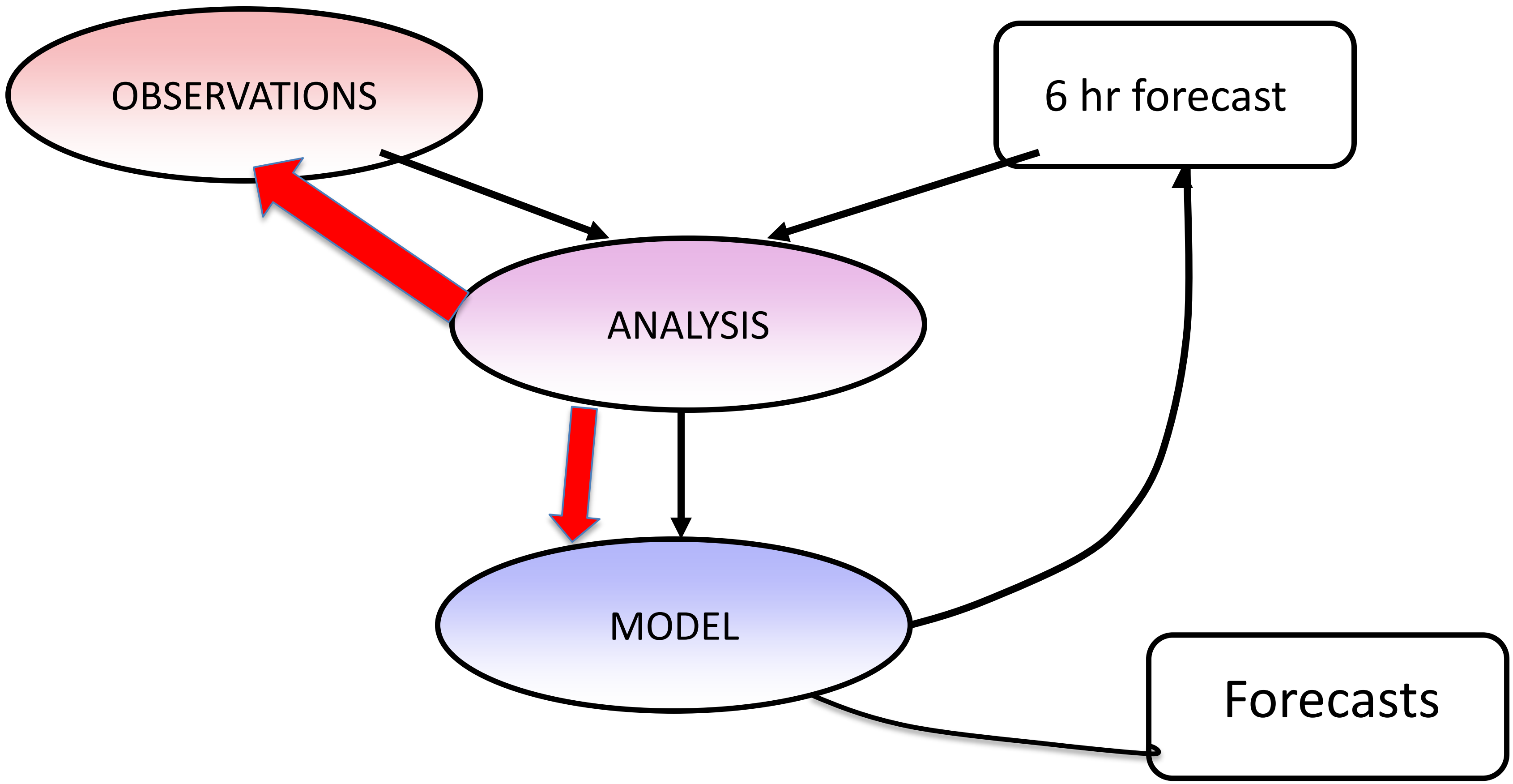
E. Kalnay, T. C. Chen, D. Hotta, Y. Ota, T. Miyoshi, Kriti Bhargava, J. Carton, T. Sluka, S.Penny

[1] UMD, [2]JMA, [3]RIKEN, [4]NCEP
With many thanks to students, friends and colleagues
from the University of Maryland

Classic Data Assimilation: For NWP we need to improve observations, analysis scheme and model



New Data Assimilation: We can also use DA to improve observations and model



The simplicity and power of EnKF should encourage the use of DA for improvements beyond its main goal

Combine optimally observations and model forecasts (mostly done! ☺)

- We should also use DA to:

Improve the observations

Improve the model

- Improve the models by parameter estimation

Example: Estimate the surface carbon fluxes as evolving parameters.

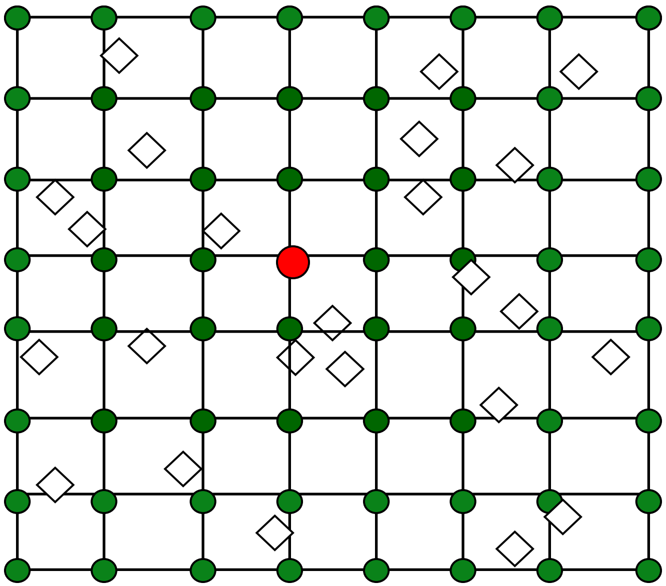
- Earth system models used by IPCC have many sub-models, but they don't include the Human System, which totally dominates the Earth system.

We should do DA of the <u>two-way coupled</u> Earth System-Human System, and use DA for parameter tuning

LETKF: Localization based on observations

Perform data assimilation in a local volume, choosing observations

The state estimate is updated at the central grid red dot

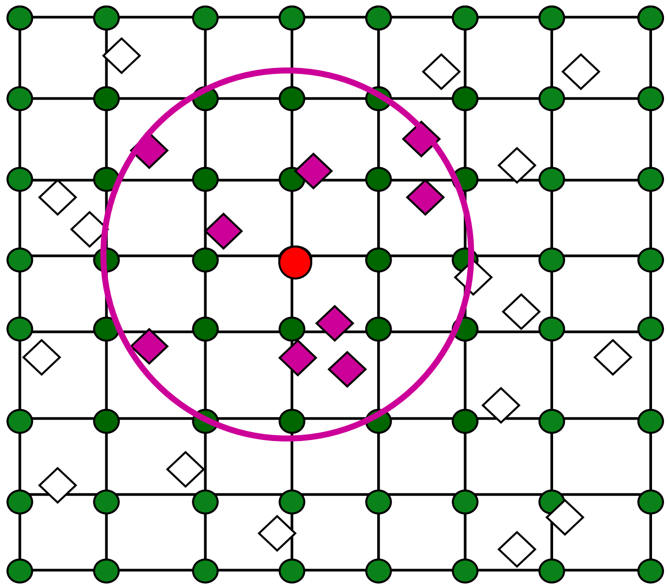


LETKF: Localization based on observations

Perform data assimilation in a local volume, choosing observations

The state estimate is updated at the central grid red dot

All observations (purple diamonds) within the local region are assimilated



The LETKF algorithm can be described in a single slide!

Local Ensemble Transform Kalman Filter (Hunt et al, 2007)

Globally:

Forecast step: $$\mathbf{x}^b_{n,k} = M_n\left(\mathbf{x}^a_{n-1,k}\right)$$

Analysis step: construct $$\mathbf{X}^b = \left[\mathbf{x}^b_1 - \overline{\mathbf{x}}^b \,|\, \ldots \,|\, \mathbf{x}^b_K - \overline{\mathbf{x}}^b\right];$$

$$\mathbf{y}^b_i = H(\mathbf{x}^b_i);\ \mathbf{Y}^b_n = \left[\mathbf{y}^b_1 - \overline{\mathbf{y}}^b \,|\, \ldots \,|\, \mathbf{y}^b_K - \overline{\mathbf{y}}^b\right]$$

Locally: **Choose for each grid point the observations to be used, and compute the local analysis error covariance and perturbations in ensemble space:**

$$\tilde{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + \mathbf{Y}^T\mathbf{R}^{-1}\mathbf{Y}\right]^{-1};\ \mathbf{W}^a = [(K-1)\tilde{\mathbf{P}}^a]^{1/2}$$

Analysis mean in ensemble space: $\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a\mathbf{Y}^{bT}\mathbf{R}^{-1}(\mathbf{y}^o - \overline{\mathbf{y}}^b)$

and add to $\mathbf{W}^a$ **to get the analysis ensemble in ensemble space.**

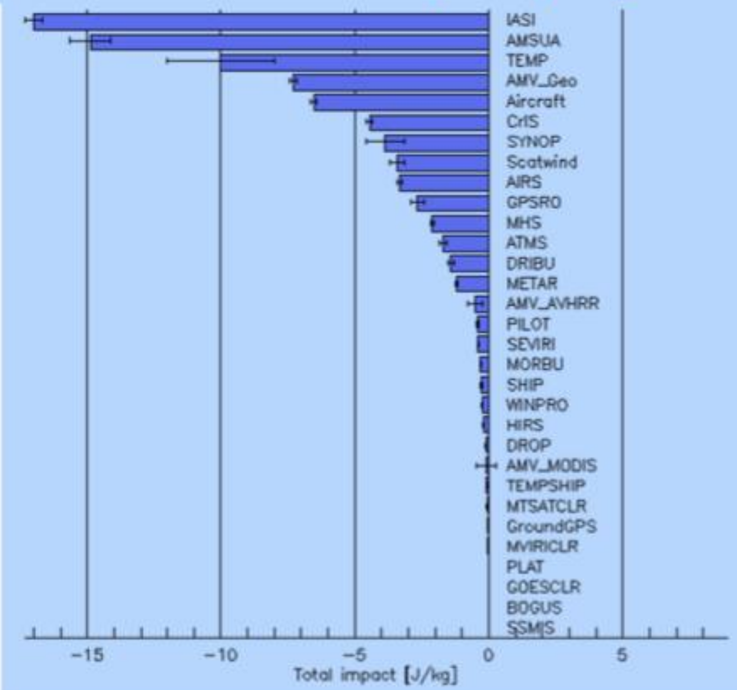
The new ensemble analyses in model space are the columns of $\mathbf{X}^a_n = \mathbf{X}^b_n\mathbf{W}^a + \overline{\mathbf{x}}^b$**. Gathering the grid point analyses forms the new global analyses. Note that the the output of the LETKF are analysis weights** $\overline{\mathbf{w}}^a$ **and perturbation analysis matrices of weights** $\mathbf{W}^a$**. These weights multiply the ensemble forecasts.**

Forecast Sensitivity to Observations (Langland and Baker, 2004)

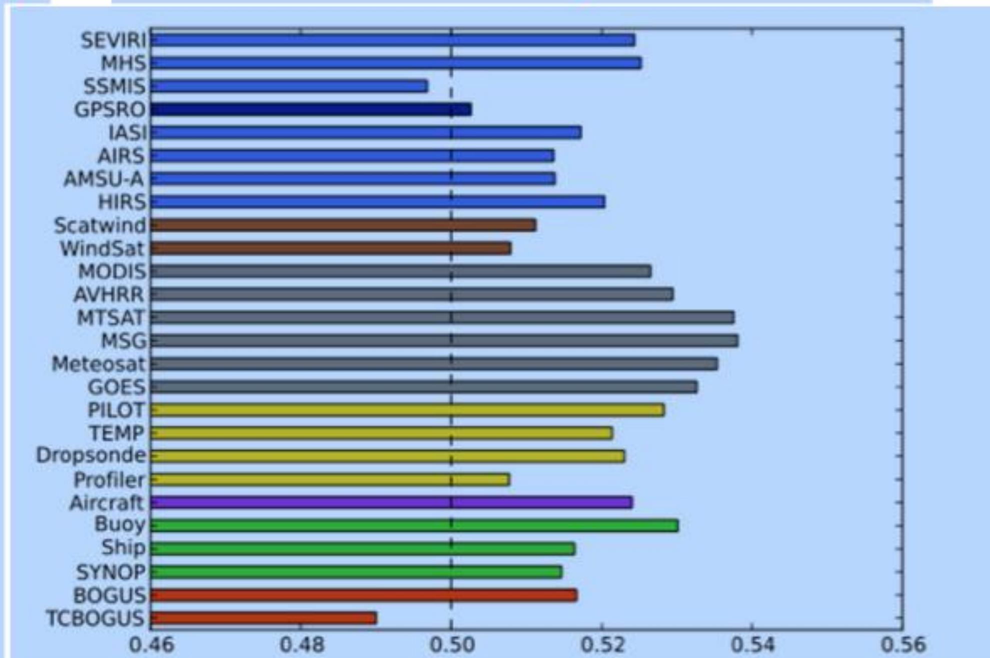
Met Office

FSOI in Global NWP

Total Observation Impact (Aug 2014)



Fraction Obs that Improve Forecast



- Infra-Red (IASI) and microwave (AMSUA) radiances now biggest impact.
- Note only ~50% of observations reduce forecast error(!).
- Estimate: need 6 months time series to assess impact for single observing site.
- EFSO methodology now being considered when no adjoint available

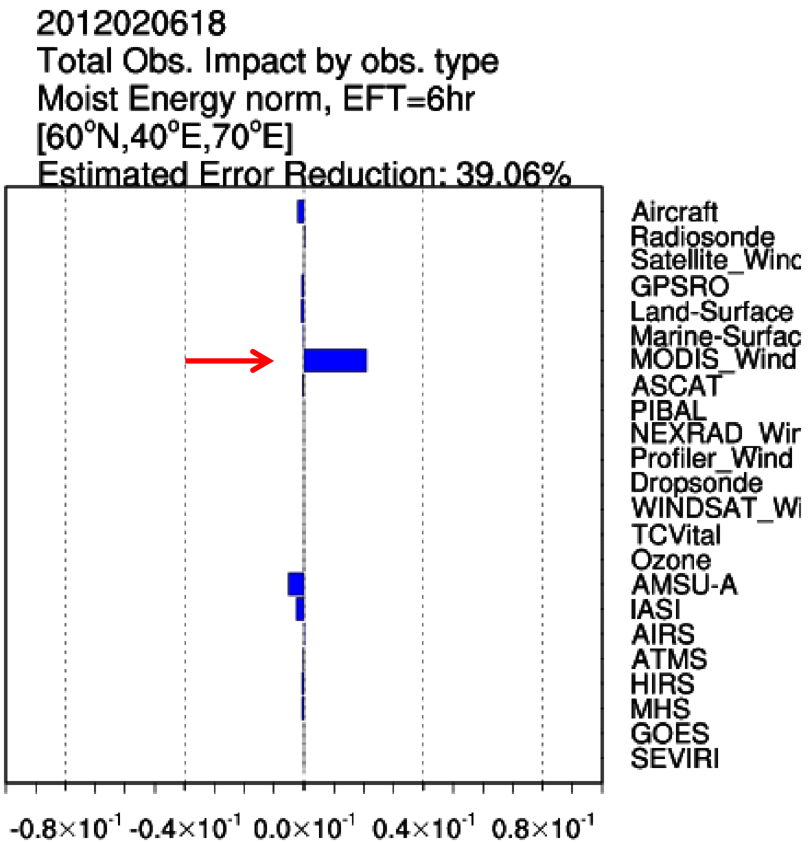
1) Improve the observations: Ensemble Forecast Sensitivity to Observations and Proactive QC

- Kalnay et al. (2012) derived EFSO.
- Ota et al. (2013) tested 24hr GFS forecasts and showed EFSO could be used to identify bad obs.
- D. Hotta (2014): EFSO can be used after only 6 hours, so that the bad obs. can be collected and withdrawn, with useful metadata, so they can be improved. The analysis is corrected with EFSO.
- We call this **Proactive QC**, much stronger than QC.
- Hotta also showed EFSO can be used to tune **R**
- Tse-Chun Chen tested impact of EFSO/PQC over 5 day forecasts: **VERY PROMISING RESULTS**

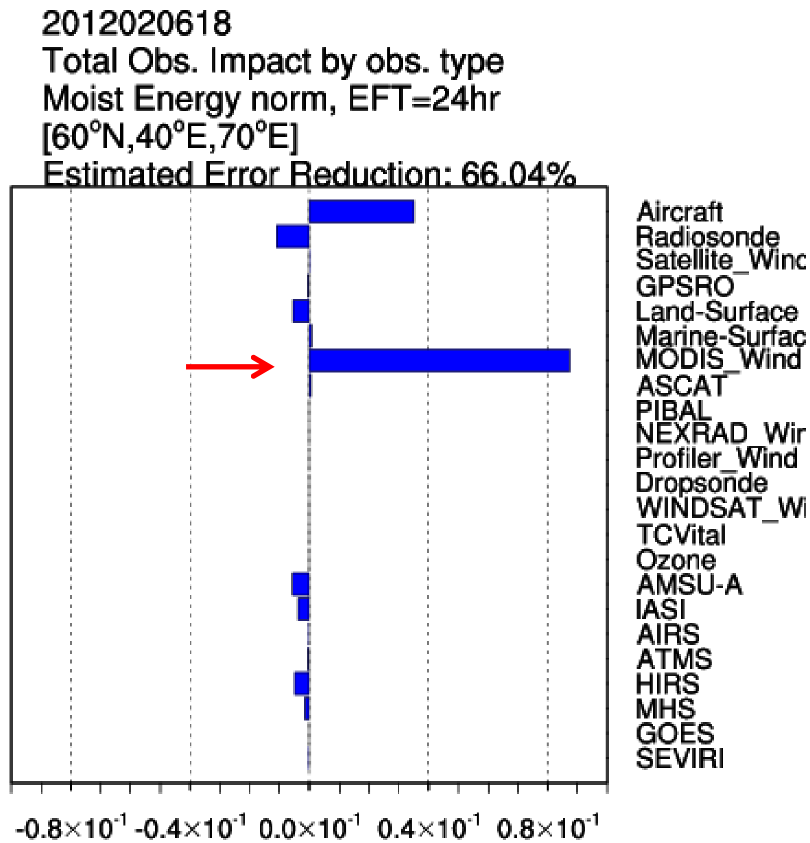
Hotta (2014)

Feb. 18 06UTC, near the North Pole (Ota et al. 2013 case). Bad obs: MODIS WINDS

FT=06 hr.



FT=24 hr.

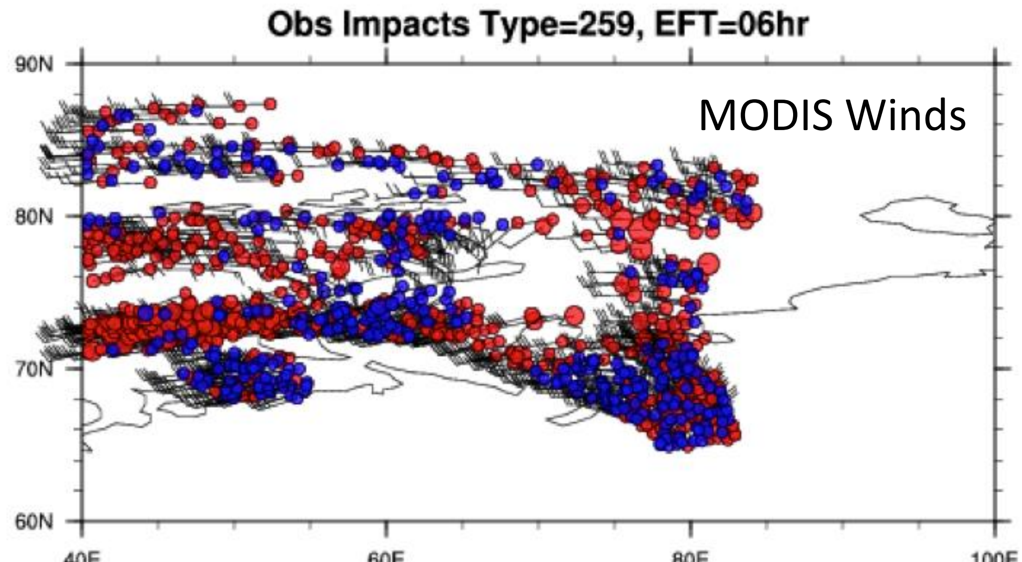


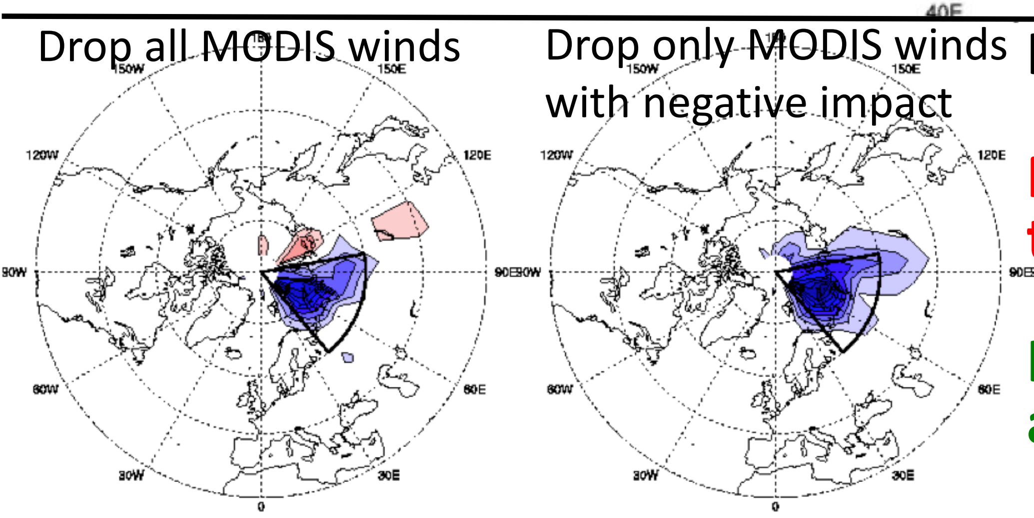
Can identify the bad observations after only 6 hours!

