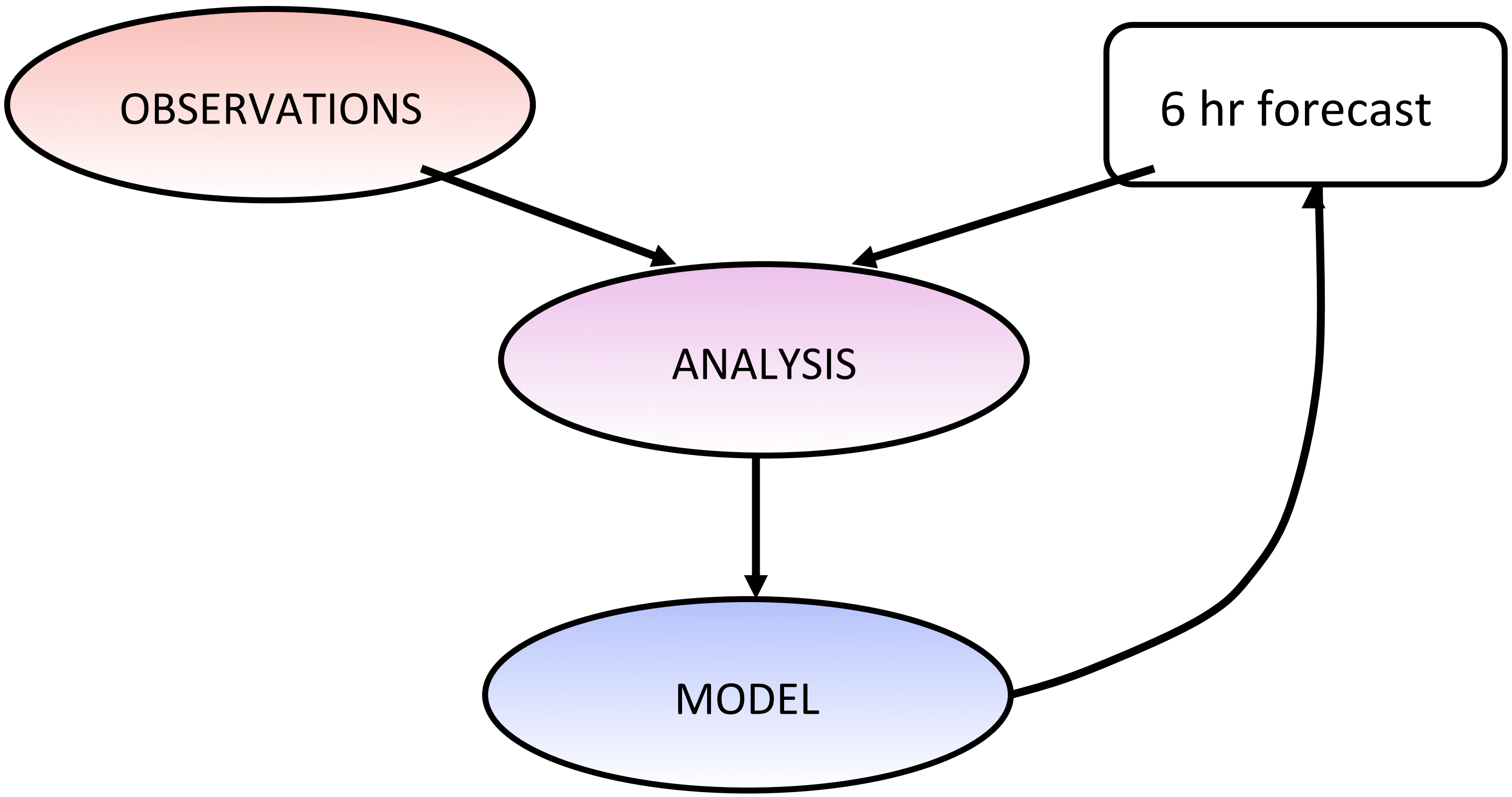
New opportunities for new reanalyses

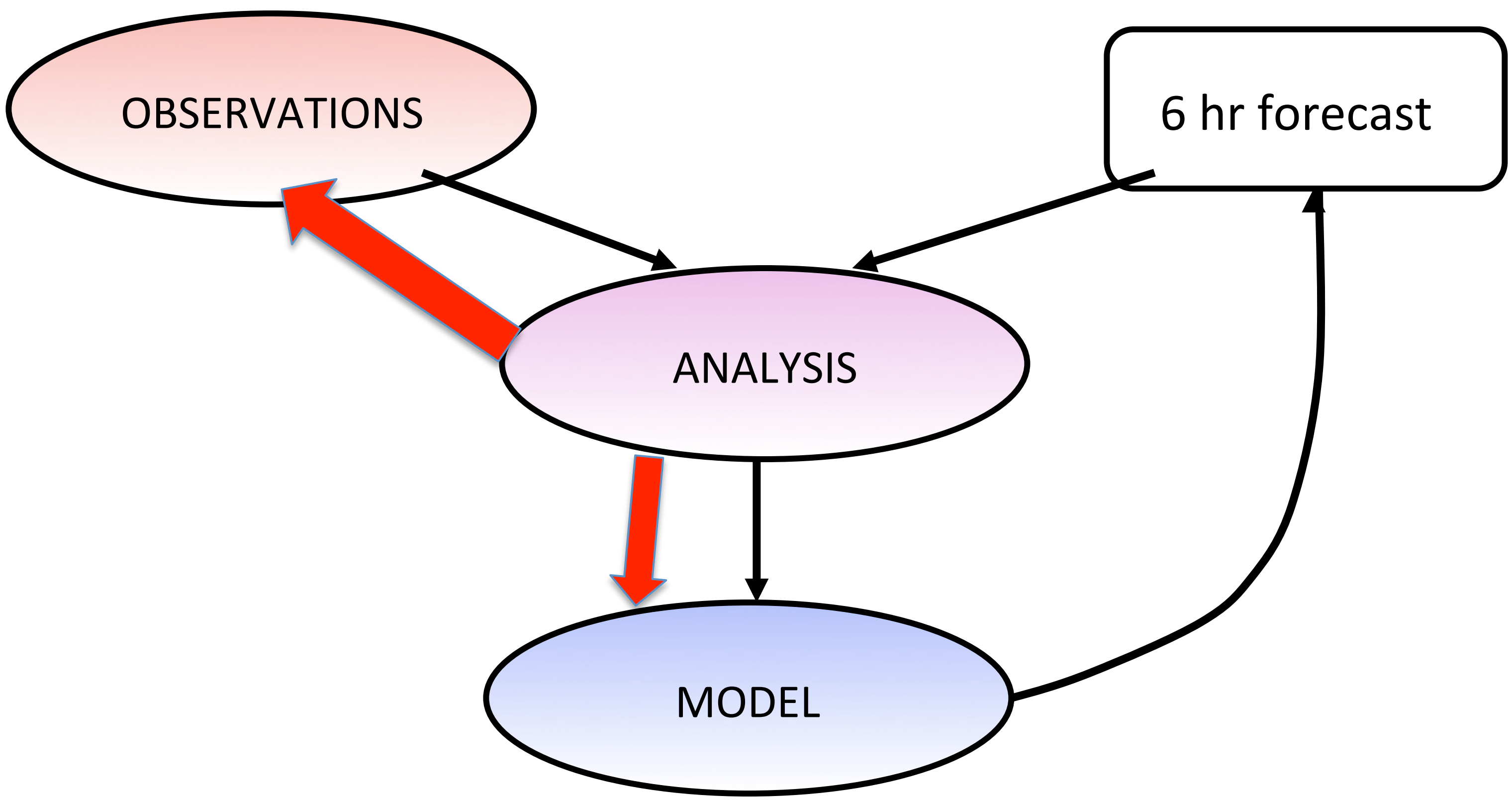
Eugenia Kalnay
with
Yan Zhou and Junye Chen
for the correction of analysis jumps
Thanks to Fanglin Yang for providing us with GDAS analysis increments!