Improve observations:

Proactive QC: Find and delete the obs that make the 6hr forecast worse using EFSO

Dr. Daisuke Hotta (2014):
EFSO is able to find whether <u>each</u> observation improves (blue) or makes the 6hr forecast worse (red)





Impact of 6hr PQC on 24hr fcst

PQC with metadata can be used to improve the algorithm!

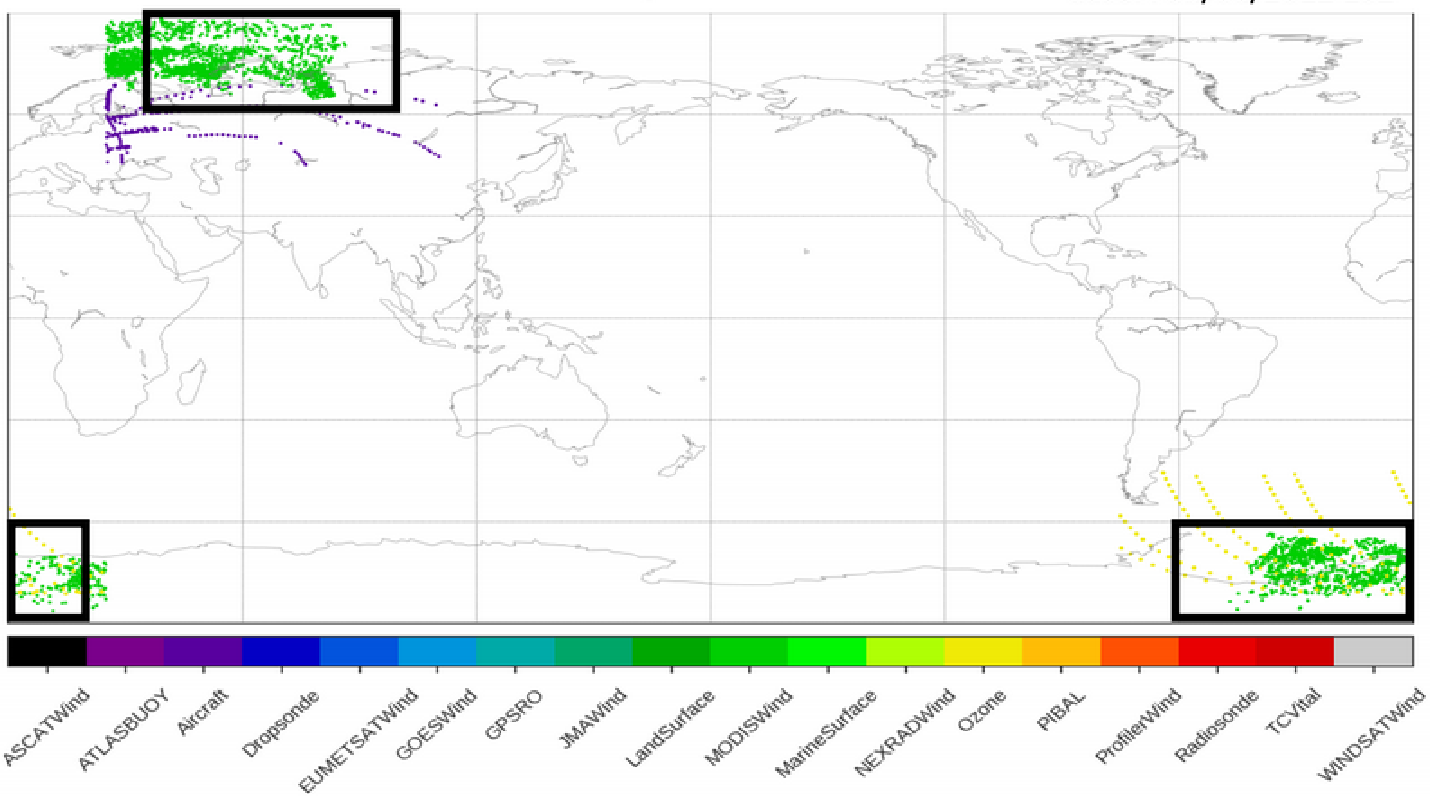
It should accelerate optimal assimilation of new instruments!

Three Data Denial Experiment Methods

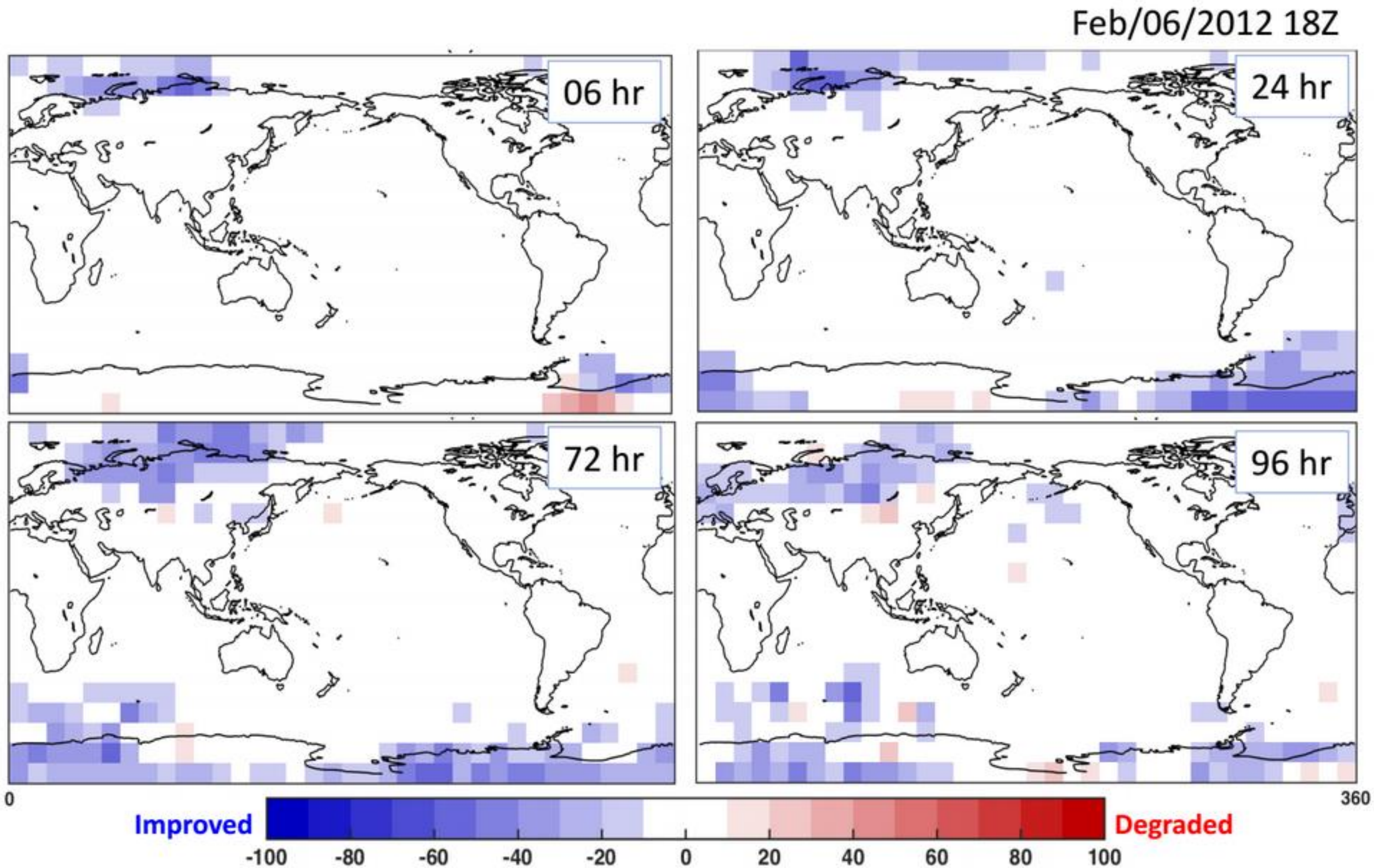
1. **Hotta** (from Hotta 2016 and Ota 2013)

- Identify **forecast error degradation regions**
- Perform EFSO w.r.t. those regions for 6-hr impact
- Reject detrimental observations only from the systems that have net detrimental impact.

Case: Feb/06/2012 18Z

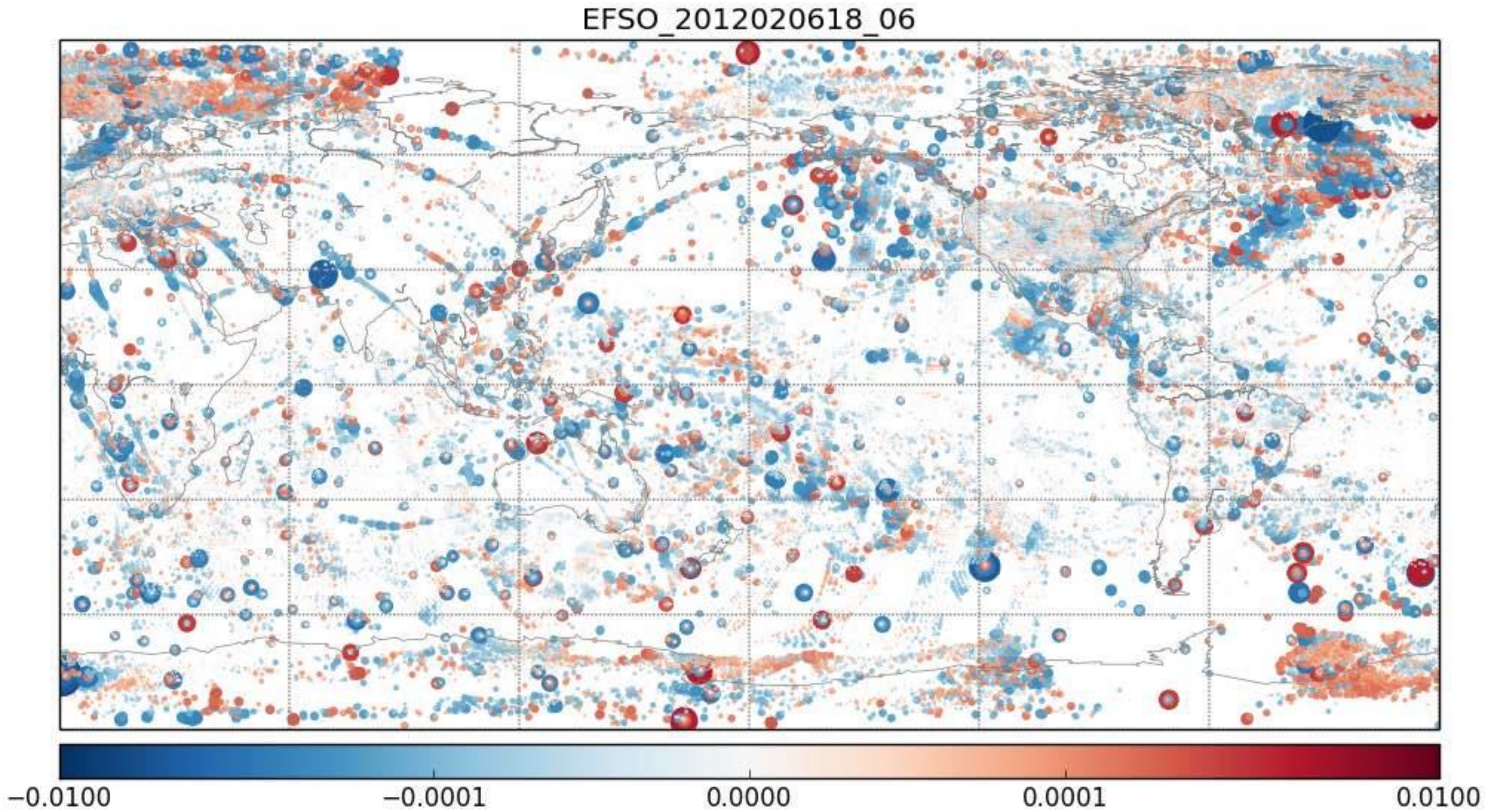


Hotta Method: Impact on the Forecasts



Improved regions **strengthen** and **propagate** with weather system

Tse-Chun Chen: new approach



EFSO applied to all observations: red – detrimental, blue – beneficial. **Threshold: Red obs withdrawn if EFSO>10^{-5}J/Kg**

Three Data Denial Experiment Methods

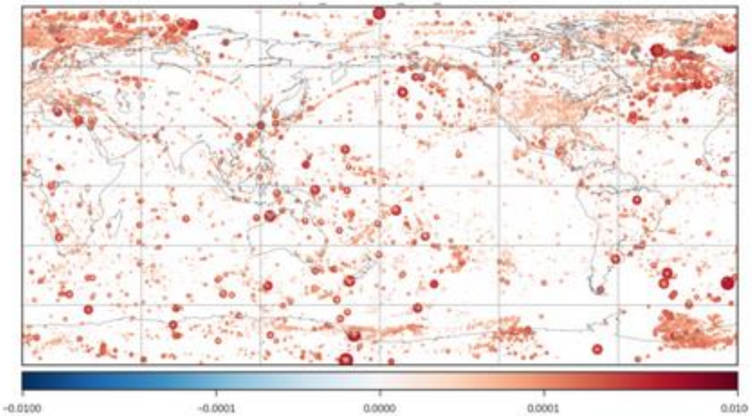
2. Threshold

- Compute global EFSO for 06-hr impact of each observation
- Reject detrimental observations with a positive (detrimental) impact larger than a 10^-5 (J/kg) threshold.

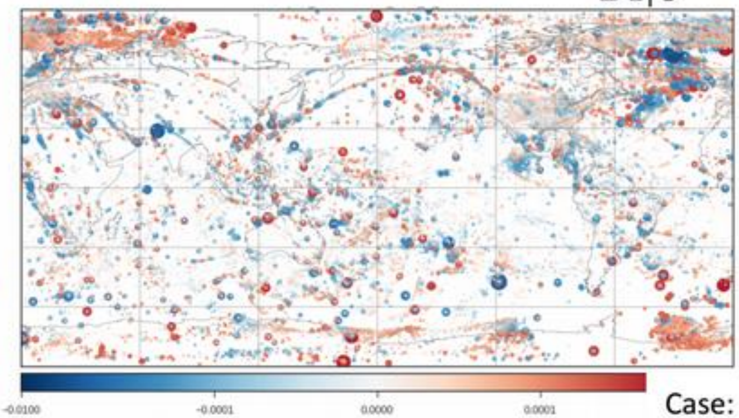
3. Assimilation in Unstable Subspace (AUS; reanalysis)

- First introduced in Trevisan (2010) with 4D-Var
- Compute the global EFSO for 06, 24-hr impact
- Assimilate only in the **beneficial growing subspace**:

$$\Delta e^2_{24|0} < \Delta e^2_{6|0} < 0$$



Threshold: 37951 rejected

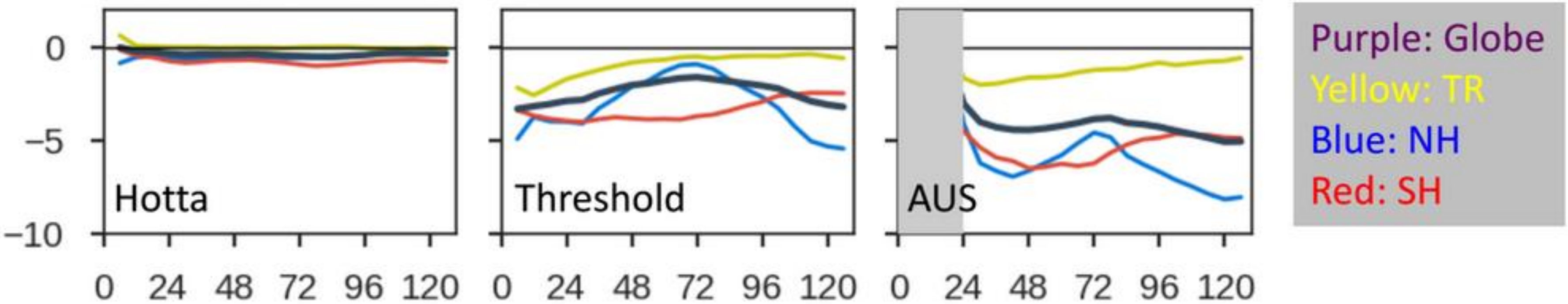


AUS: 287289 rejected

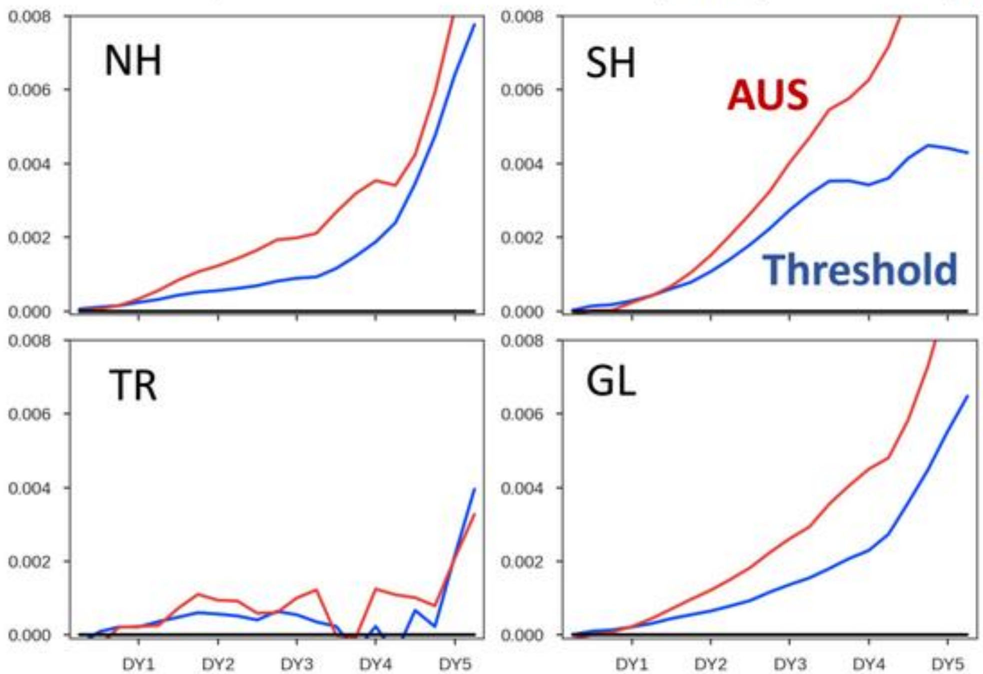
Case: Feb/06/2012 18Z
Color: 06hr MTE impact (J/kg)
Size: Magnitude of impact

Offline Experiment: 18 cases

MTE relative improvement (%)



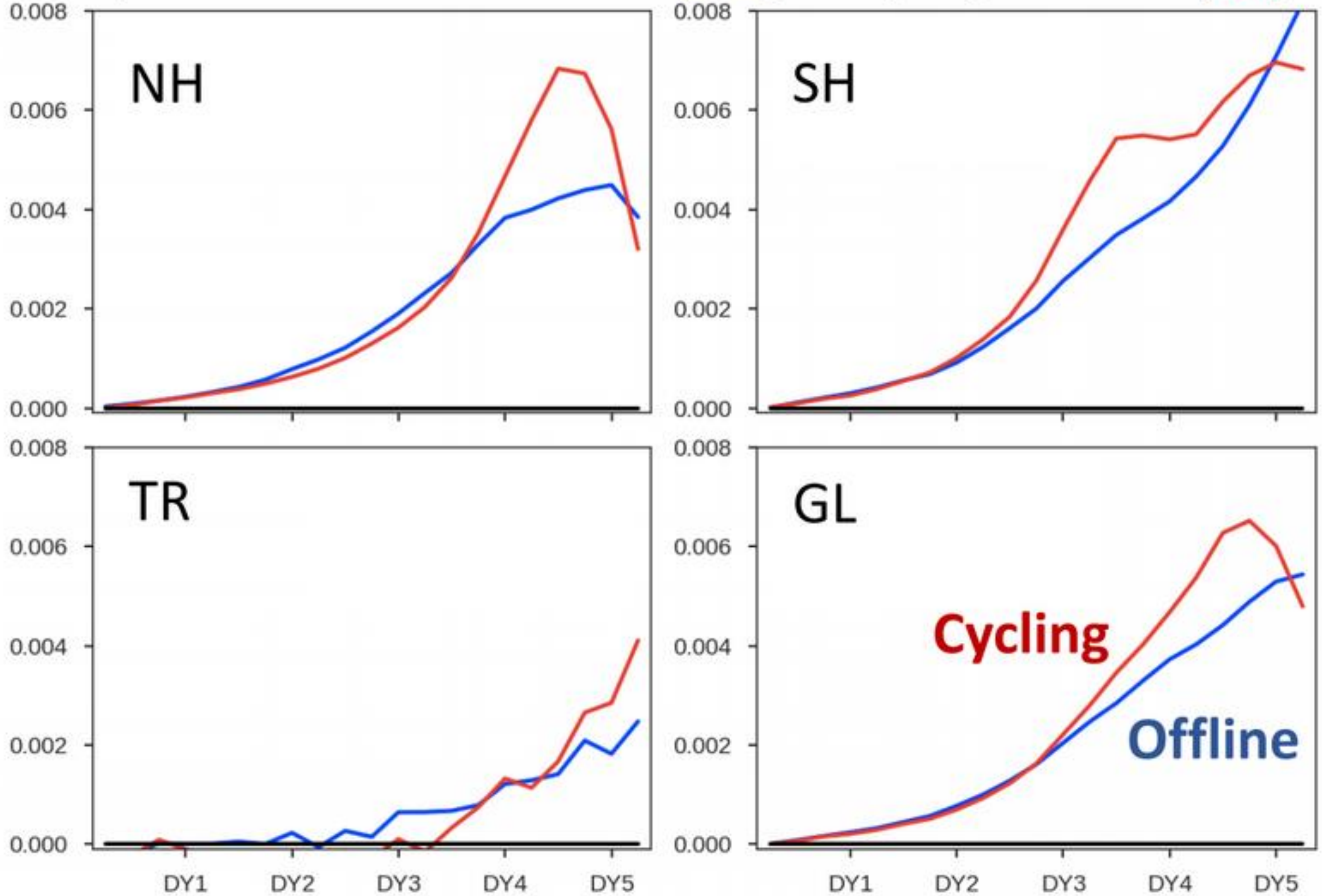
Z500 ACC Improvement: Threshold (blue) v.s. AUS (red):



- PQC corrects analysis and the subsequent forecast.
- All three methods improves model forecasts on average.
- The **AUS** and **Threshold** method have forecast improvements larger than **Hotta** method.

Cycling PQC Experiment: 10 days

Z500 ACC Improvement: Offline Threshold (blue) v.s. Cycling Threshold (red)



Improvement by cycling PQC maximizes around 3-5 day forecasts by accumulated beneficial effect of past PQCs.

Operational Implementation

Using GFS early analysis saves 3 hours of waiting.

Estimated PQC correction using **same Kalman gain K:**

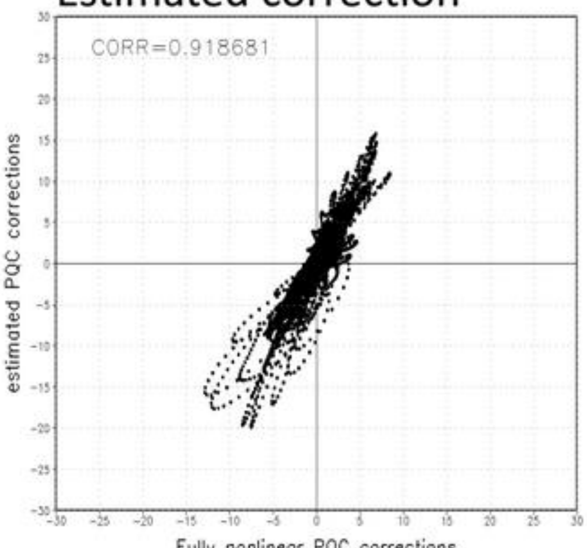
- K is actually depending on H, which determines by observations

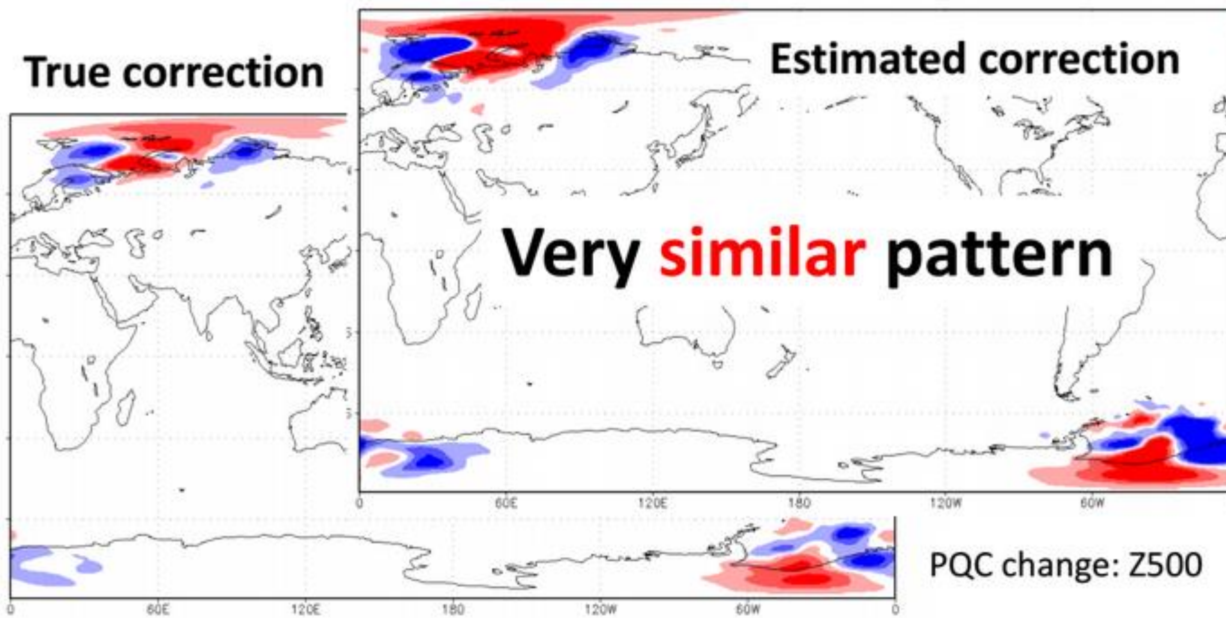
$$\bar{\mathbf{x}}_0^{a,\mathrm{deny}} - \bar{\mathbf{x}}_0^{a} \approx -\mathbf{K}\delta\bar{\mathbf{y}}_0^{ob,\mathrm{deny}}$$

$$\mathbf{K} \approx \frac{1}{K-1}\mathbf{X}_0^a\mathbf{X}_0^{aT}\mathbf{H}^T\mathbf{R}^{-1} \approx \frac{1}{K-1}\mathbf{X}_0^a\mathbf{Y}_0^{aT}\mathbf{R}^{-1}$$

(Hotta, 2016)

True correction v.s.
Estimated correction





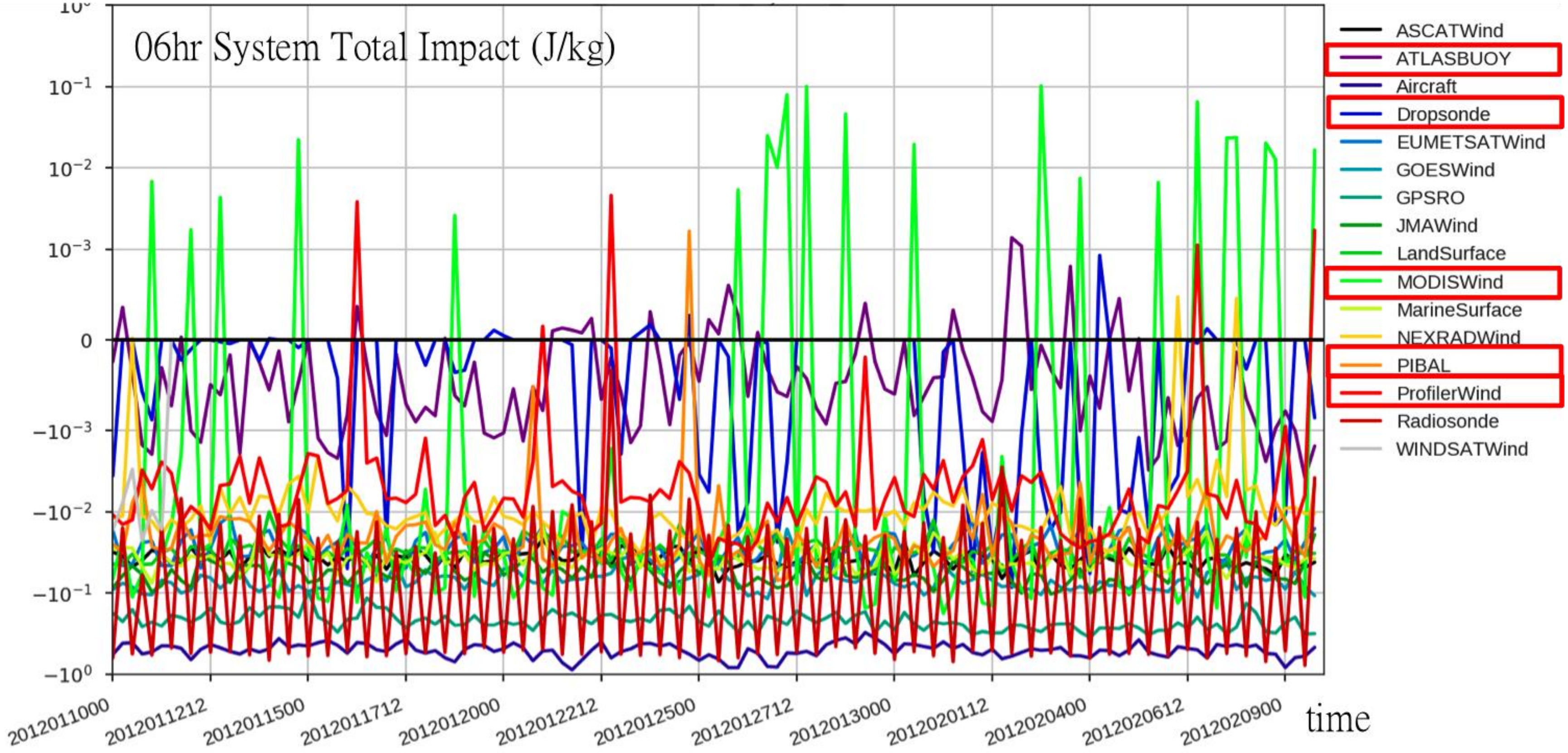
PQC change: Z500

Other advantages of EFSO/PQC

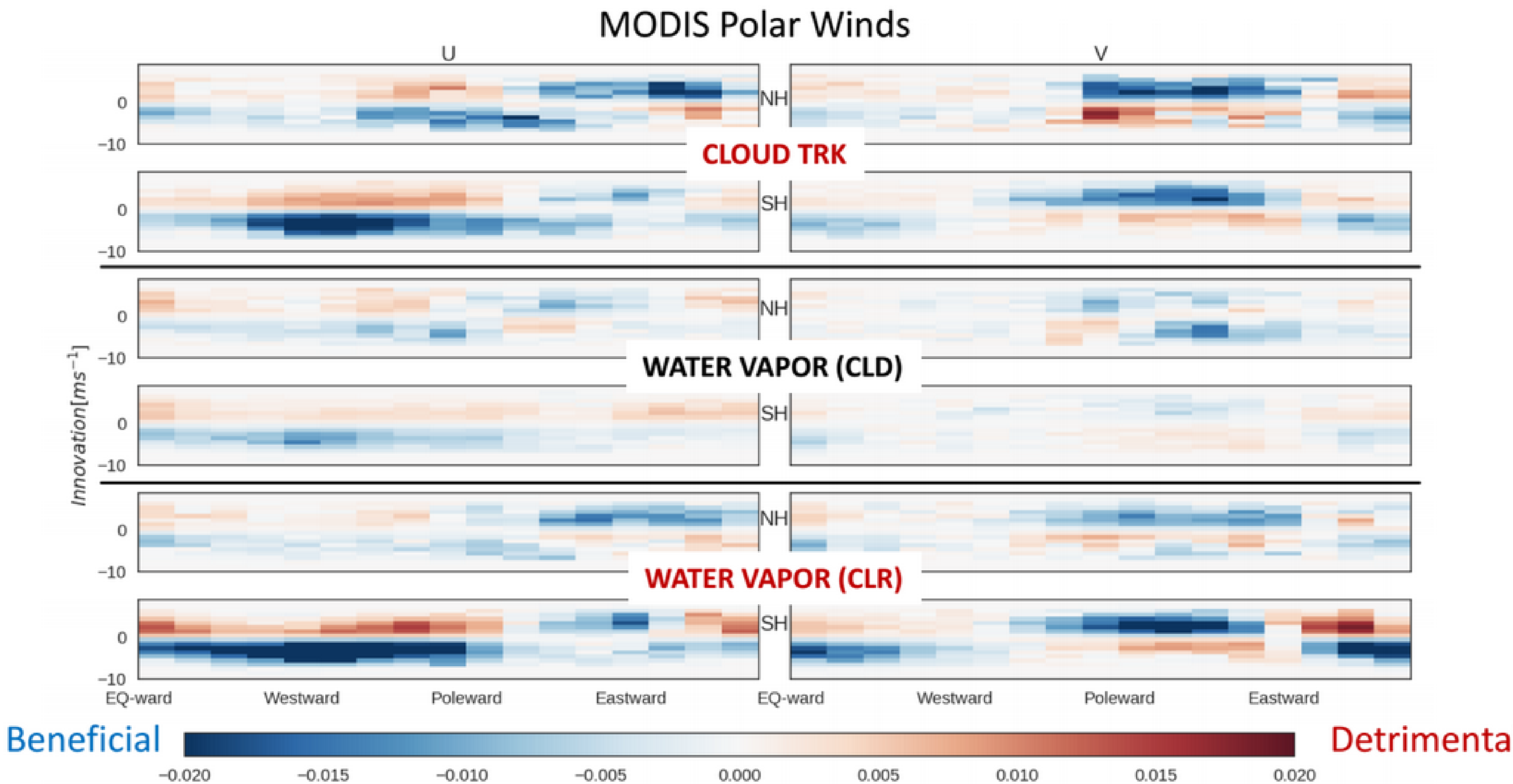
- It can be used to determine whether new instruments are improving the analysis regardless of how many other observations there are.
- EFSO can be used as a clear track of the impact of all observing systems.
- It provides the ability to do a quick QC. For example, Chen found that the detrimental MODIS winds had clear biases.

Alarm bells could be produced in operations!

- Improve NWP by using Ensemble Forecast Sensitivity to Observation (Kalnay et al 2012)
- MODIS winds and Profiler Winds are sometimes detrimental

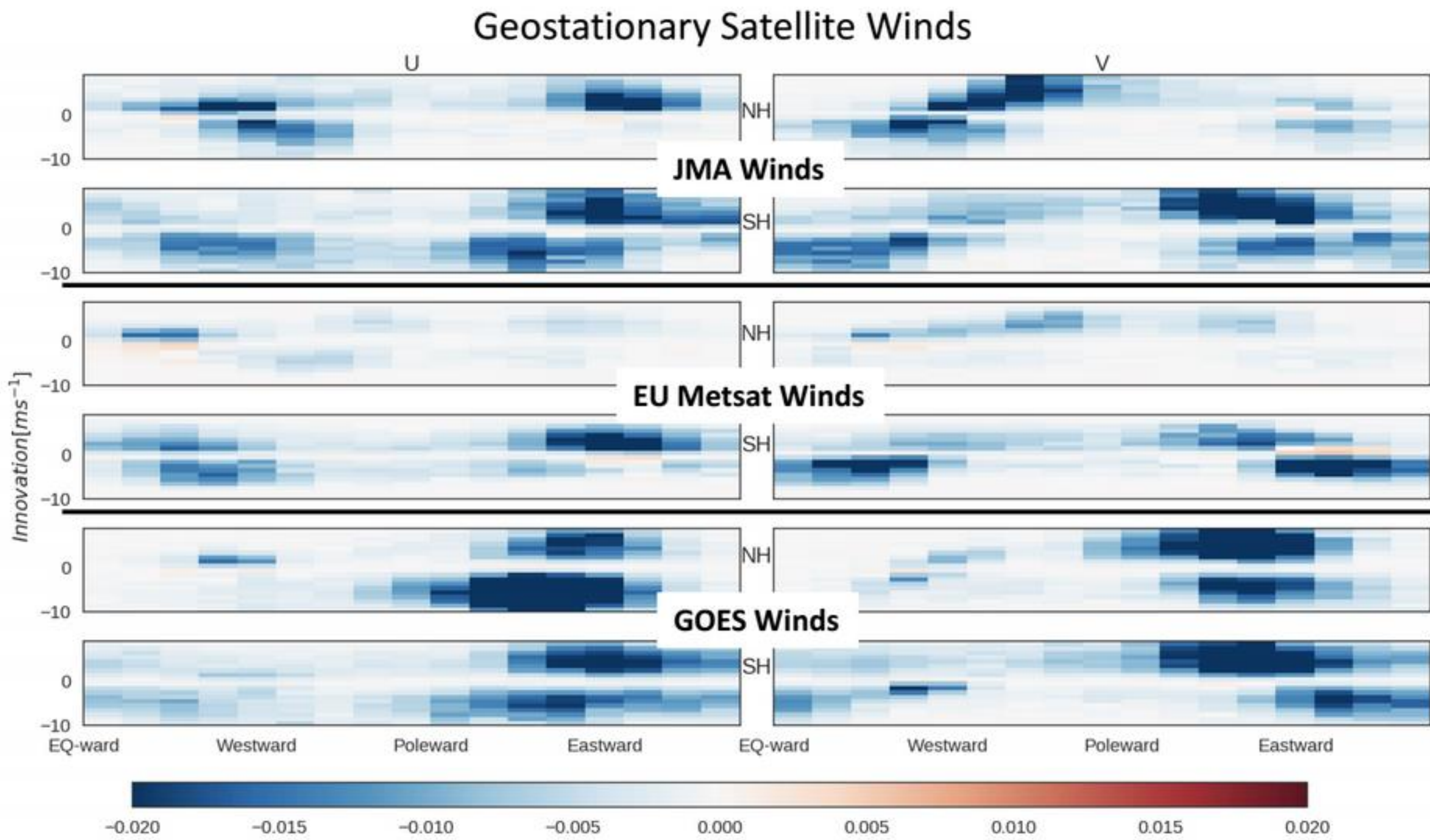


Biases: Innovation and Wind Direction



- Prevailing **positive innovation bias** in U comp.
- Cloud tracking winds (top) and Water vapor tracking (bottom) resemble each other in both hemisphere

Biases: Innovation and Wind Direction



- No such biases for Geostationary Satellite Winds

Summary: Proactive QC based on EFSO

- We found an efficient way to determine for each observation if it beneficial or detrimental, and can avoid large “skill dropouts” due to detrimental observations.
- We are working with the MODIS winds scientists to find and correct the problem that MODIS winds show.
- This method can also be used to implement the assimilation of new instruments much more efficiently than the present approach of computing many 5-day forecasts to try to find whether there is a tiny positive impact.