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

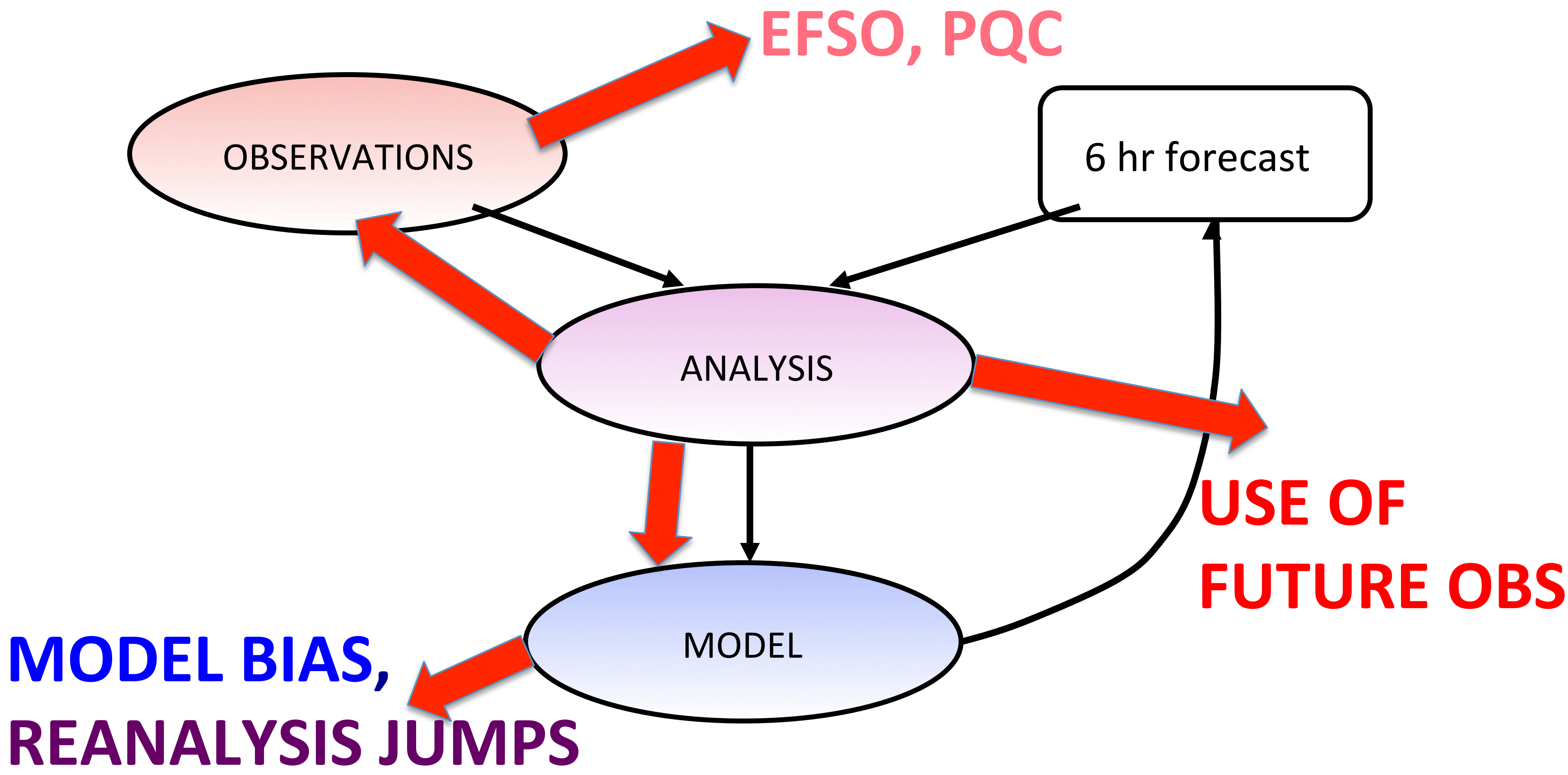
These improvements are done **independently**



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



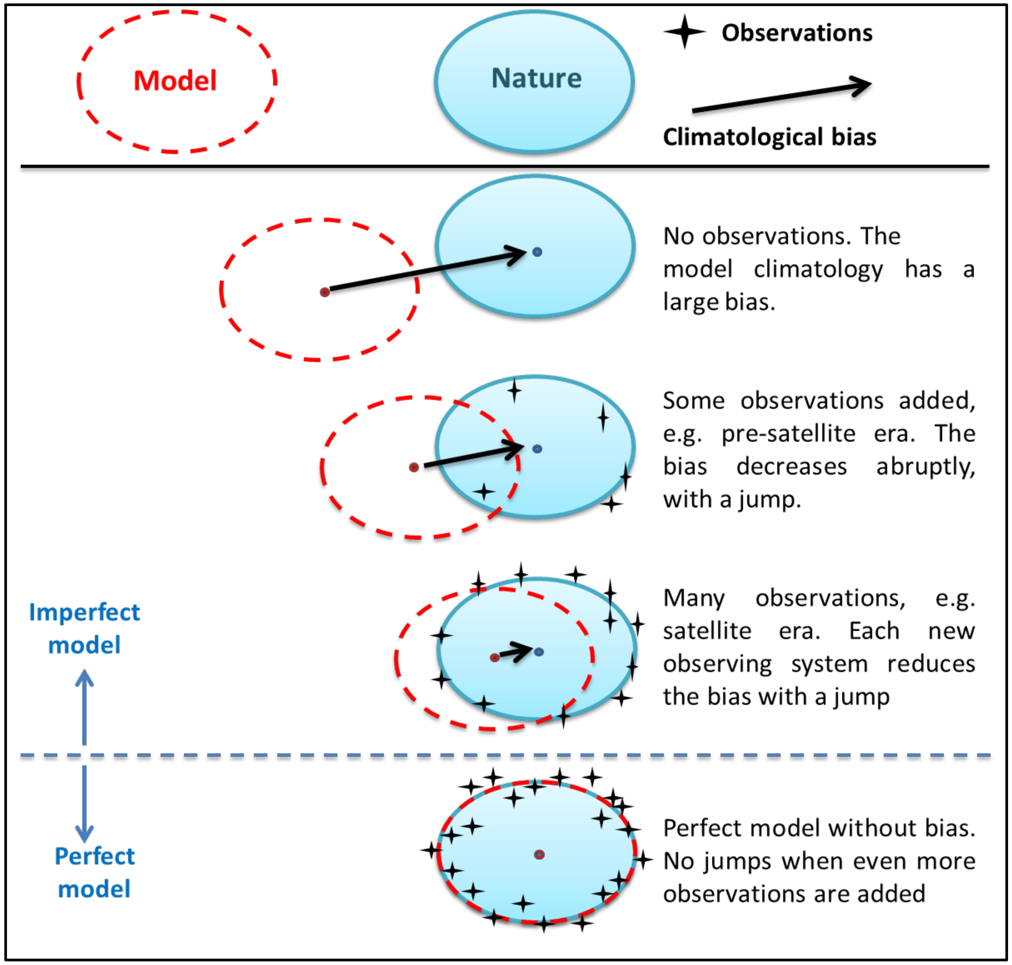
New Opportunities for Reanalysis: We can improve and use future observations, correct model bias, minimize Reanalysis Jumps with new observations



New opportunities for reanalysis: Outline

- Model bias and reanalysis jumps
- How to estimate and correct model bias
- How to minimize reanalysis jumps
- Proactive QC: Find and delete obs that are flawed but passed the regular QC.
- Increase the accuracy of the analysis by using future data, not just past data.
- Strongly coupled data assimilation.

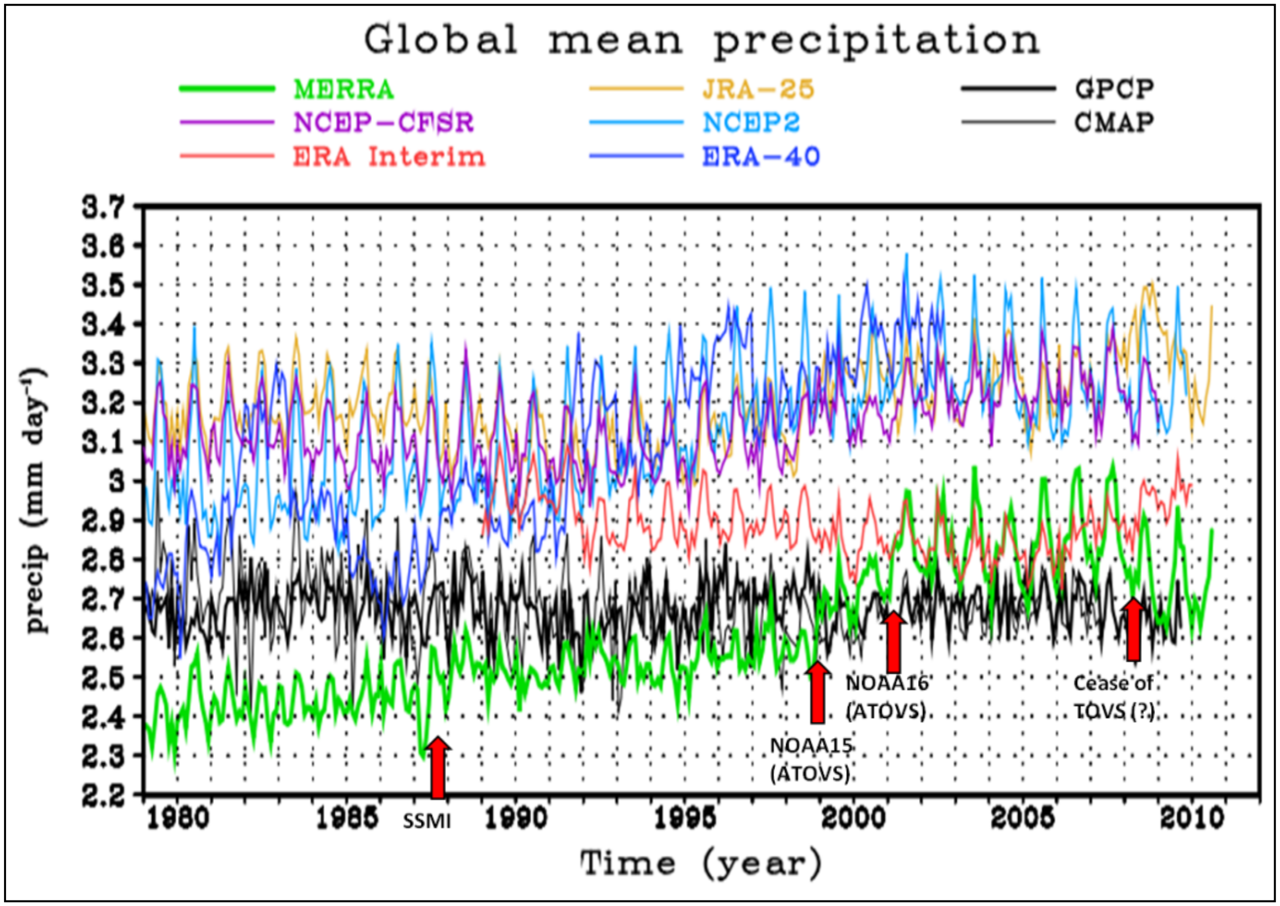
Why do we get reanalysis jumps? Model bias!