2) Ensemble Forecast Sensitivity to Error Covariances

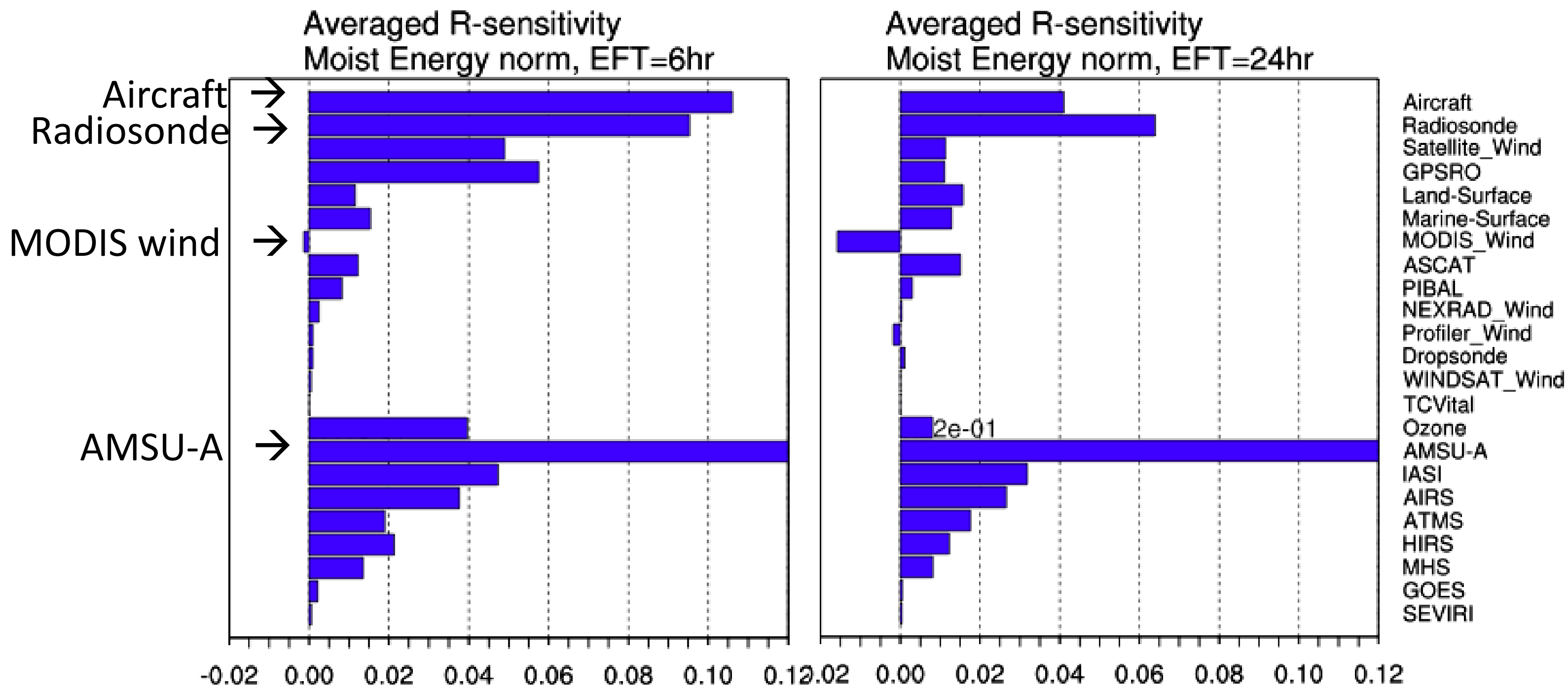
Hotta (2014)

- Daescu and Langland (2013, *QJRMS*) proposed an adjoint-based formulation of forecast sensitivity to **B** and **R** matrix.
- Daisuke Hotta formulated its ensemble equivalent for **R** using **EFSO** by Kalnay et al. (2012) :

$$\left[\frac{\partial e}{\partial \mathbf{R}}\right]_{ij} \approx \frac{\partial e}{\partial y_i} z_j \approx -\frac{1}{K-1}\left[\mathbf{R}^{-1}\mathbf{Y}_{\mathbf{0}}^{\mathbf{a}}\mathbf{X}_{\mathbf{t|0}}^{\mathbf{fT}}\mathbf{C}\left(\mathbf{e}_{\mathbf{t|0}} + \mathbf{e}_{\mathbf{t|-6}}\right)\right]_i \left[\mathbf{R}^{-1}\delta y^{oa}\right]_j$$

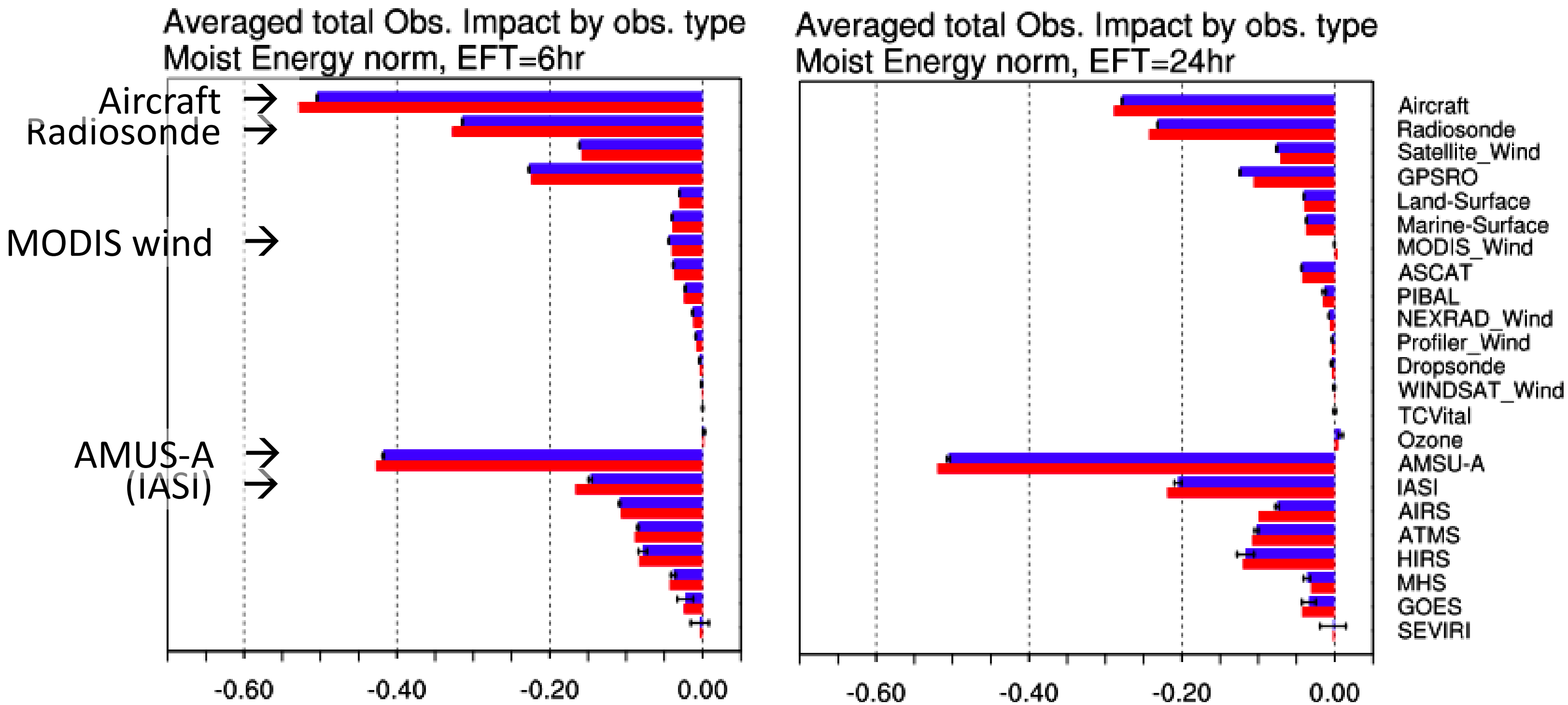
where $\mathbf{z}$ is an "intermediate analysis increment" in observation space

R-sensitivity results from GFS / GSI-LETKF hybrid



- Positive value: error increases as σ_o^2 increases → should decrease σ_o^2
- Aircraft, Radiosonde and AMSU-A: large positive sensitivity
- MODIS wind : negative sensitivity
- → **Tuning experiment**:
  - Aircraft, Radiosonde and AMSU-A: scale σ_o^2 by 0.9
  - MODIS wind: scale σ_o^2 by 1.1

Tuning Experiment: Result
EFSO before/after tuning of R



- Aircraft, Radiosonde and AMSU-A: significant improvement of EFSO-impact
- IASI: Significant improvement in EFSO although its error covariance is untouched!
- Very promising results for quick testing of new observing systems!

3) How can we estimate and correct model bias?

Kriti Bhargava, Eugenia Kalnay, Jim Carton, with Fanglin Yang, Mark Iredell

- The best current estimate of nature is the Analysis.
- The First Guess (6hr forecast) contains the initial forecast errors (before they grow nonlinearly).
- Analysis - First Guess (6hr forecast)= Analysis Increments (**AI**) = 6hr model errors.
- **The time average of AI is the best estimate of the error growth due to model bias in 6 hr.**
- However, the analysis increment may also contain **observation biases**.

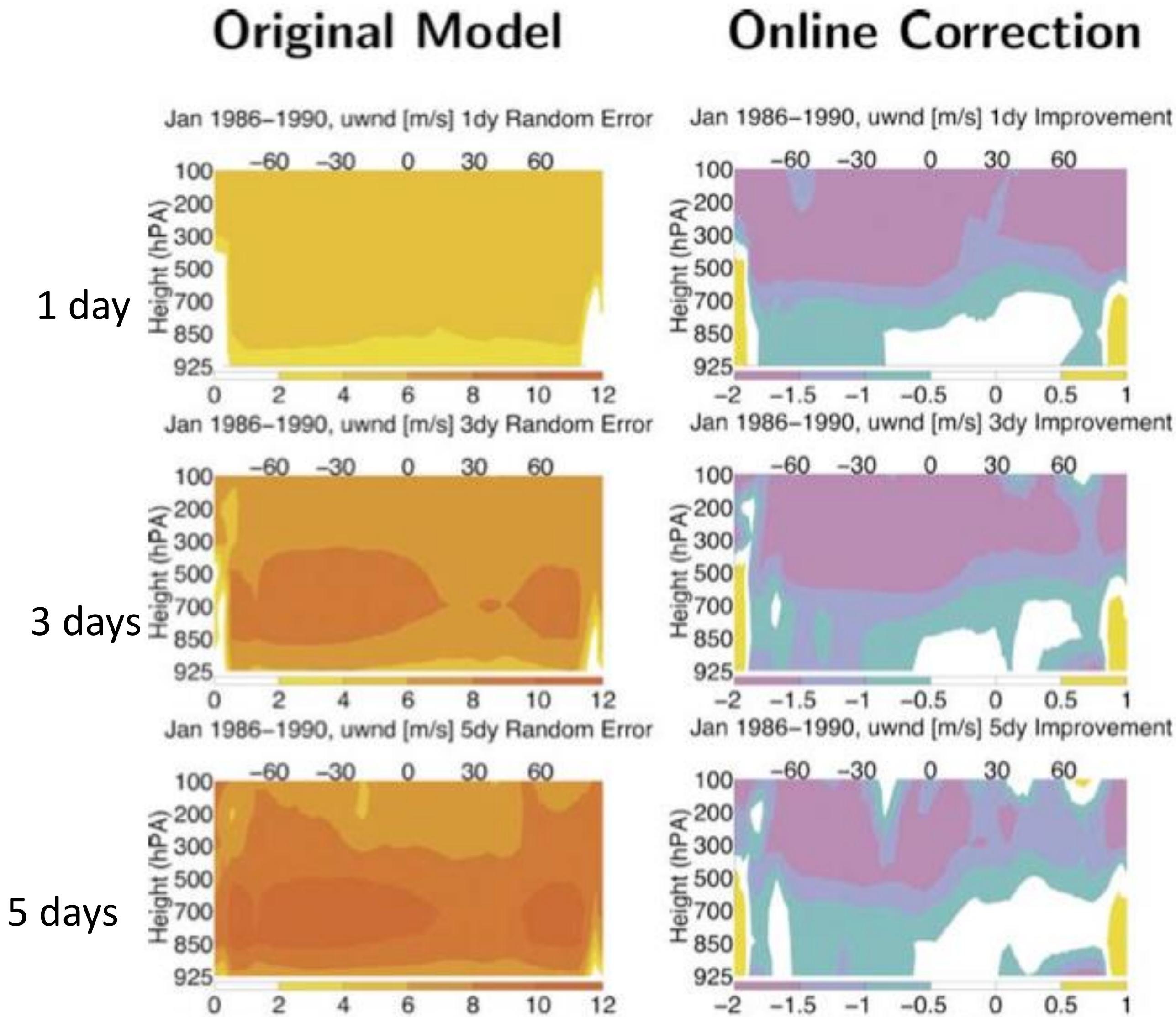
Danforth and Kalnay (2007, 2008a, 2008b)

- Danforth, Kalnay and Miyoshi (DKM-2007) estimated the 6hr errors of the SPEEDY model.
- Estimated the average SPEEDY model error (bias) by averaging:

  Reanalysis R1 – 6 hour forecast » $\overline{\mathrm{AI}}$

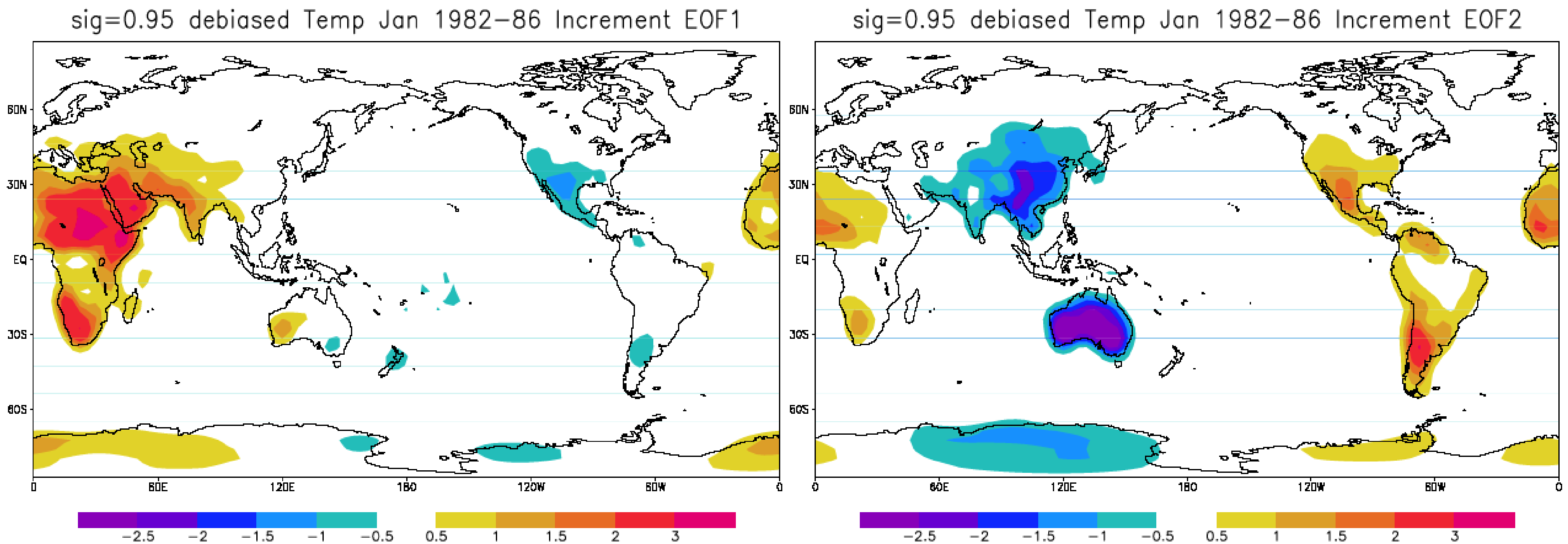
- They corrected the SPEEDY model with $\overline{\mathrm{AI}}/6hr$
- This significantly improved both the forecasts systematic errors and the random errors!

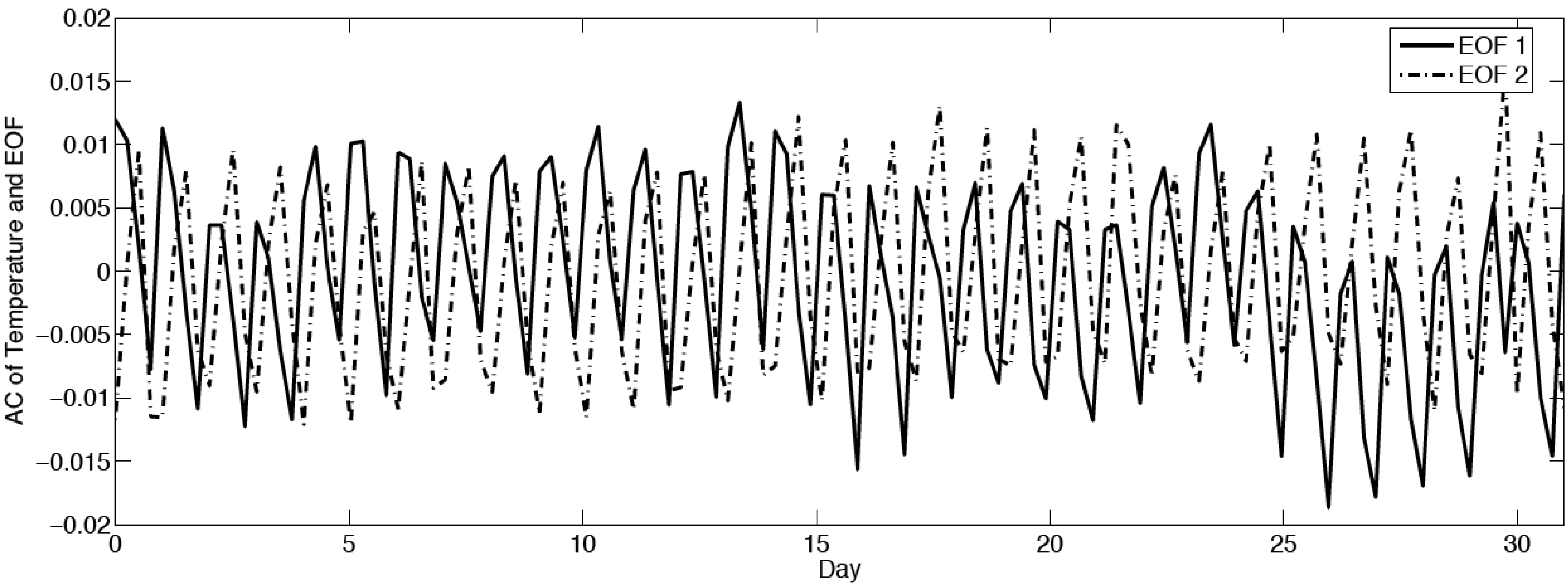
Both bias and random errors were significantly smaller when correcting the model with the model bias!



Random errors were reduced by the online model correction!

The 2 leading EOFs of the error anomalies gave the diurnal cycle errors





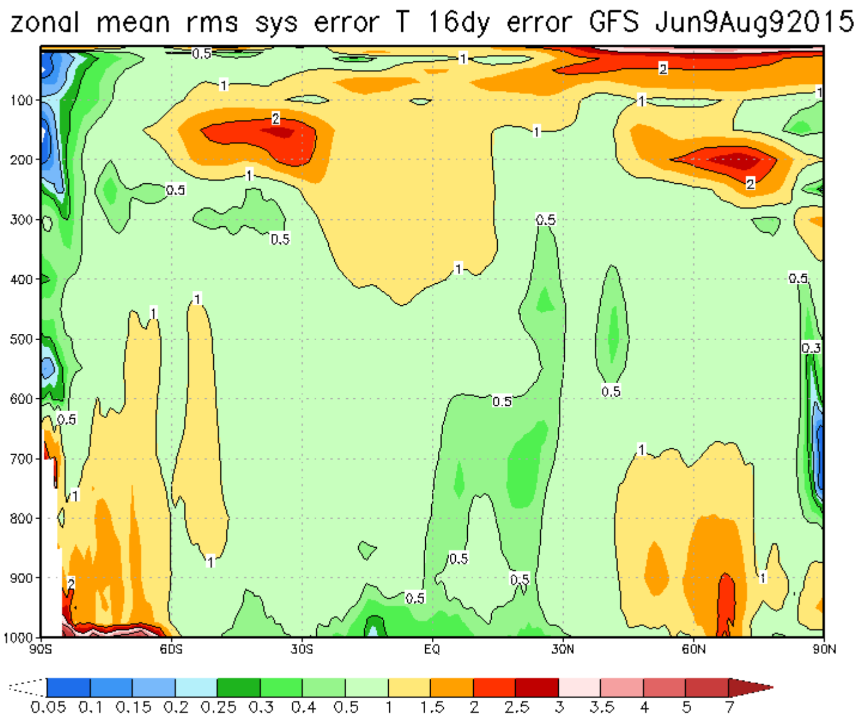
Can we estimate and correct model bias and random forecast errors in the NCEP/GFS?