A schematic of “climate jumps” associated with observing system changes

- **The climatological bias between the forecast model and the nature decreases with a *jump* when a new observing system was assimilated.**
- **The purpose of Yan Zhou’s dissertation is to find a solution to minimize the “climate jumps” associated with observing system changes.**

Example: MERRA global mean precipitation



Global monthly mean precipitation (mm/day) time series for MERRA (green), several other reanalyses, and GPCP and CMAP (black) (Chen et al., 2012)

- **Jumps in the MERRA global mean precipitation time series appeared simultaneously with introducing or ceasing different types of satellite observations, like SSM/I and ATOVS (red arrows)**

How can we estimate and correct model bias?

- 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 = Analysis Increments (**AI**) = - Initial (linear) model errors
- **Time average of AI is the best estimate of the error growth due to model bias in 6 hr**
- Danforth, Kalnay and Miyoshi (DKM-2007) estimated the 6hr errors of the SPEEDY model.

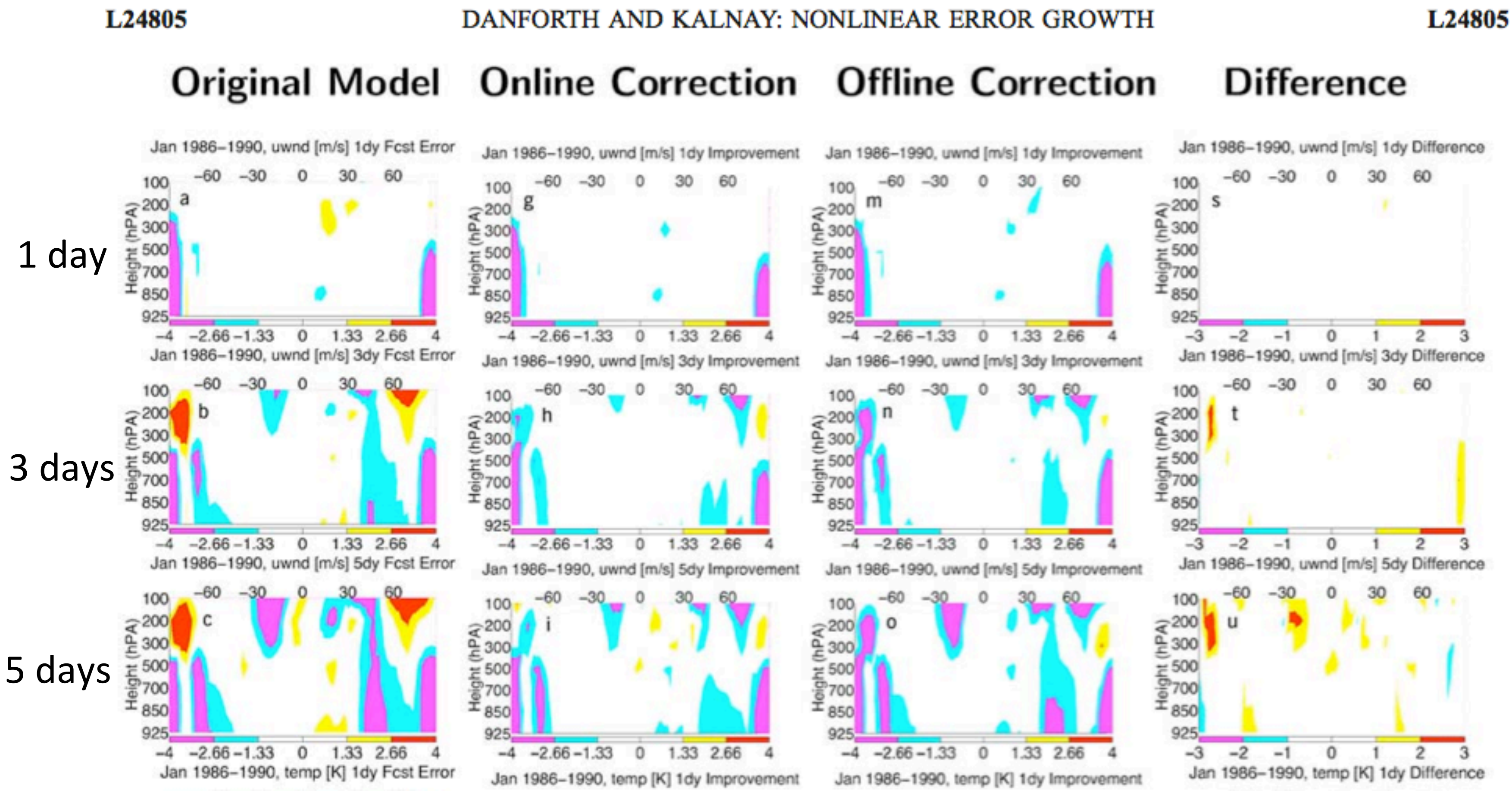
DKM-2007 results

- Estimated the monthly mean 6hr forecast bias
- Corrected the model by adding (–bias/6hr) to each variable time derivative, at each grid point.

Results

- The bias correction after 3 or 5 days was the same as the best *a posteriori bias* correction.
- But the random errors were **smaller**.
- The dominant EOFs of the 6hr debiased forecast errors were the errors in the diurnal cycle.
- It was possible to estimate the systematic errors for anomalies (e.g., ENSO, lows over land or over ocean)

The model corrected online did about as well as the model statistically corrected off-line

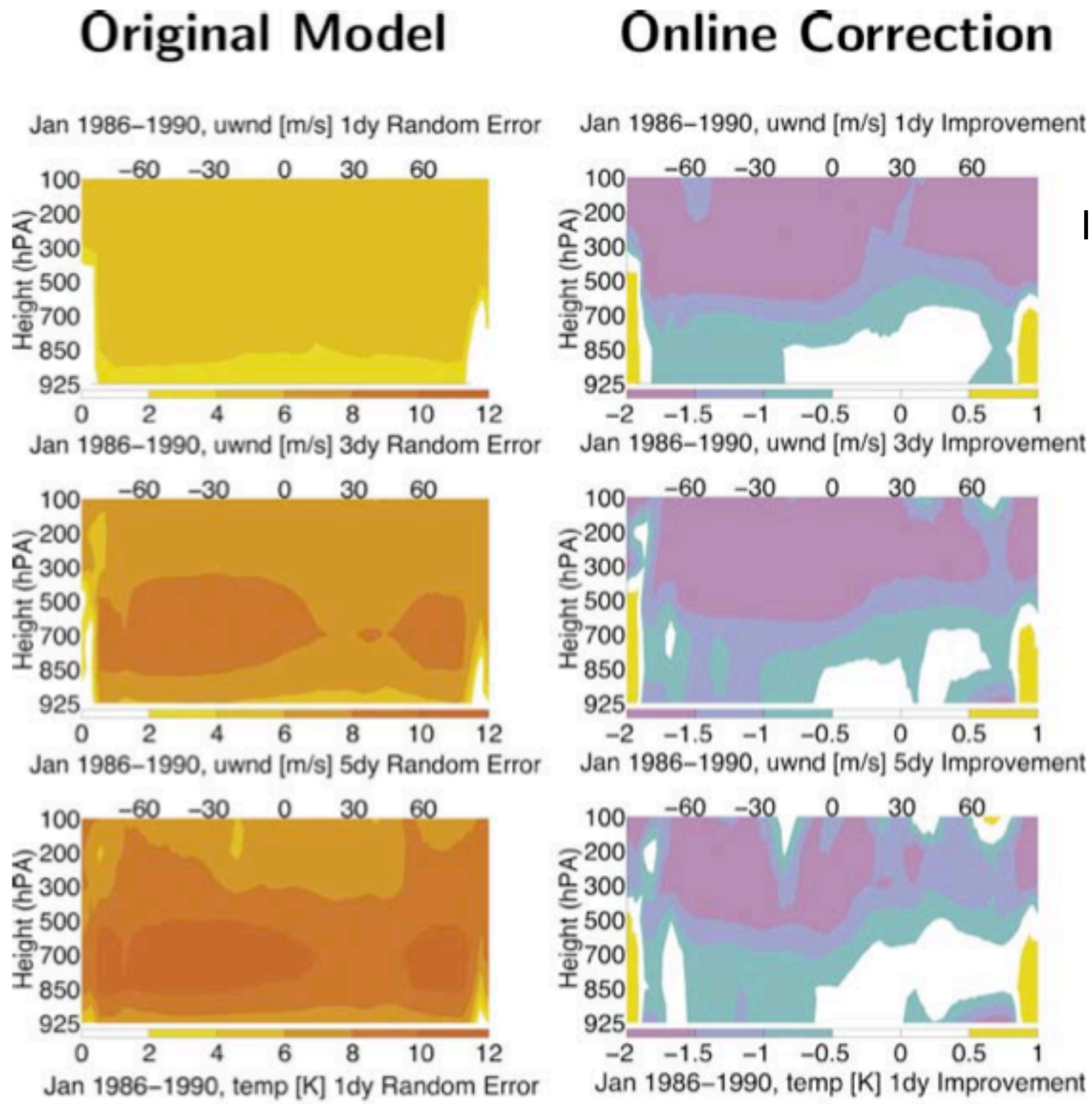


But the random errors were significantly smaller!

1 day

3 days

5 days

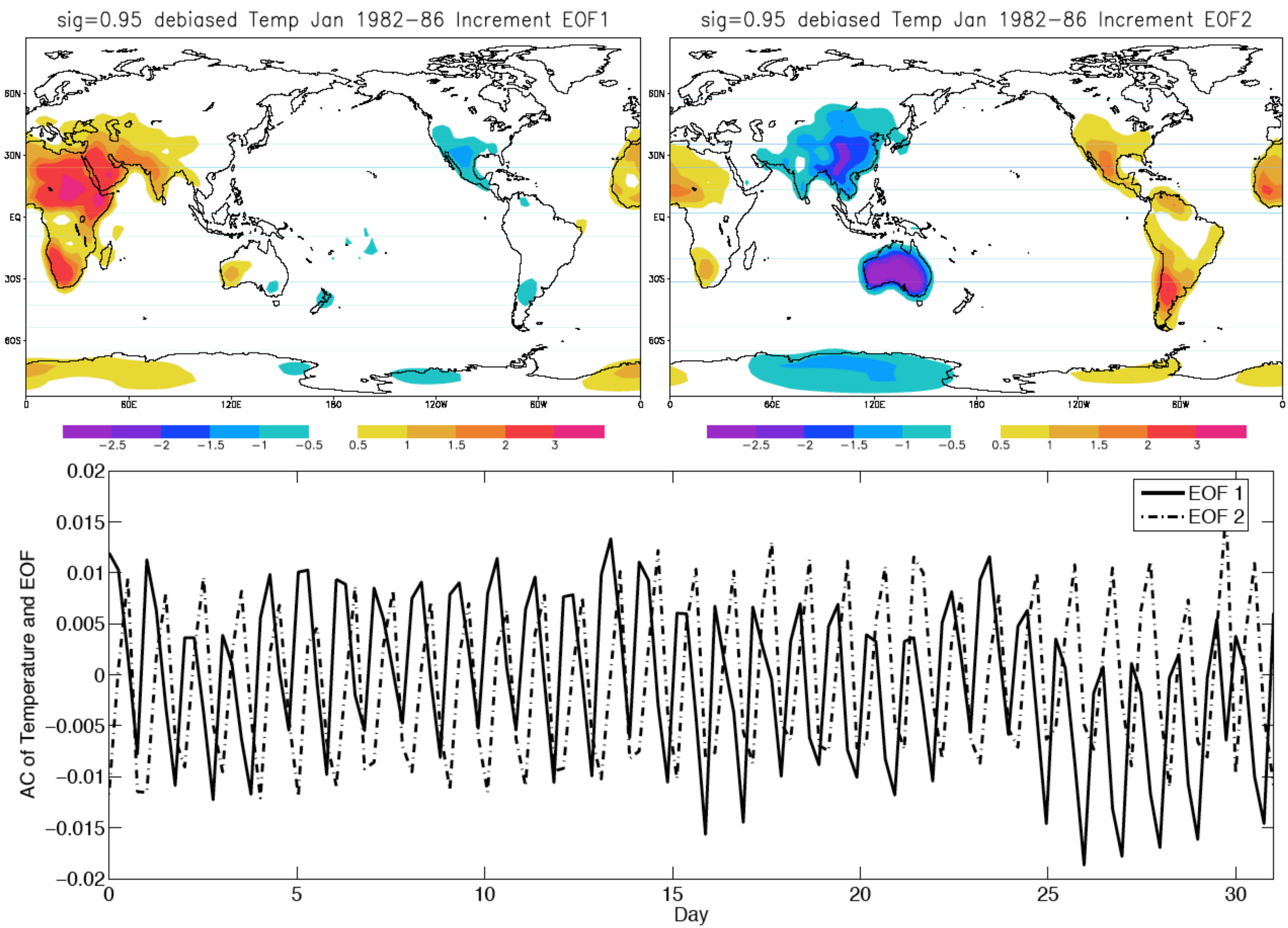


Improvements

How to find the diurnal cycle model errors using EOFs from a Reanalysis (Danforth et al., 2007)

Estimated the average SPEEDY model error (bias) by averaging over several years the 6 hour forecast (started from reanalysis) minus the reanalysis.