- The systematic errors in the GFS (and all NWP models) are not negligible.
- They are statistically corrected *a posteriori* (offline).
- We aim to correct the GFS (online) adding the average AI/6hr to each forecast variable, like Danforth and Kalnay (2008).
- This should not only improve the forecasts but also facilitate testing model improvements.
- If the observations are biased, correcting them should reduce the Analysis Increments

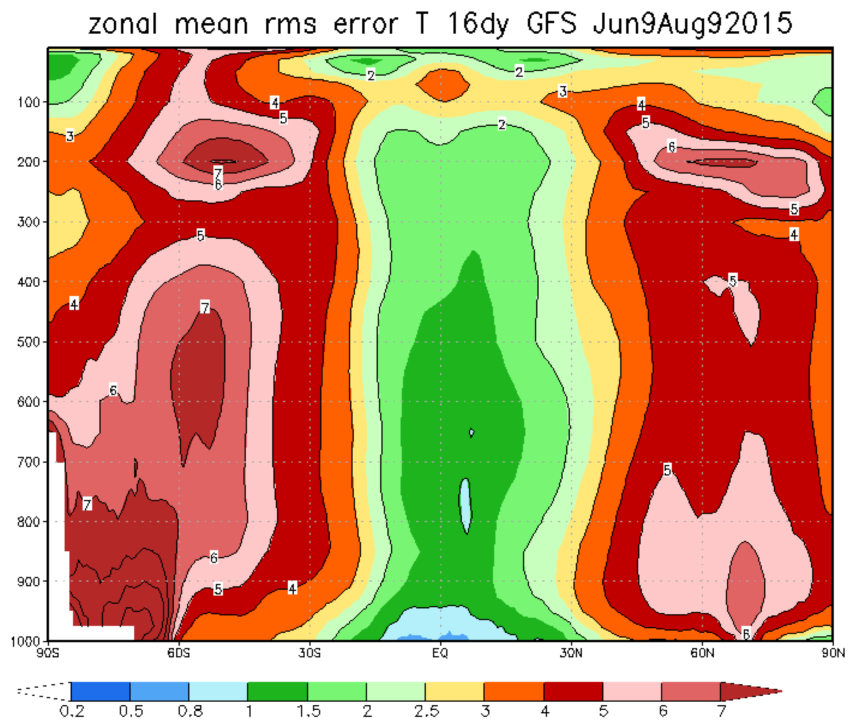
Systematic model errors - GFS

Systematic error range ~1/3 Total error range after 2 weeks

RMS Systematic errors GFS



RMS Total errors GFS



ΔT(systematic) ~ 0.5 -3K

ΔT(total) ~ 1.5 -9K

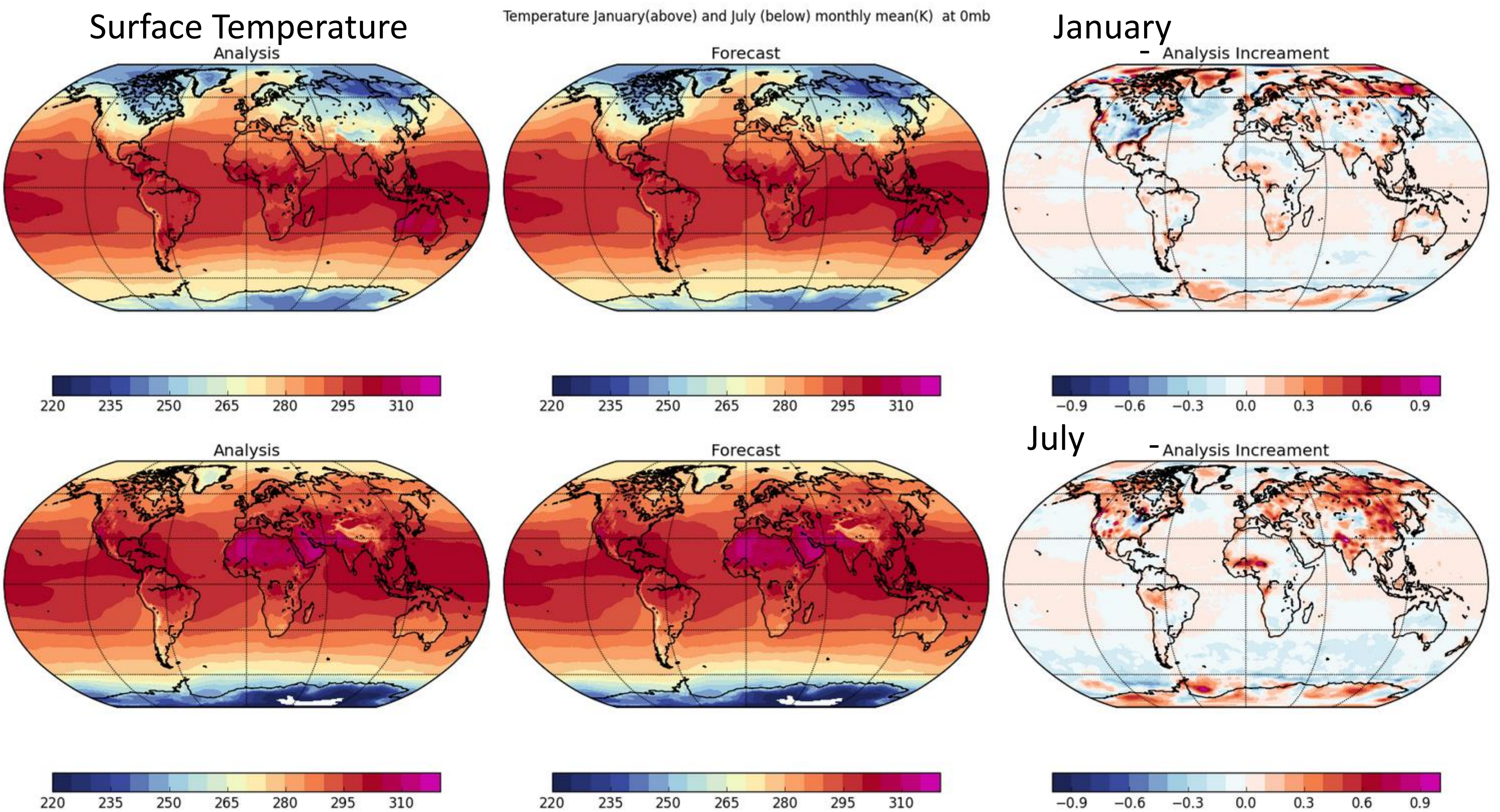
Image courtesy: Glenn White

Application to GFS

Bhargava, Kalnay, Carton

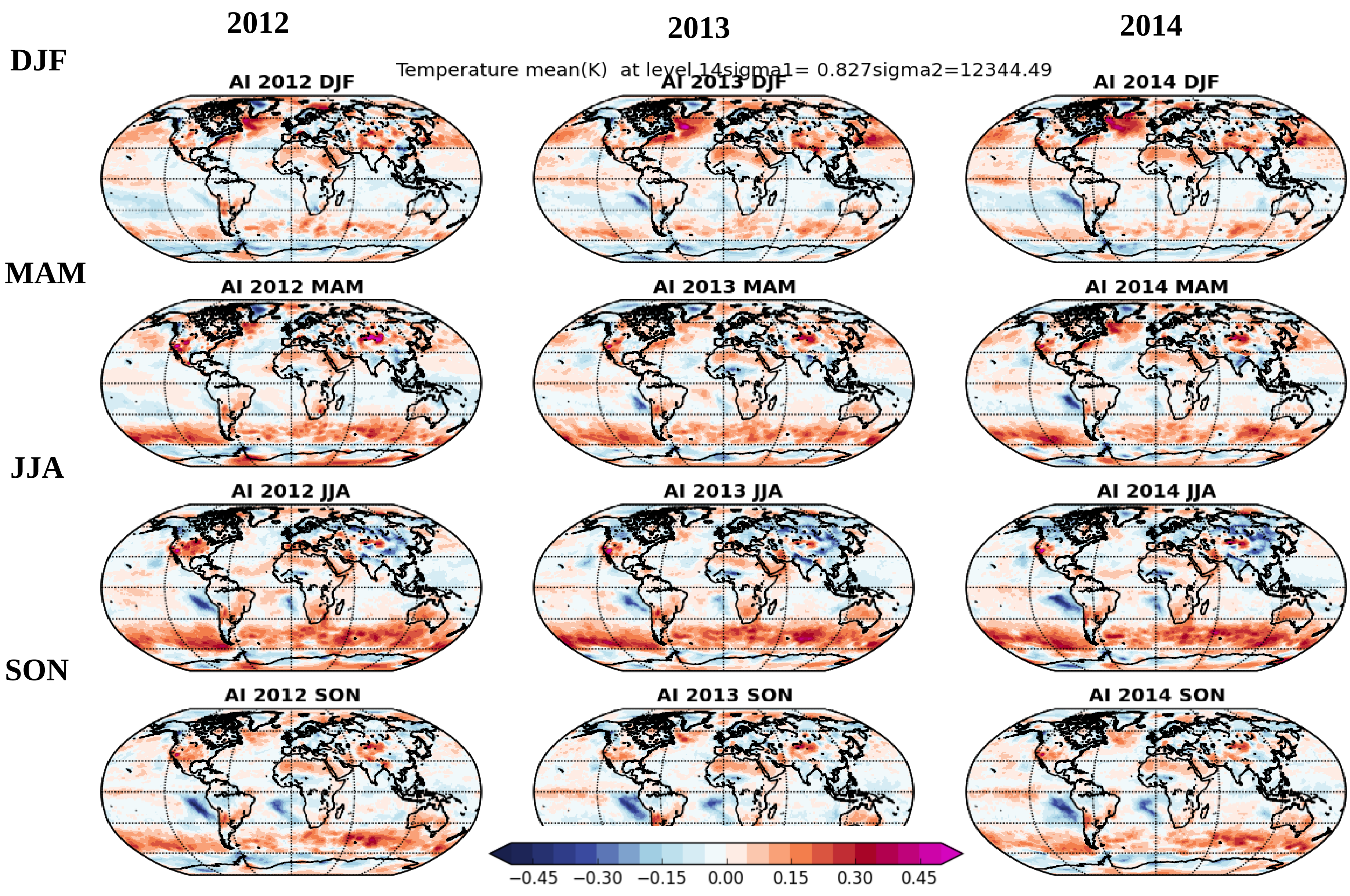
- We obtained T254 6hr forecasts and analyses for 2012, 2013, 2014 from Dr. Fanglin Yang
- We estimate the GFS systematic errors
  - Mean
  - Diurnal
- Check robustness: compare 2012, 2013, 2014
- Explore low dimensional approaches (e.g. diurnal cycle)
- Explore error sensitivity to resolution

First results: 2014 Analyses, Forecasts and Bias



The analysis and 6hr forecasts are almost identical,
but the AI are well defined.

Seasonal Mean Bias: T (K) at ~850 mb for 2012, 2013, 2014



Findings

- Estimate the GFS systematic mean errors ✓
- Check the robustness of the seasonal averaged AI (2012 vs 2013 vs 2014)✓ Errors are robust
- Explore the errors in diurnal cycle
- Check if the low dimensional approaches can be used to correct the diurnal cycle errors
- Validate if errors can be explored at a resolution lower than operational

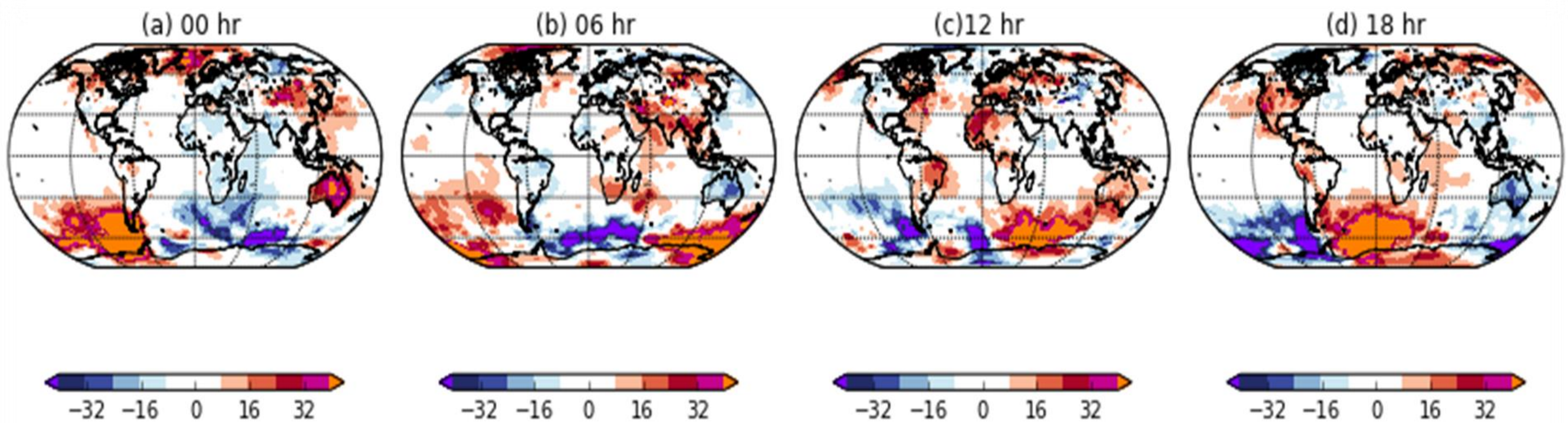
Diurnal cycle error estimation

- Compare the AI at 00, 06, 12 and 18Z
- Compute Empirical Orthogonal Functions (EOFs) of the AI anomaly
- Check how well the diurnal cycle errors are represented by the leading modes

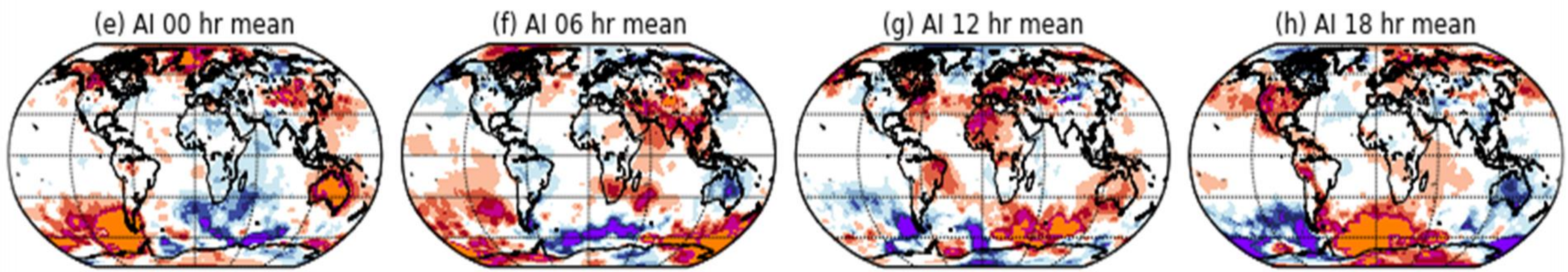
First 4 vs 120 modes: P_s (mb) Sept’14

First 4 EOFs of AI capture the diurnal cycle errors almost perfectly

Top: 4 modes



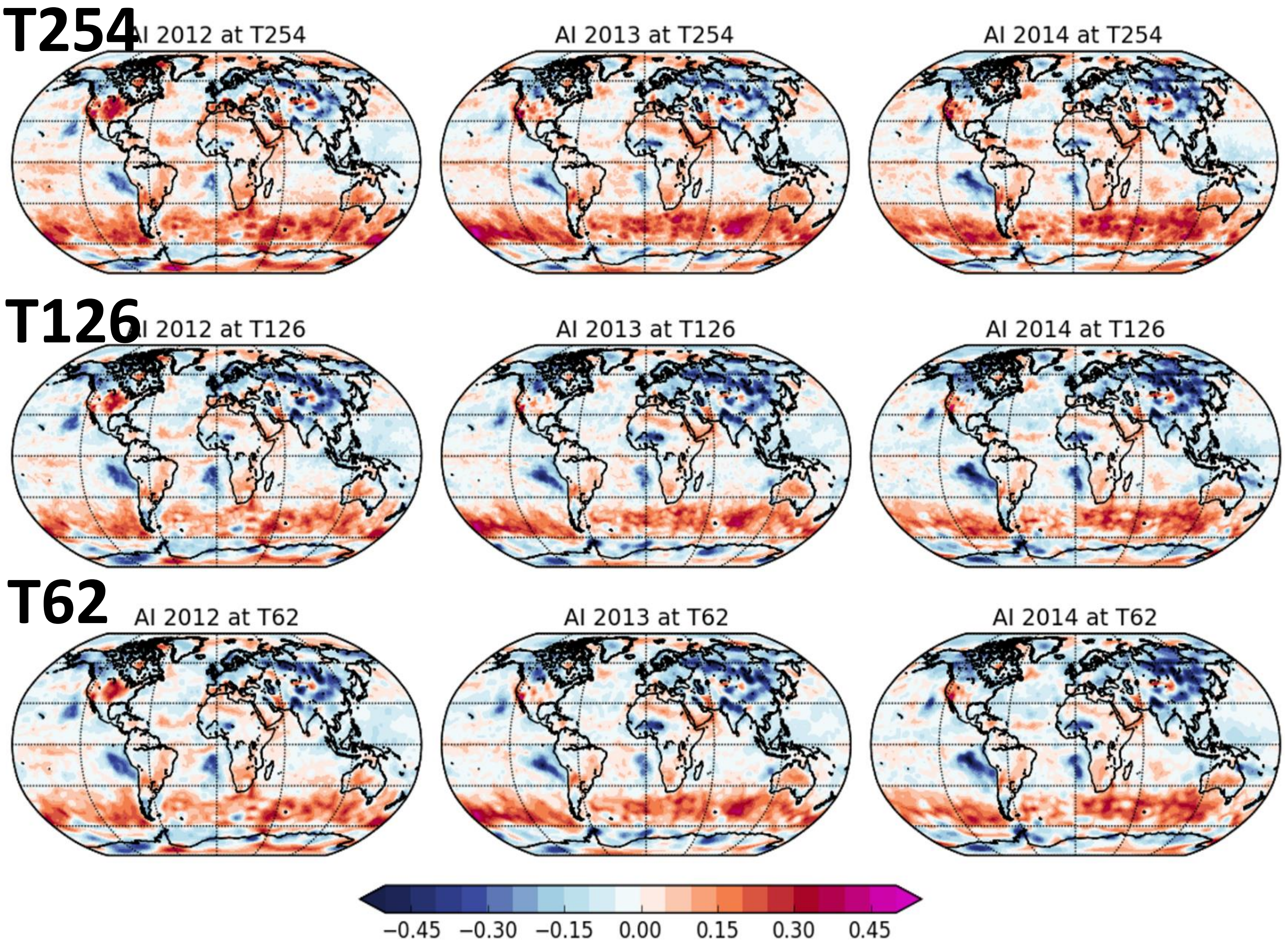
Bottom: 120 modes



Findings

- Estimate the GFS systematic mean errors ✓
- Check the robustness of the seasonal averaged AI (2012 vs 2013 vs 2014)✓ Errors are robust
- Explore the errors in diurnal cycle ✓
- Check if the low dimensional approaches can be used to correct the diurnal cycle errors. ✓ Yes, need only 4/120 modes and should be able to correct the diurnal cycle!
- Check if errors can be explored at a resolution lower than operational

Bias is independent of resolution: it is large scale

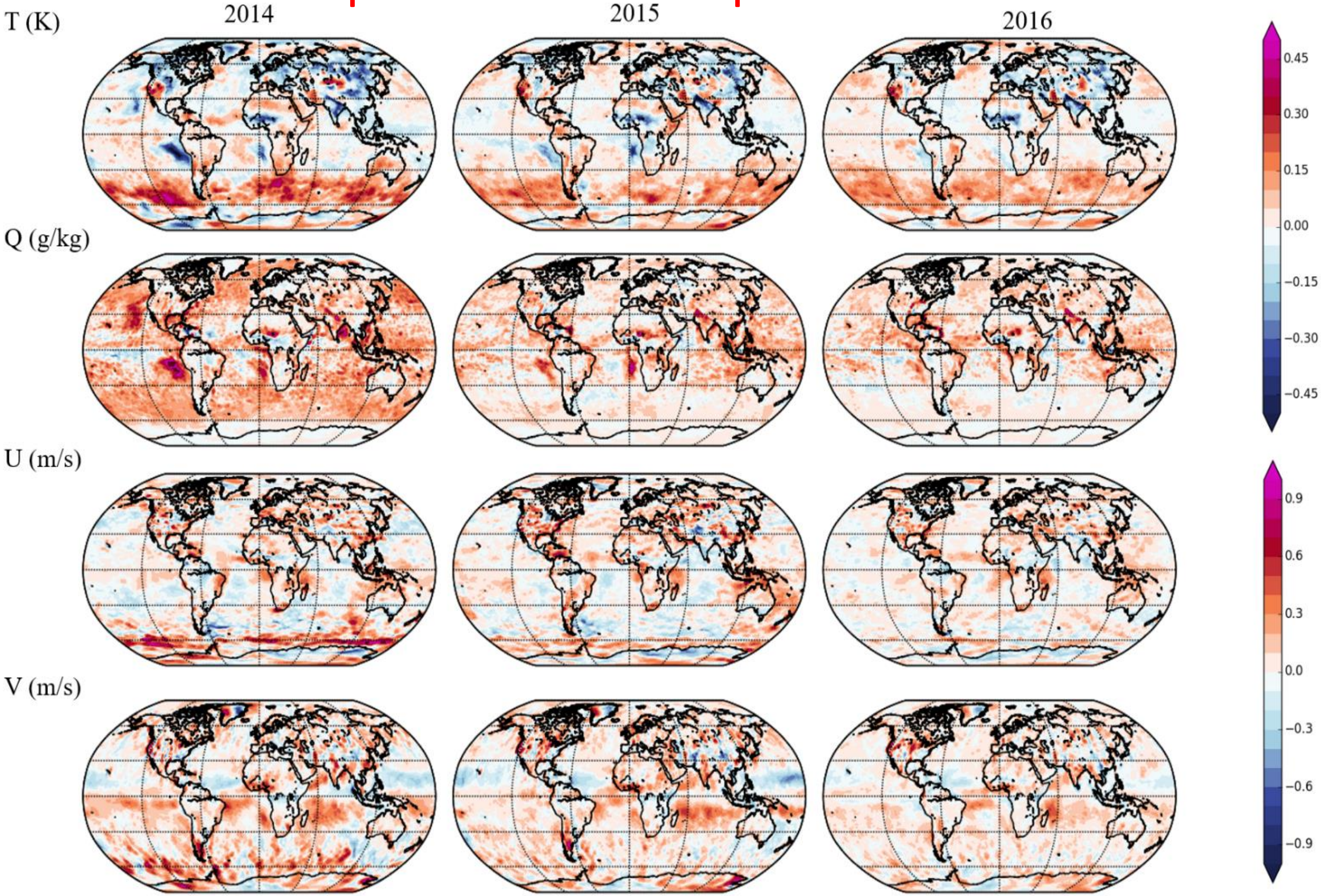


Projecting July 2014 mean Temperature AI at T62 (top), T126 (middle) and original T254 (bottom)

Errors reduced from 2014 to 2015, 2016 over oceans

What produced this improvement?