Then they computed the EOFs of the anomaly in the model error, and found two dominant EOFs representing the model error in representing the diurnal cycle:



sig=0.95 debiased Temp Jan 1982−86 Increment EOF1
sig=0.95 debiased Temp Jan 1982−86 Increment EOF2
60N
30N
EQ
30S
60S
0
60E
120E
180
120W
60W
−2.5 −2 −1.5 −1 −0.5
0.5 1 1.5 2 3
EOF 1
EOF 2
AC of Temperature and EOF
Day
0.02
0.015
0.01
0.005
0
−0.005
−0.01
−0.015
−0.02
0 5 10 15 20 25 30

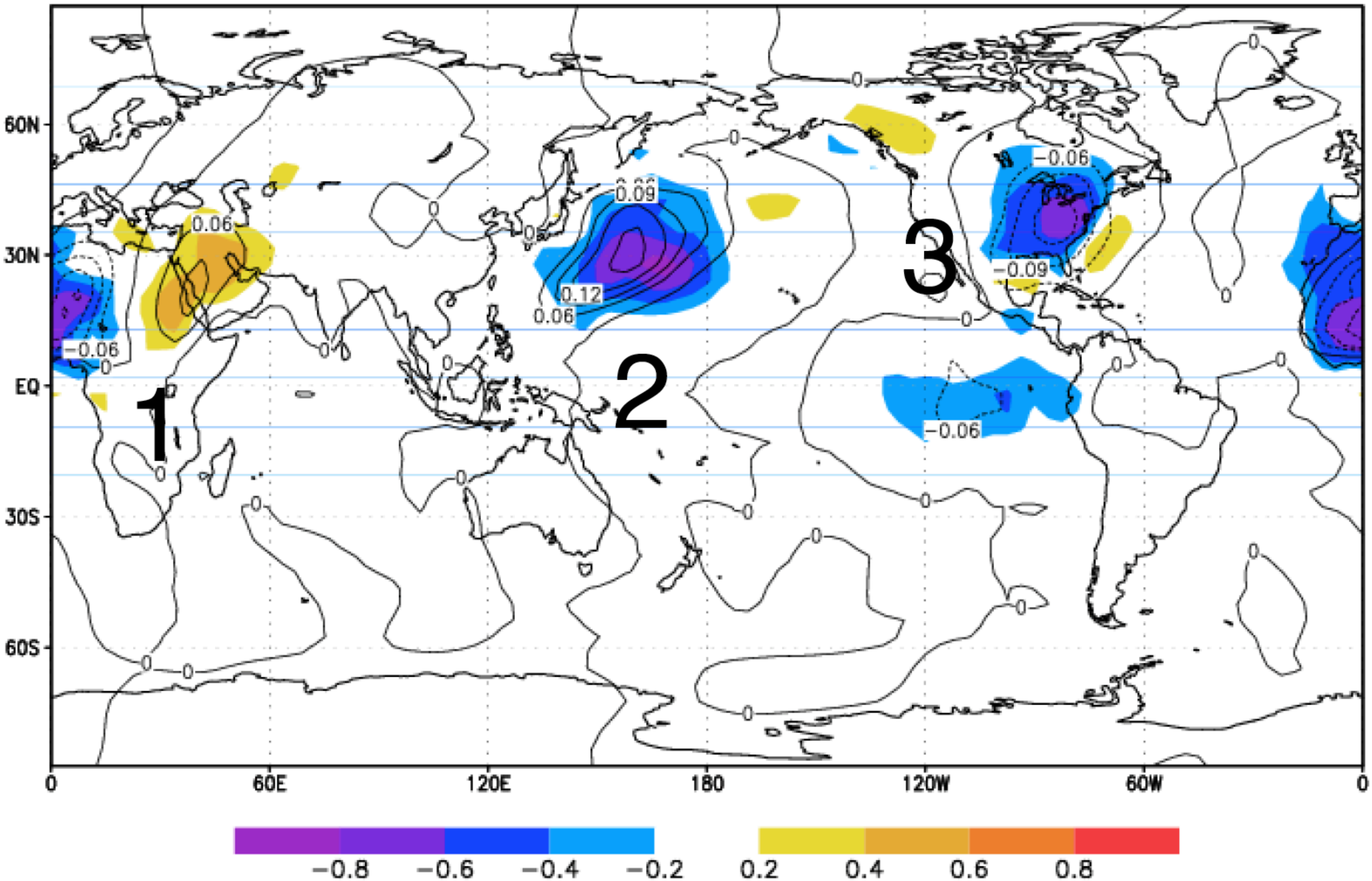
How to find the state dependent errors using coupled SVD's (Danforth et al., 2007)

Three leading coupled SVD's of the covariance of 6 hr forecast errors and corresponding model state anomaly for T at sigma=0.95. Contours: state anomaly, colors: heterogeneous correlation with forecast errors.