June at 850mb



- 01/14/2015 12Z: T1534 Semi-Lagrangian GFS Major Upgrade (NWS TIN)
  - **Model Changes**
    * Upgrade from current operational T574 Eulerian (~23km) to T1534 Semi-Lagrangian (~13km)
    * Use high resolution daily RGT SST instead of weekly OI SST, and use daily sea ice analysis
    * Extend high resolution forecast from 8 days to 10 days.
    * Use McICA radiation approximation
    * Reduced drag coefficient at high wind speeds
    * Hybrid EDMF PBL scheme and TKE dissipative heating
    * Retuned ice and water cloud conversion rates, background diffusion of momentum
    * Retuned orographic gravity-wave forcing and mountain block etc
    * Change from Lagrangian to Hermite interpolation in the vertical to reduce stratosph
    * Restructured physics and dynamics restart fields and updated sigio library
    * Consistent diagnosis of snow accumulation in post and model
    * Compute and output frozen precipitation fraction
    * Divergence damping in the stratosphere to reduce noise
    * Added a tracer fixer for maintaining global column ozone mass
    * Stationary convective gravity wave drag
    * New blended snow analysis to reduce reliance on AFWA snow
    * Changes to treatment of lake ice to remove unfrozen lake in winter
    * Modified initialization to reduce a sharp decrease in cloud water in the first model time ste
    * Correct a bug in the condensation calculation after the digital filter is applied
    * Replace Bucket soil moisture climatology by CFS/GLDAS
    * Add the vegetation dependence to the ratio of the thermal and momentum roughness
    * Fixed a momentum roughness issue
    * Accumulation bucket changed from 12 hour to 6 hour between day 8 and day 10
  - **GSI Changes**
    * convert GFS GSI to vertical structure
    * increase horizontal resolution of ensemble from T254 to T574
    * reduce number of second outer loop iterations from 150 to 100.
    * changes in radiance assimilation: upgrade to CRTM v2.1.3
    * move to enhanced radiance bias correction scheme
    * correct bug in AMSU-A cloud liquid water bias correction term
    * assimilate new radiances: F17 an F18 SSMIS, MetOp-B IASI
    * turn off known bad channels: AQUA AIRS channels 321, NOAA-19 AMSUA channel 7, NOAA-19 MHS channel 3
    * increase ATMS observation errors: increase channels 6 - 10 from 0.3 K to 0.4 K, increase channels 11 - 12 from 0.4 K to 0.45 K
    * turn on cloud detection channels for monitored instruments: NOAA-17, -19 HIRS, GOES-13 and -14 sounders
    * changes in assimilation of atmospheric motion vectors (AMV): assimilate NESDIS GOES hourly AMVs, improve AMV quality control
    * improve GPS RO quality control

- 05/11/2016 12Z: Data Assimilation and Model Upgrade (NWS TIN)
  - **Data Assimilation Upgrade**
    * Upgrade the 3D Hybrid Ensemble-Variational to 4D Hybrid Ensemble-Variational Data Assimilation System
    * Multivariate Ozone update
    * Assimilate all-sky (clear and cloudy) radiances
    * Bias correct aircraft data
    * Modify relocation and storm tracking to allow hourly tropical cyclone relocation
    * other upgrades (e.g. CRTM, Data selection/thinning, AMV winds, etc.)
  - **Model Upgrade**
    * Corrections to land surface to reduce summertime warm, dry bias over Great Plains
    * Hourly output fields through 120-hr forecasts
    * Improved icing probability products and new icing severity product
    * add five more levels from 10 hPa to 1 hPa in post-processed pgb files

14/1/2015: Use high resolution daily RGT SST instead of weekly OI SST, and use daily sea ice analysis

We found the change that improved T and Q over oceans. The AI approach could be used to test and attribute these changes.

Source:http://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html

Findings

- Estimate the GFS systematic mean errors ✓
- Check the robustness of the seasonal averaged AI: (2012 vs 2013 vs 2014)✓ Errors are robust
- Find errors in diurnal cycle ✓
- Check if the low dimensional approaches can be used to correct the diurnal cycle errors. ✓ Yes, need only 4/120 modes and should be able to correct the diurnal cycle!
- Check if errors can be explored at a resolution lower than operational. ✓ Yes, the errors project on low wave numbers <<T62 (large scales)
- In 2015-2016 the errors over ocean were smaller: We traced this to the replacement of weekly OI SST with daily high resolution Real Time Global RTG SST. ✓

Proposed plans for GFS correction in collaboration with EMC

- Apply online AI/6hr corrections to GFS
- Examine if it improves bias and random error
- Compare online correction results with standard operational statistical bias correction
- Facilitate testing new parameterizations of the physics: They should reduce the AI
- Compare the 2014 online correction with the impact of the use of improved SST in 2015
- Examine the systematic errors in the CFS
- **This should facilitate GFS improvements at NCEP**

4)Strongly Coupled Data Assimilation

Travis Sluka

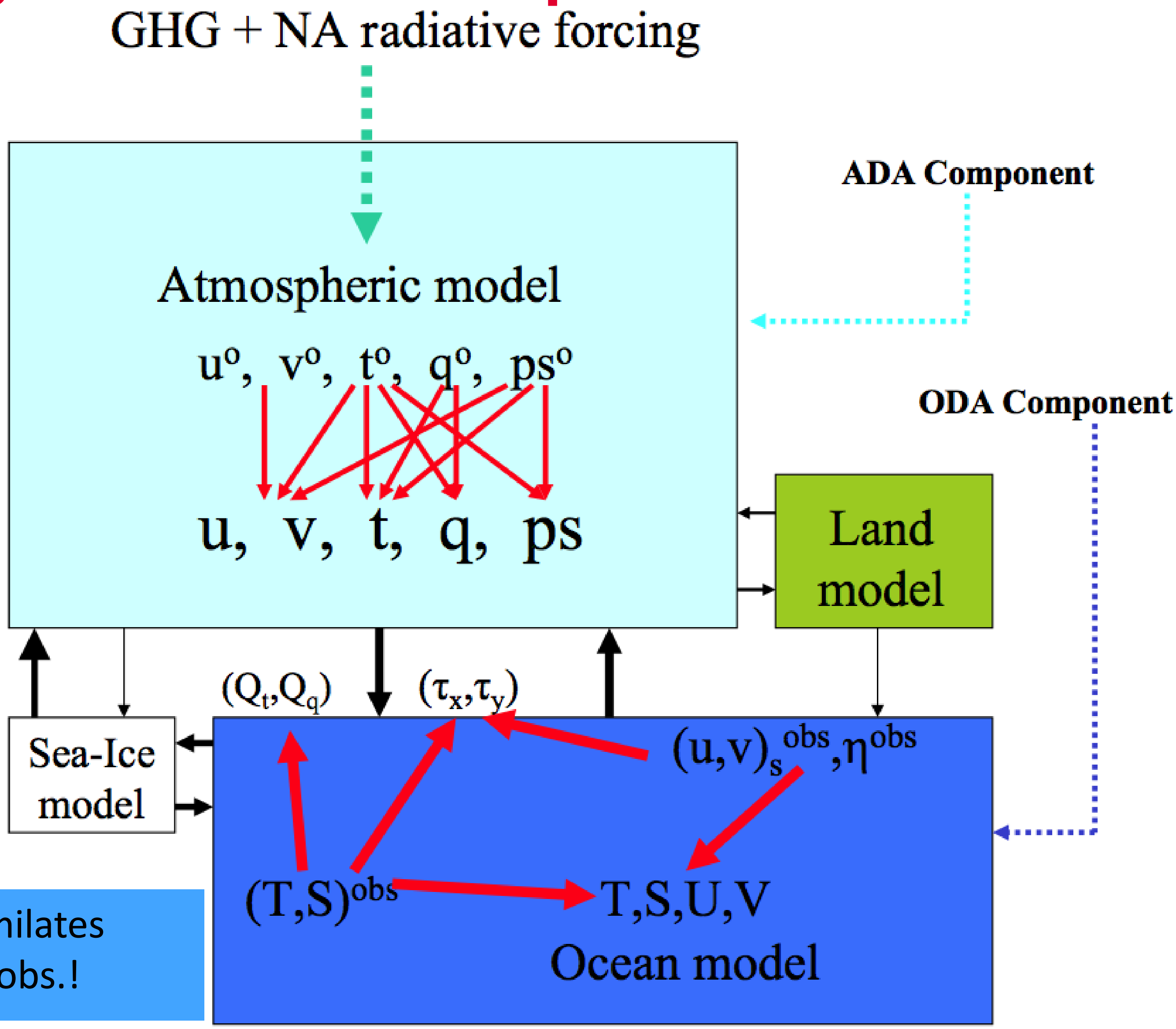
with Steve Penny, Eugenia Kalnay and Takemasa Miyoshi

University of Maryland

4) How should we do coupled ocean-atmosphere data assimilation?

- Should we do coupled data assimilation?
- Yes: e.g., see Tamara Singleton thesis (in a toy coupled ocean-atmosphere model, strongly coupled DA was best)
- Current approaches assimilate separately the ocean and the atmosphere observations, and then couple the models (weak coupling)
- We proposed strong coupling: the ocean "sees" the atmospheric observations, and the atmosphere "sees" the ocean observations (Sluka, Penny, Miyoshi

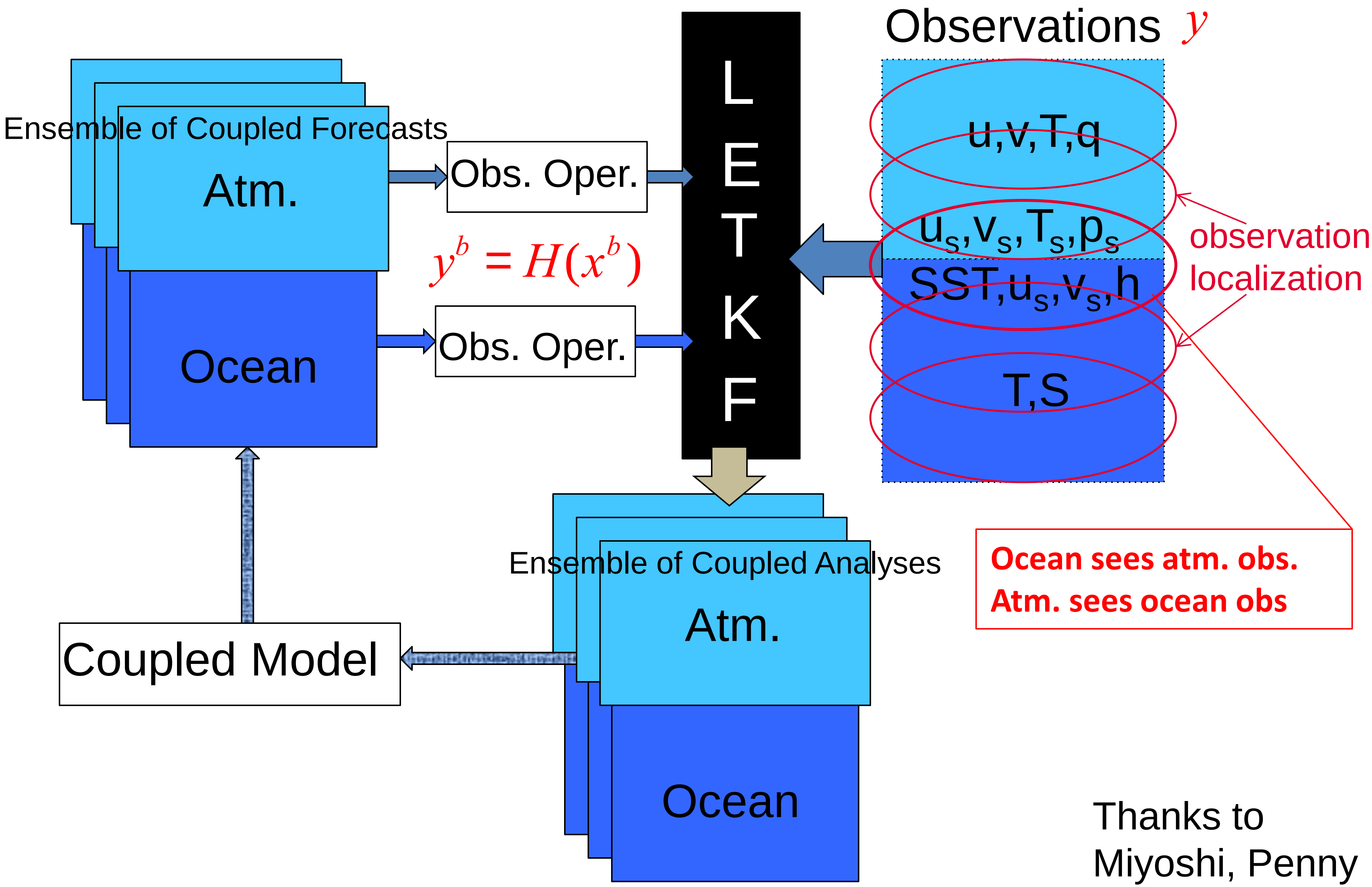
Data Assimilation: STANDARD (WEAK) COUPLING
S. Zhang et al.: GFDL Coupled Ocean-Atm EnKF



Atmosphere assimilates only atm. obs.!

Ocean assimilates only ocean obs.!

Strongly coupled LETKF assimilation



Thanks to
Miyoshi, Penny

Impact of strong coupling of the ocean-atmosphere LETKF (**Sluka et al.**, GRL, 2016)

- **SPEEDY-NEMO** coupled model. Perfect model OSSE.
- Standard (weak) coupling as a control
- Test strong coupling: the ocean sees the atmospheric observations and the atmosphere sees the ocean observations

Experiments: 1) Only atmos. obs.

(2) Only ocean obs.)

- CONTROL: Weakly coupled data assimilation: Only the atmosphere assimilates atmos. observations.
- Strongly coupled DA: ocean also assimilates atmospheric observations (and vice versa).

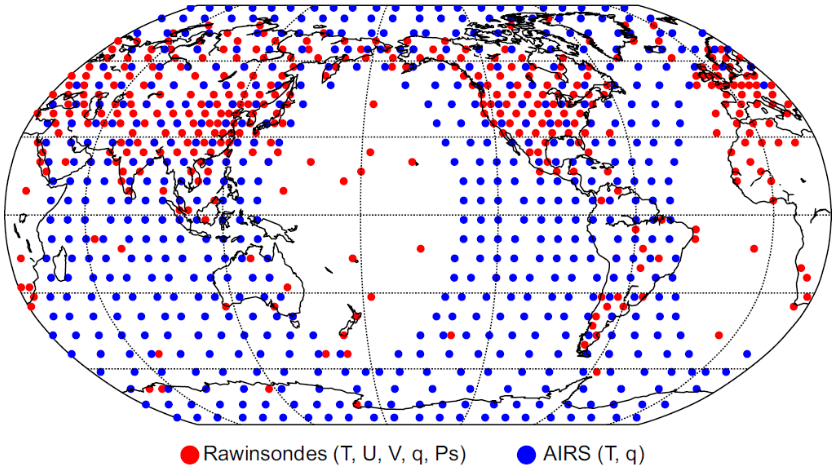
Sluka et al., GRL, 2016

SPEEDY-NEMO OSSE

Using the fast SPEEDY-NEMO (one year run takes only 12 hours on 1 core)

- Perfect model OSSE conducted first using **only atmospheric observations**

6 hr ATM observations



SPEEDY-NEMO

- T30 atmosphere
- 2 degree ocean
- Coupling every 6 hours

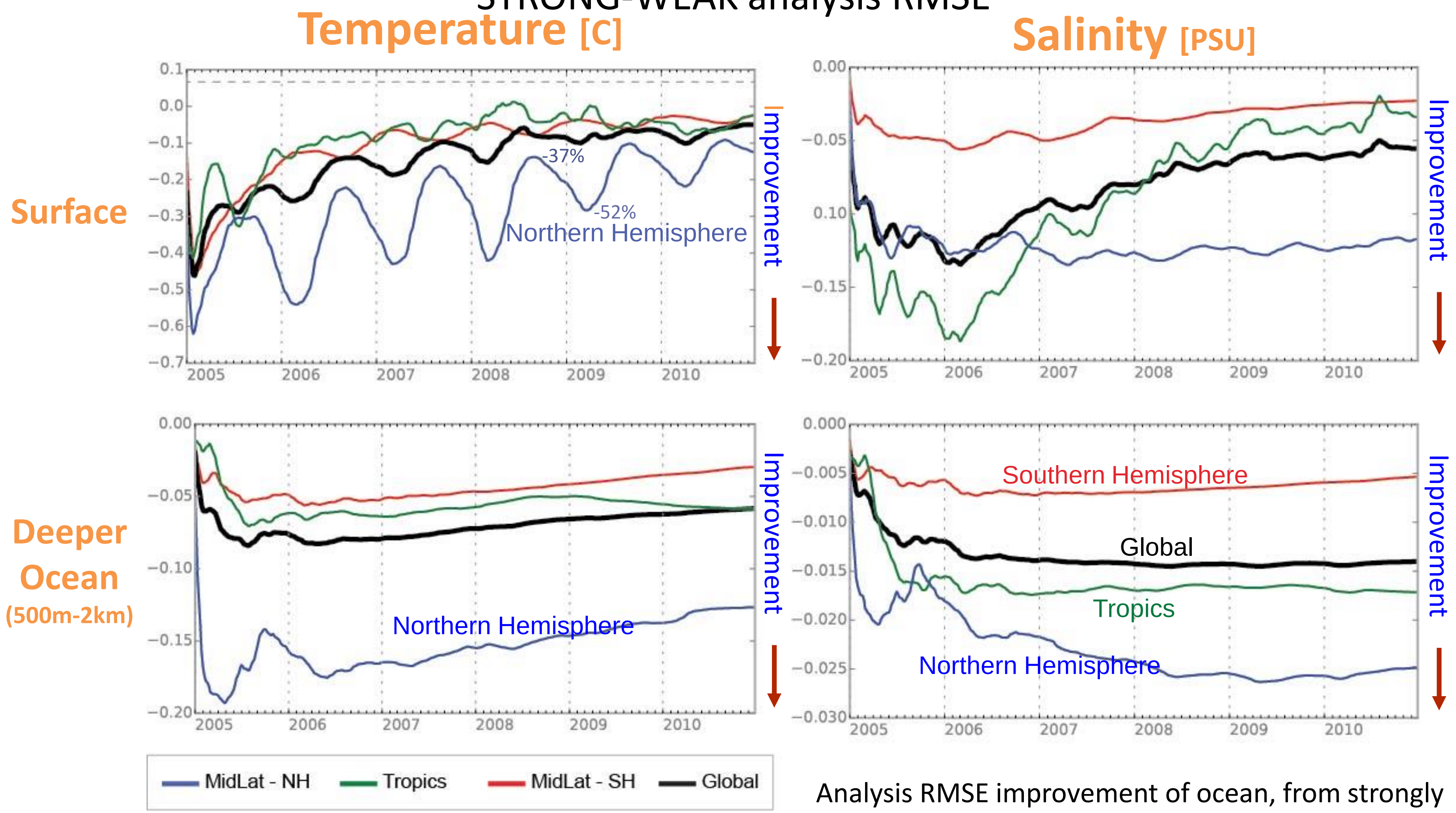
Experiment parameters

- 40 ensemble members
- Localization: 1000km Horiz.
- Relaxation to prior spread: 90% for OCN, 60% for ATM

Sluka et al., GRL, 2016

SPEEDY-NEMO Strongly Coupled DA

STRONG-WEAK analysis RMSE



Analysis RMSE improvement of ocean, from strongly coupled DA of simulated **atmospheric observations**

Sluka et al., GRL, 2016

SPEEDY-NEMO Strongly Coupled DA

STRONG-WEAK analysis RMSE

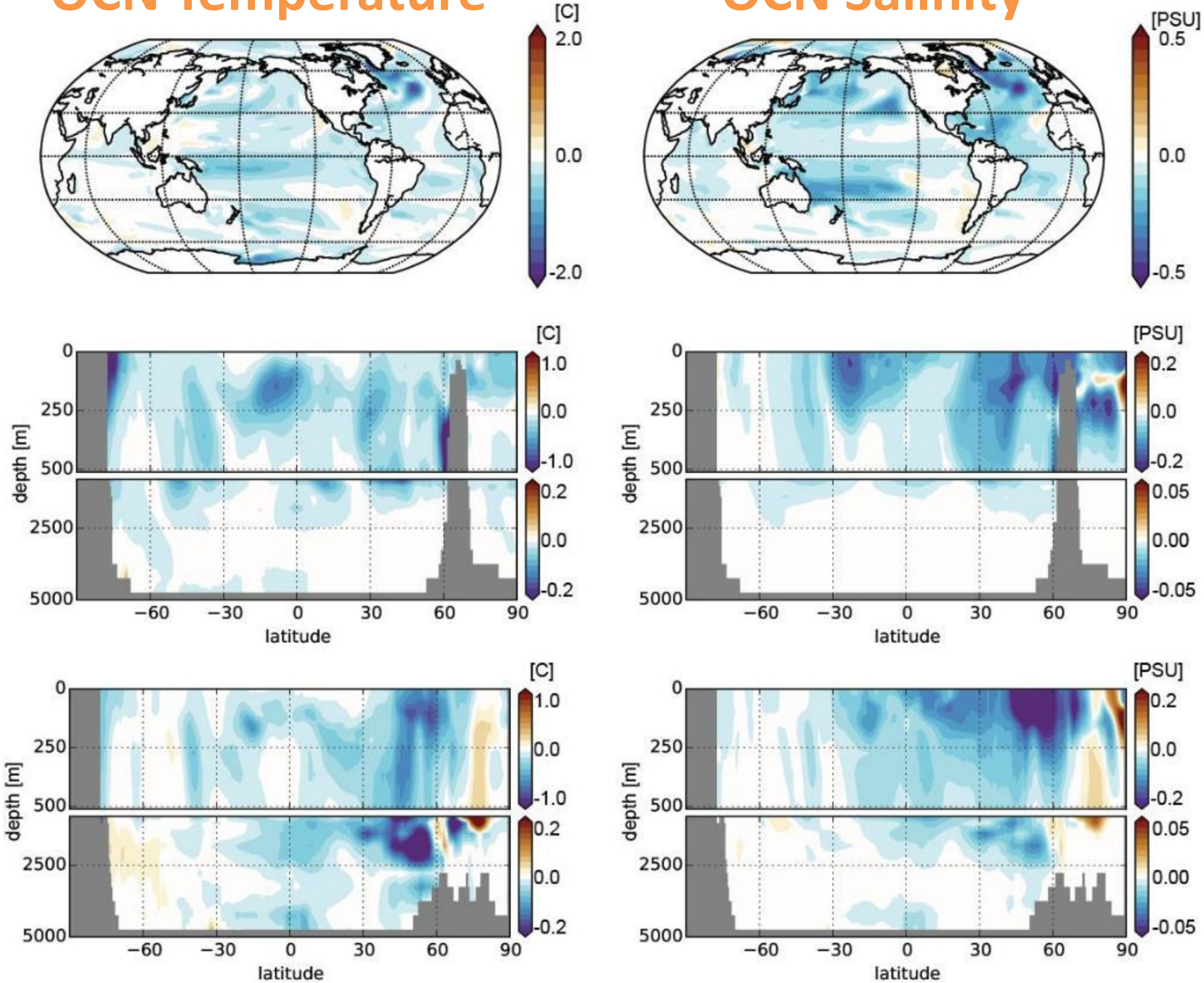
OCN Temperature

OCN Salinity

Upper 500m

Pacific

Atlantic

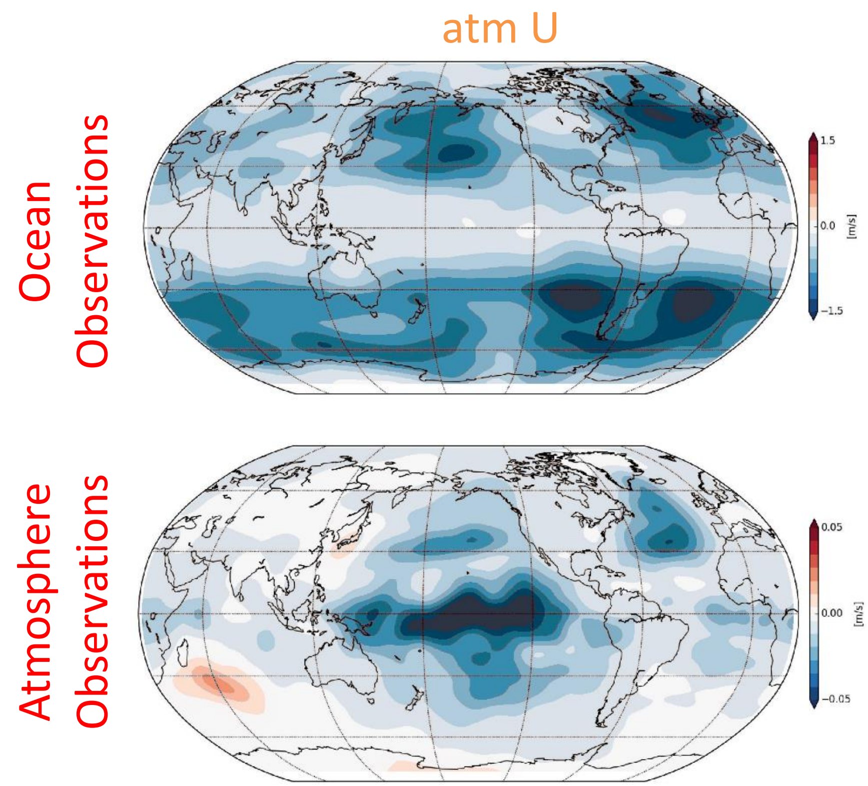


SPEEDY-NEMO Strongly Coupled DA

STRONG-WEAK analysis RMSE

STRONG-WEAK, **blue is good**

- The opposite experiment (assimilating OCN obs into the atmosphere) shows improvement as well
- Interesting! A **coupled ocean** drives the atmosphere in the tropics, and so, **ocean obs** dominate in the extratropics!
- Ocean observations affect the ATM where OCN coupling cannot have an impact.
- And ATM OBS impact where ATM coupling cannot have an impact



Now Sluka is testing strongly coupling the NCEP CFS (Coupled Forecasting System) with real observations

- Weak coupling experiment: JJA 2005. Atmosphere assimilates all atmospheric observations except radiances every 6hrs. Ocean assimilates profiles (buoys) every 24hrs, at 12Z, no SST relaxation.
- Strong coupling: Like the weak coupling, but the ocean also assimilates surface ship atmospheric T and Q every 24 hrs.
- Uses LETKF with 50 member ensemble

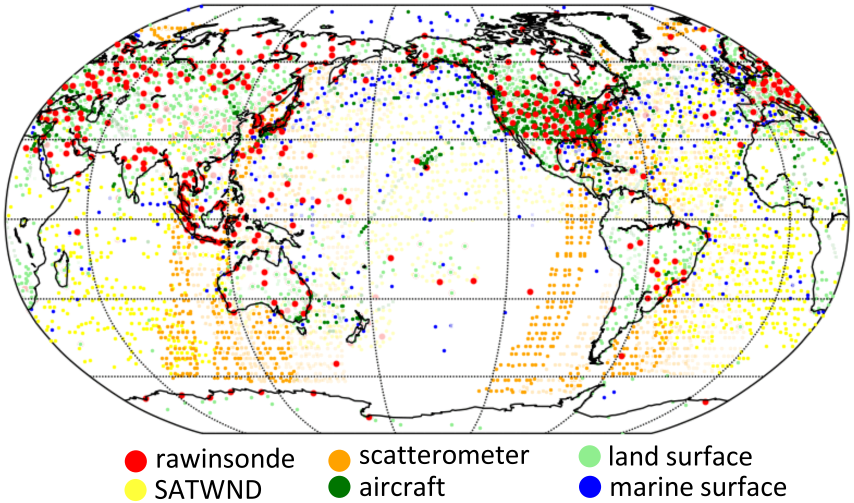
CFSv2-LETKF

- Combined **existing** GFS-LETKF (Lien, 2013) and MOM-LETKF (Penny, 2013)

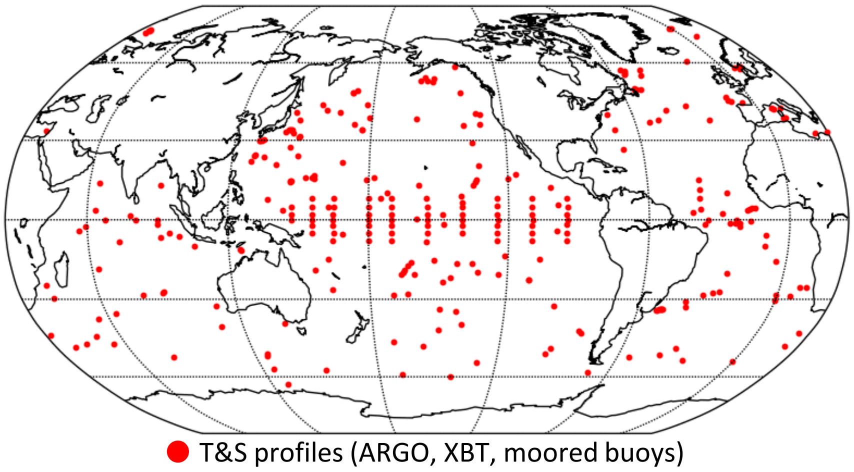
- T62/L64 atm 0.5deg ocn (reduced resolution ATM)
- 50 member ensemble (initialized from CFSR, run freely for 6 months to develop sufficient spread)

- observations from operational ATM PREPBUFR and OCN profiles used by GODAS

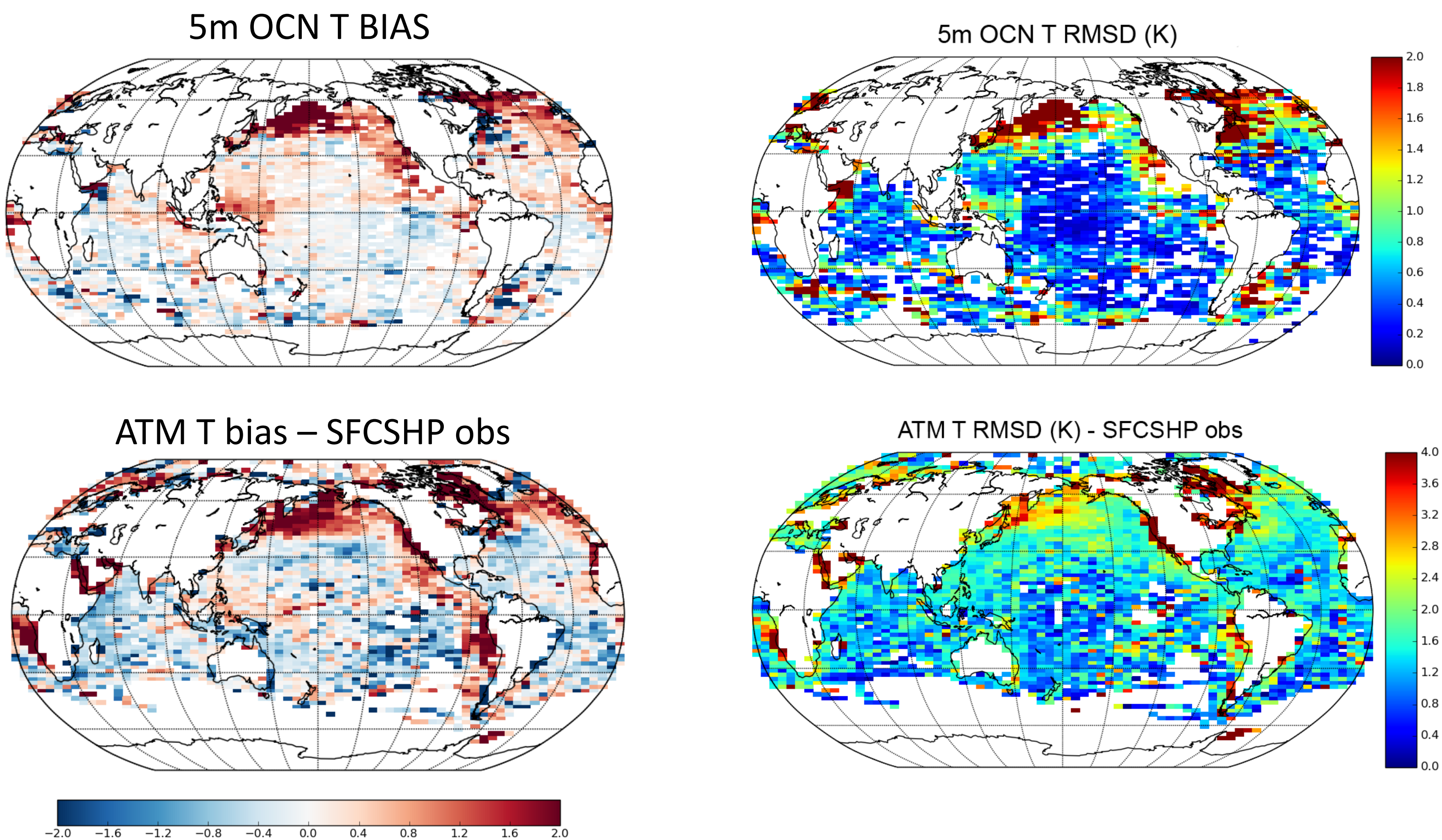
ATM



OCN



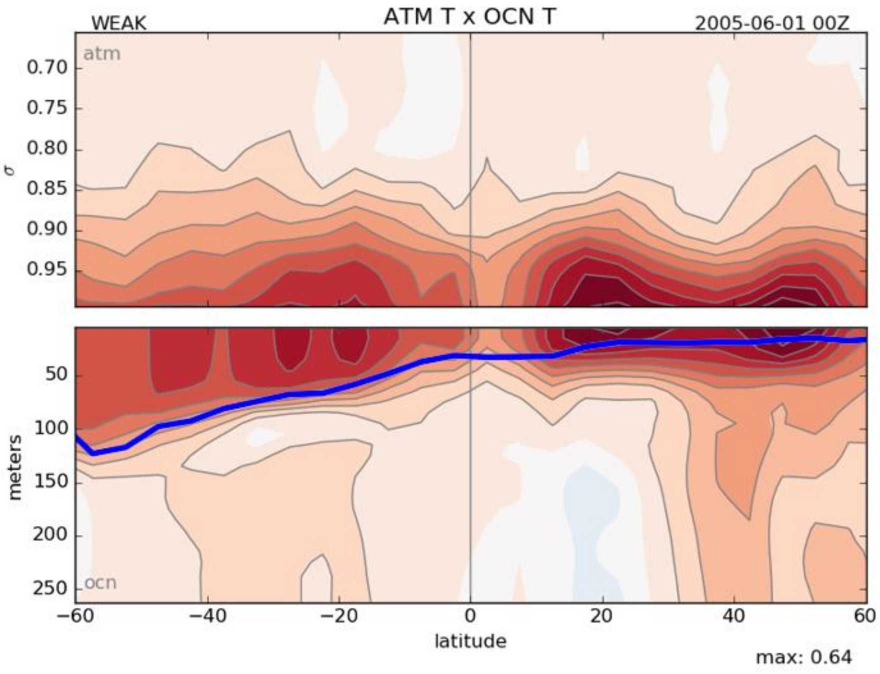
Weakly Coupled DA - JJA



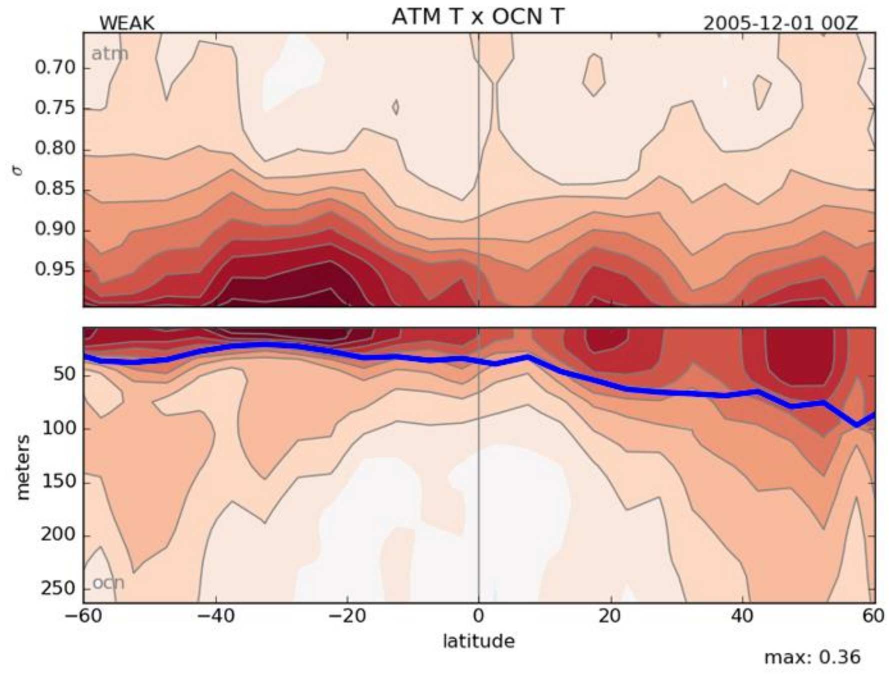
Weakly Coupled DA – cross covariances

- Cross correlations given by the ensemble for a single date
- ATM and OCN temperature max correlation of 0.36, highest values in that hemisphere's summer, below 850mb and above top of thermocline
- June values likely artificially large due to insufficient spin up time for the ocean

June



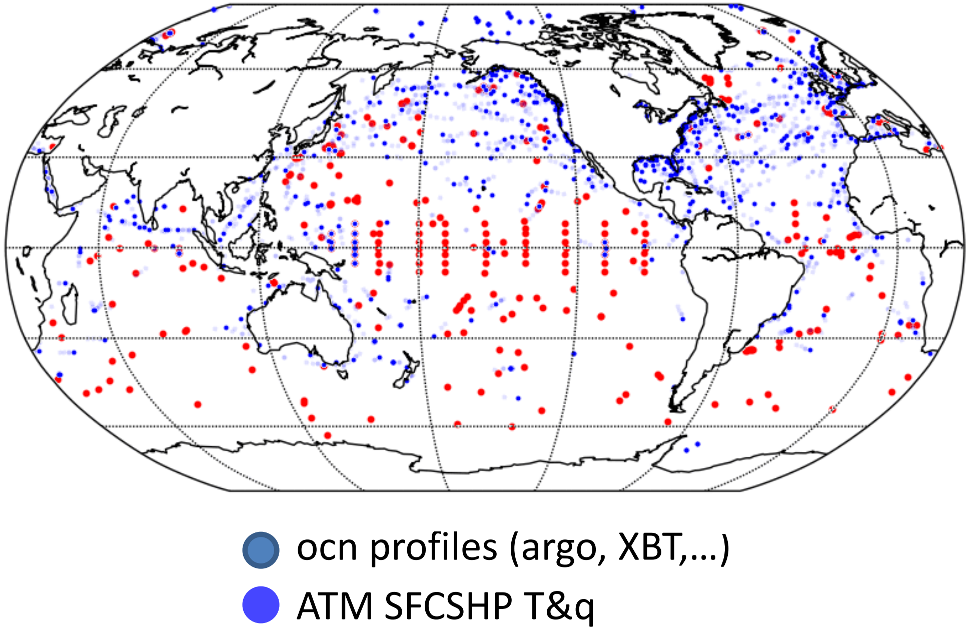
December



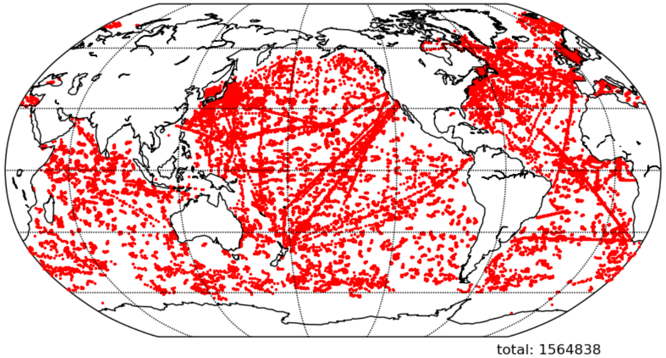
Mixed Layer depth (depth of $T_{10m} \pm 0.2°C$)

Strongly coupled DA

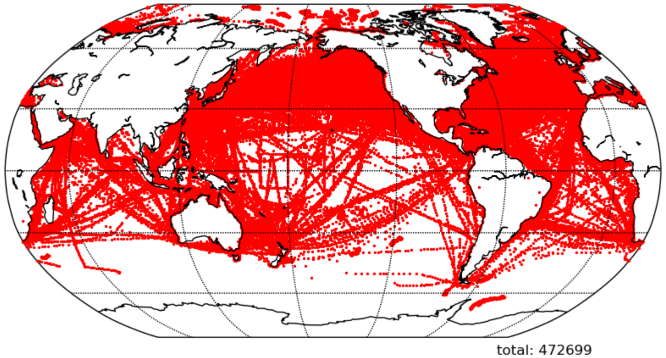
- 1 way strongly coupled DA
- Strongest cross correlations are between OCN_T and ATM_T/ATM_q, so…
- OCN assimilates surface ship T and q as well, given by the **SFCSHP** section of the PREPBUFR