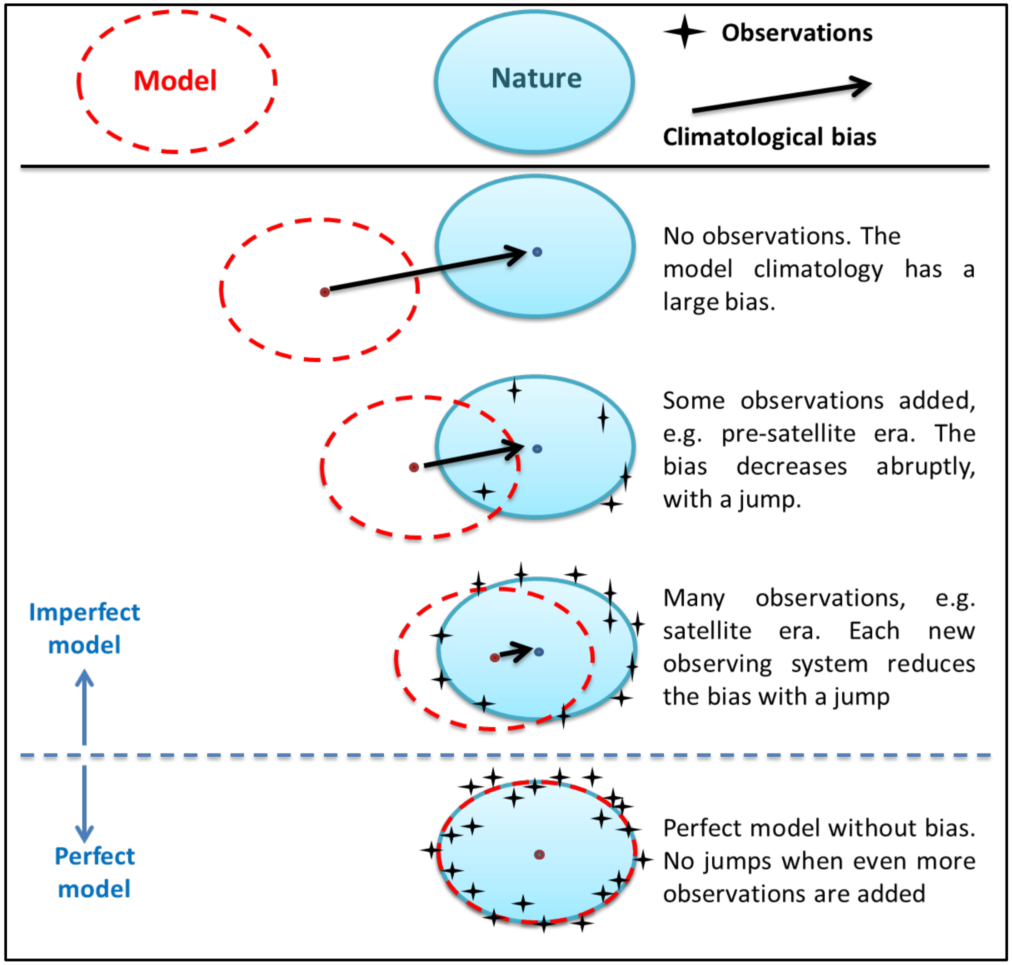
Over land, the corrections suggest the anomalous temperatures are too strong, and over ocean too weak and too far to the west.

This can be extended to improving forecasts using coupled SVD's

sig=0.95 Temp Jan 1982-86 Correlation Maps



Why do we get reanalysis jumps? Model bias!



- **The climatological bias between the forecast model and the nature decreases with a *jump* when new obs are assimilated. These jumps are the worst deficiency of reanalyses, especially long reanalyses.**
- **One solution <u>is not to include new observations</u> (Compo et al., 2009)!**
- **Another solution is to estimate and correct the jumps.**

Analysis increments (AI) estimate model errors. Average analysis increments estimate model biases.

- Fanglin Yang has kindly provided us with the GDAS AI for the years 2014 and 2013.
- We (Jim Carton, Kriti Bharghava and I) plan to
  - Study the monthly means of the AI estimation of the model bias.
  - Test whether we can apply the correction AI/6hr to the GFS model
  - Look at the diurnal cycle errors: compute the EOFs of the AI after taking out their mean (bias)

How can we minimize the jumps when we add new observing systems? (Yan Zhou's thesis)

Yan Zhou tested 3 plausible methods to avoid jumps

1- DKM2007 (Based on Danforth-Kalnay-Miyoshi 2007)

2- MERRA (Based on Junye Chen's idea for MERRA)

3- Climatological (suggested as a baseline by B. Hunt)

All 3 methods attempt to find the average change in analysis climatology that the new instrument introduces, and to add it to the analysis previous to the new instrument in order to correct its bias.

The best results were obtained with DKM2007. Next with MERRA. The simple climatological correction was the worst.

How can we minimize the jumps when we add new observing systems? (Yan Zhou's thesis)

- Yan Zhou tested 3 methods:

  N=with new obs; O=only old obs

  $\mathrm{AI}_{\mathrm{N}}^{\mathrm{N}}$ Analysis with New obs, First Guess with New obs

  $\mathrm{AI}_{\mathrm{N}}^{O}$ Analysis with Old obs, First Guess with New obs

  – DKM2007: $\overline{\mathrm{AI}}_{\mathrm{N}}^{\mathrm{N}} - \overline{\mathrm{AI}}_{\mathrm{N}}^{O}$ BEST

  – MERRA: $\overline{\mathrm{AI}}_{\mathrm{N}}^{\mathrm{N}} - \overline{\mathrm{AI}}_{O}^{O}$ IN BETWEEN

  – Climatology: $\overline{\mathrm{A}}_{\mathrm{N}}^{\mathrm{N}} - \overline{\mathrm{A}}_{O}^{O}$ WORST

How can we minimize the jumps when we add new observing systems? (Yan Zhou's thesis)

- The best method she found (DKM2007) can be easily carried out **during** the reanalysis:
- When starting a new obs system, for 1-2 years:
  - Compute the New AI (with new obs system)
  - Compute the Old AI (without the new obs system but using the same first guess as the New AI)
  - Time average of (New AI-Old AI)= $\Delta\overline{\mathrm{AI}} = \overline{\mathrm{New\ AI}} - \overline{\mathrm{Old\ AI}}_{\mathrm{New\ FG}}$
  - **This is the correction in the model bias introduced by the new observations**.
- This should be added to the reanalysis done <u>before</u> the introduction of the new observations.
- It should minimize the reanalysis jumps.
- Cheaper than doing two reanalyses with and without new obs (the "MERRA approach).

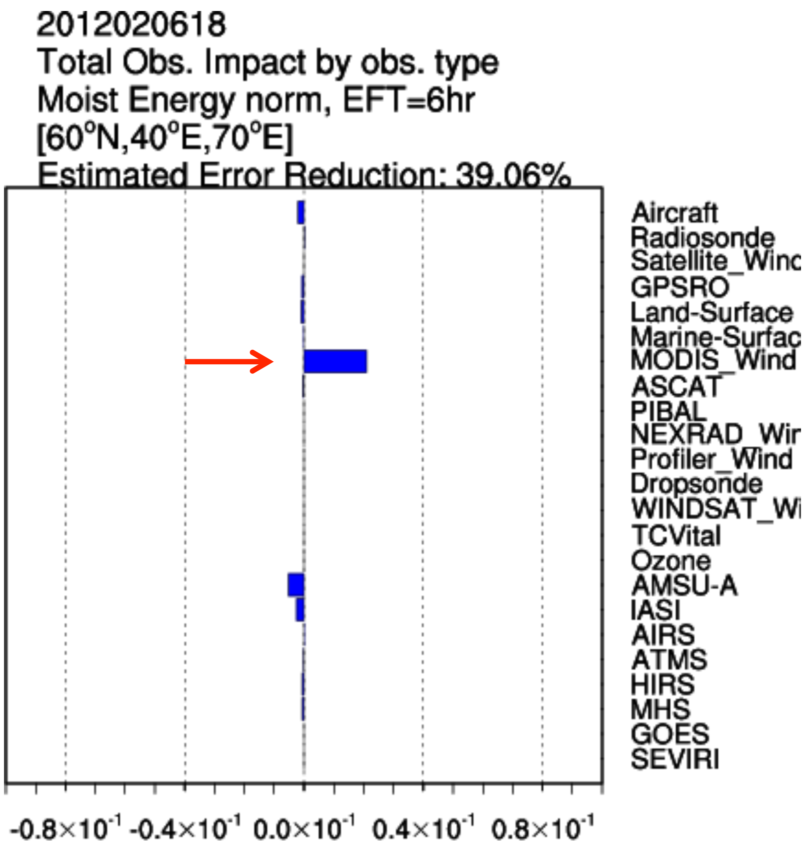
Improve the observations: Ensemble Forecast Sensitivity to Observations and Proactive QC