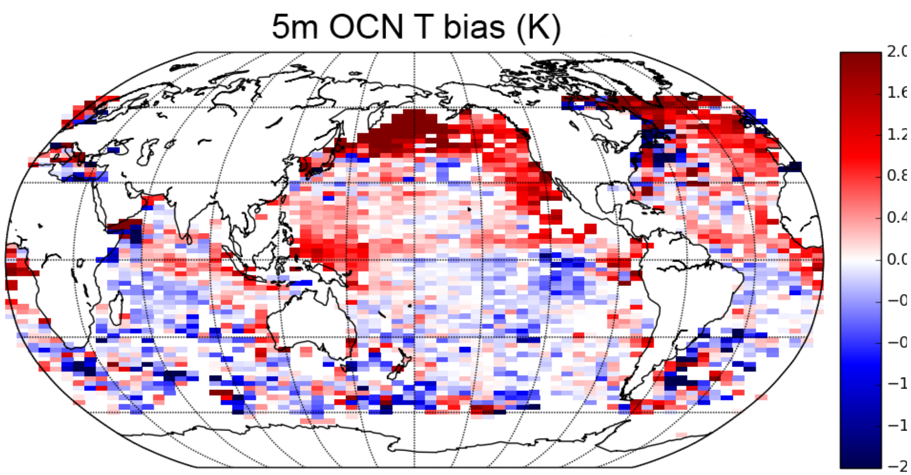
OCN obs (JJA)



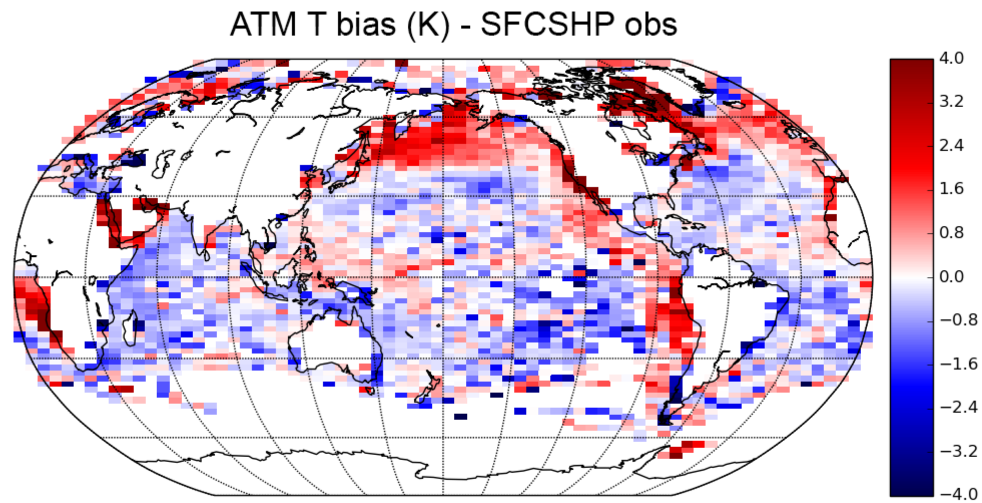
ATM SFCHSP obs (JJA)



Sluka: First results testing weakly coupling the NCEP CFS with real observations



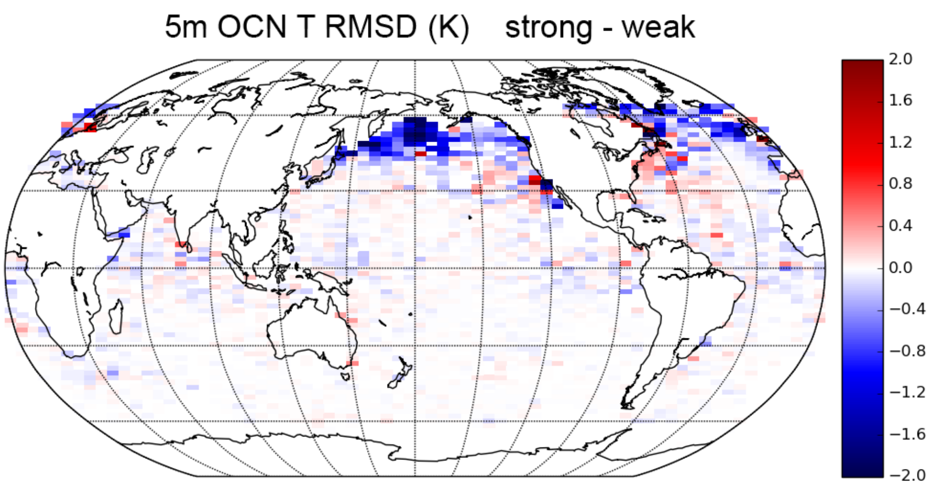
Weakly Coupled DA Ocean 5m T bias

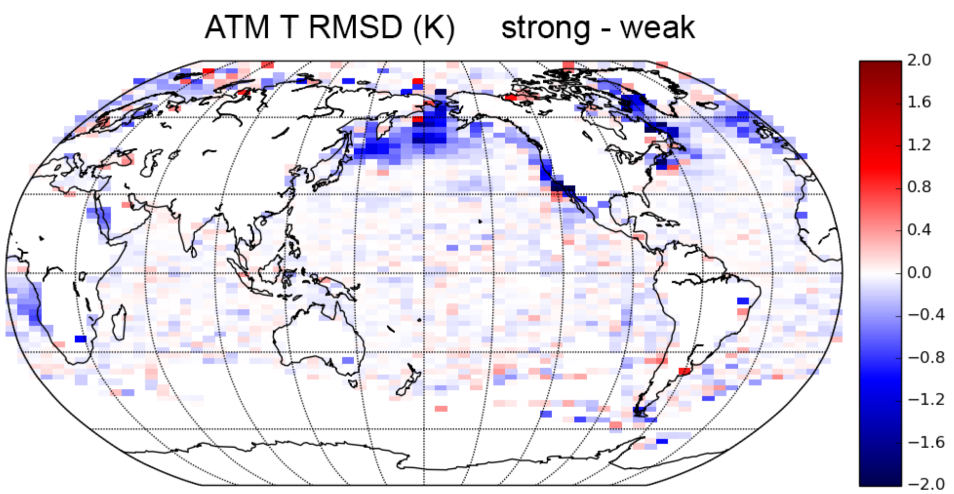


Weakly Coupled DA Atmospheric surface T bias

There is a strong positive temperature bias in the weakly coupled DA in the Pacific and Atlantic oceans, especially near the coasts.

Difference in the RMS errors between strong and weak coupled data assimilation. Blue: Strong is better





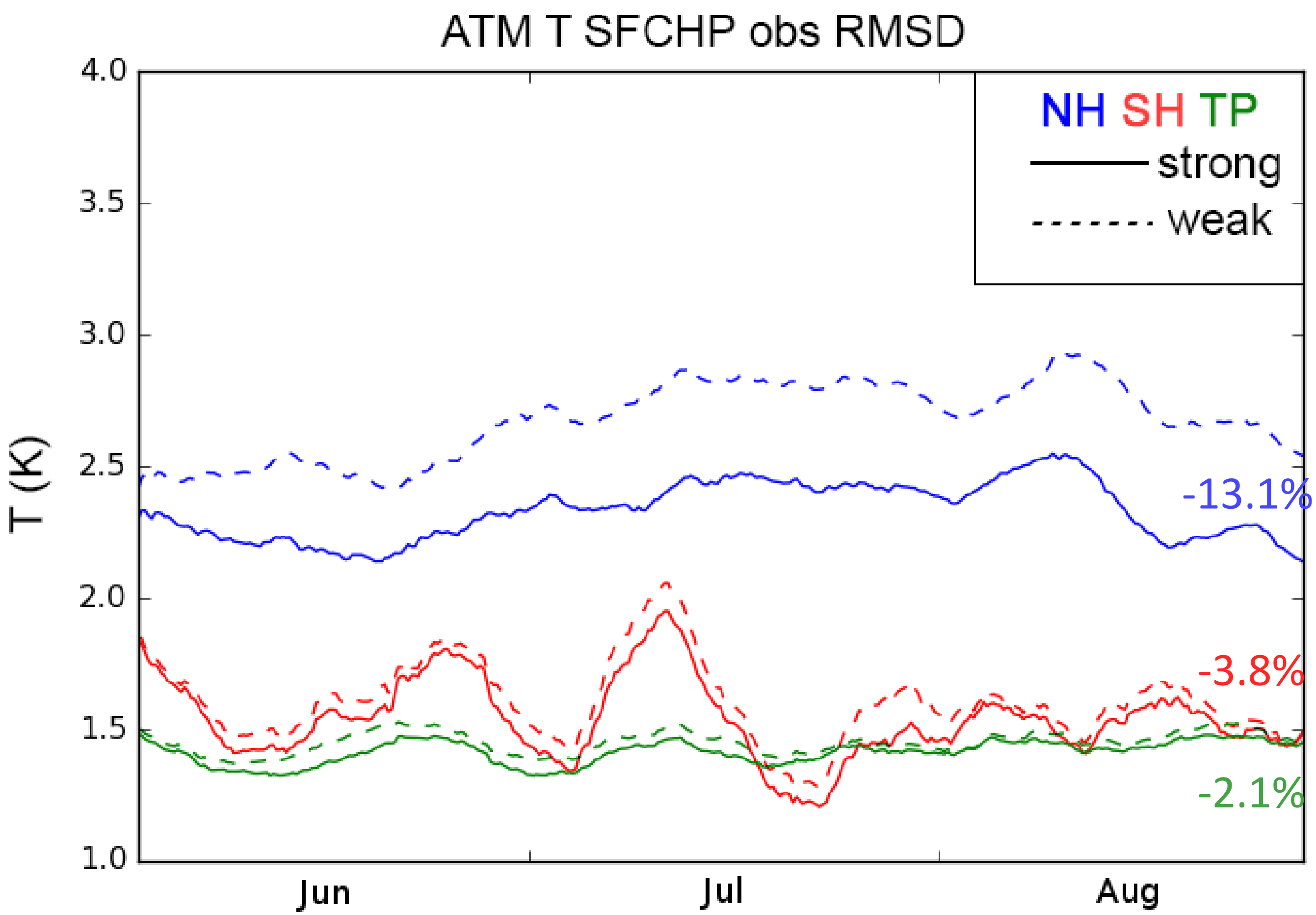
Strong-Weak Coupled DA Atmos. Surface T RMS error

Strong-Weak Coupled DA Ocean 5m T RMS error

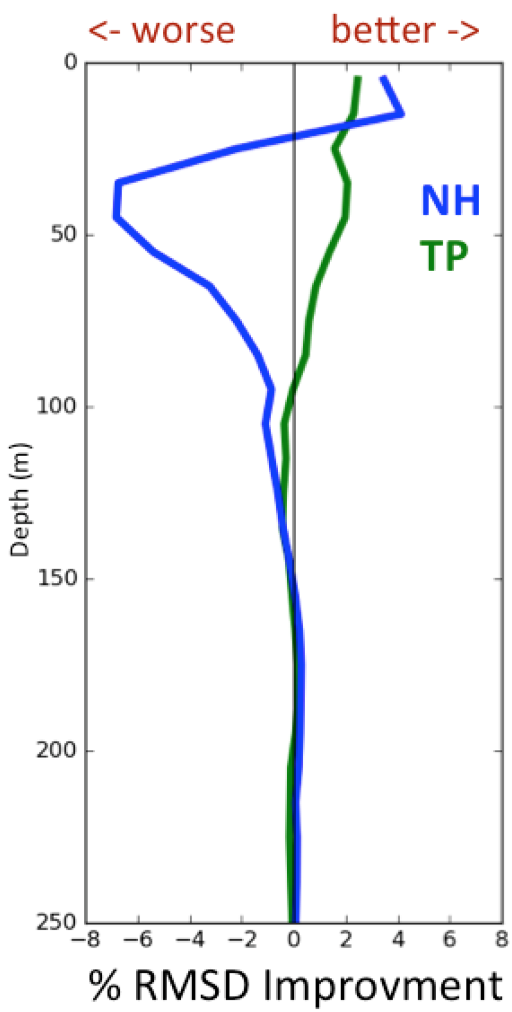
- The ocean improved its bias because it assimilated surface atmospheric observations.
- The improved coupled ocean model in turn reduced the atmospheric errors.

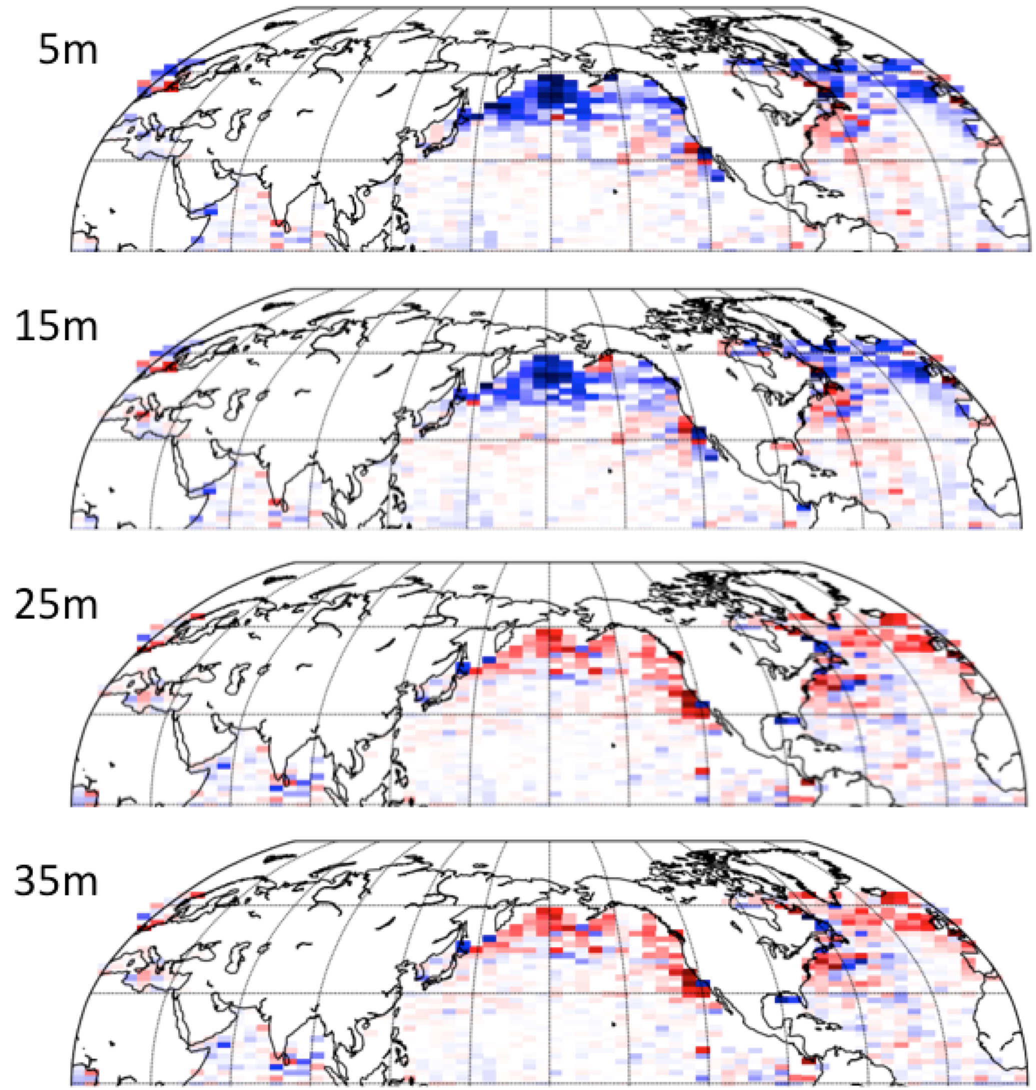
Strongly Coupled CFS - results

- Errors in 6 hour background for ATM T are greatly reduced in the NH



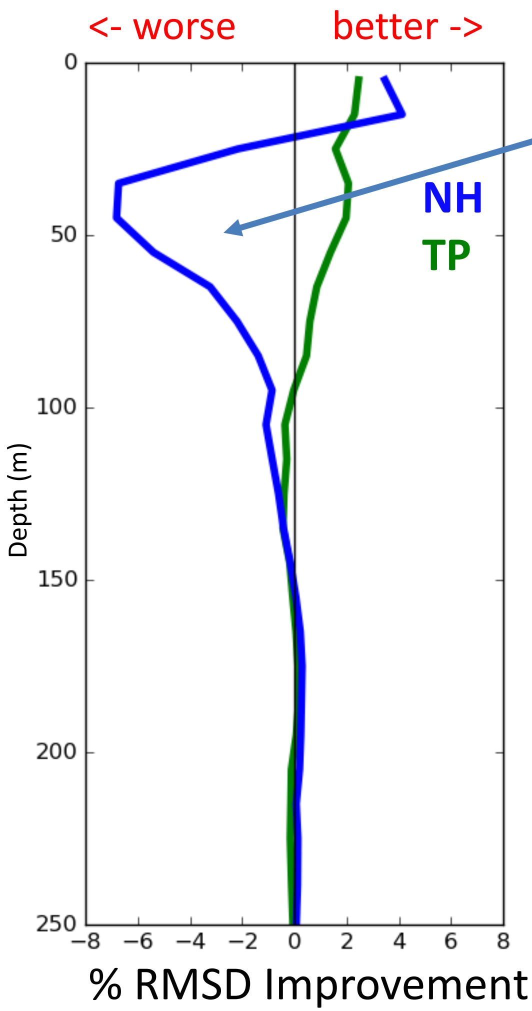
Strongly Coupled CFS - results



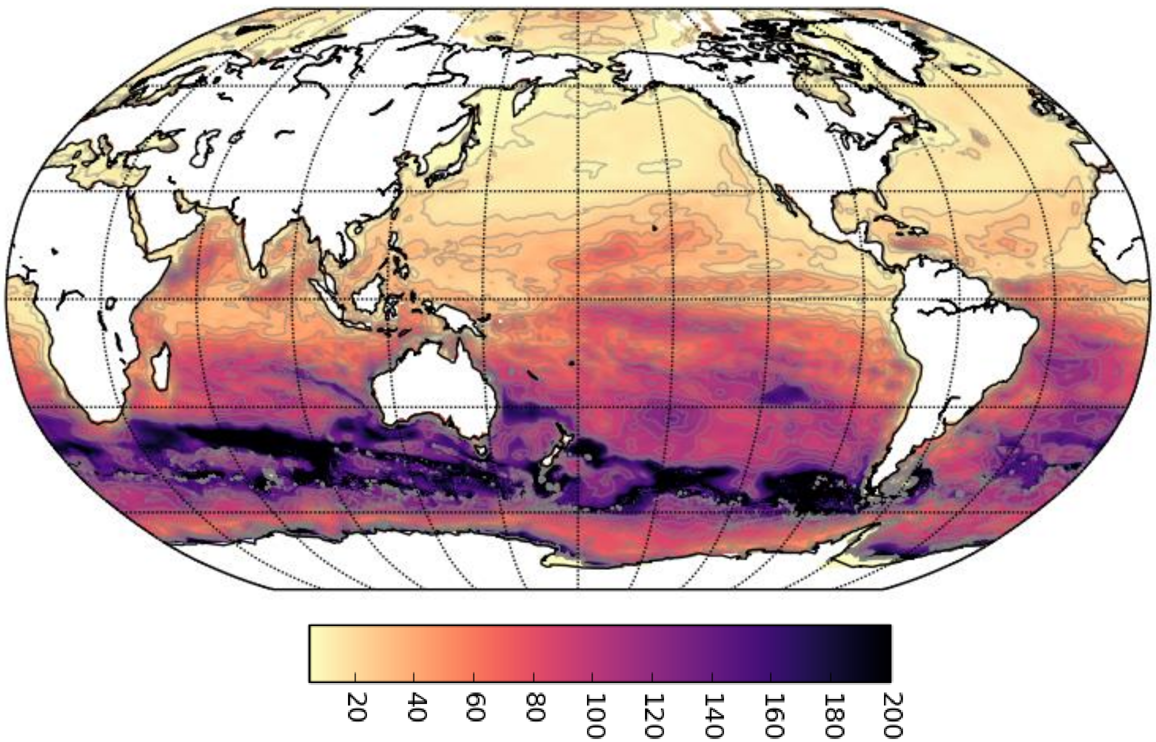


Strongly Coupled CFS - results

Caused by naïve fixed vertical localization of ATM observations into ocn (σ=50m). Need to limit impact to mixed layer only.



Mixed Layer depth (JJA)



Ultimate Goal…

- **CFSv3 -** NCEP transitioning to **gain hybrid-GODAS**, based on LETKF for the **ocean**.
- Increased potential after that for an operational strongly coupled hybrid-LETKF **global** DA system