- Kalnay et al. (2012) derived EFSO
- Ota et al. (2013) tested 24hr 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 withdrawn and collected with useful metadata so they can be improved.
- We call this **Proactive QC**, much stronger than QC.
- Hotta also showed EFSO can be used to tune **R**
- G.-Y. Lien (2014) tested EFSO to identify useful observations of precipitation, with good results.

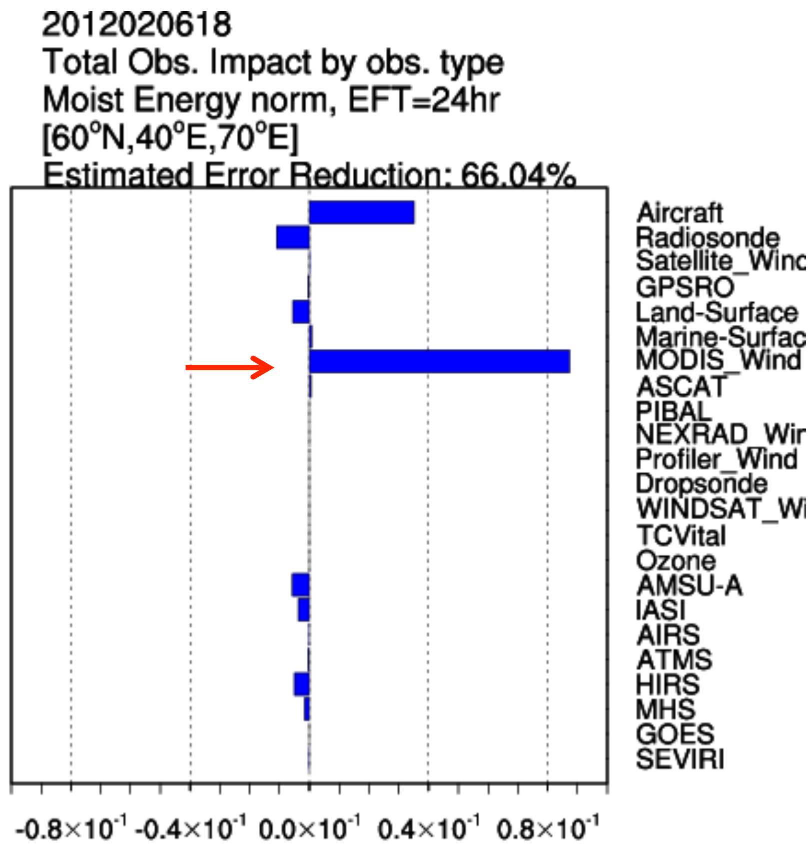
Hotta (2014)

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

FT=06 hr.



FT=24 hr.

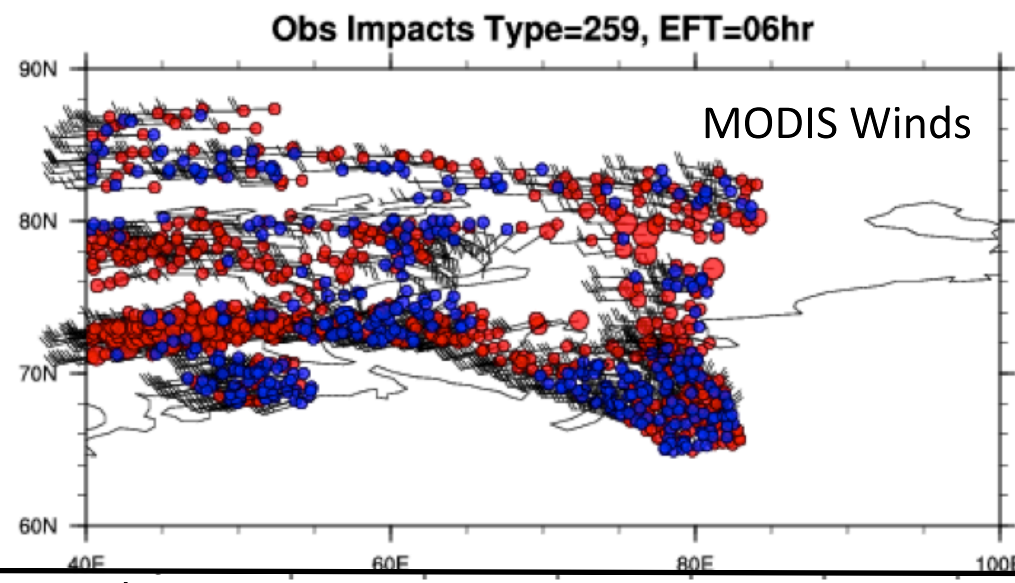


Can identify the bad observations after only 6 hours!

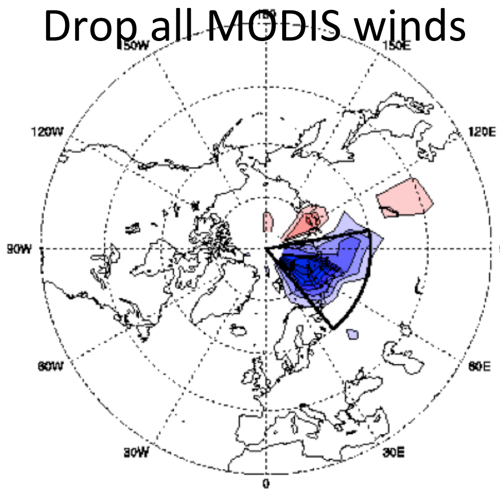
Improve observations:

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

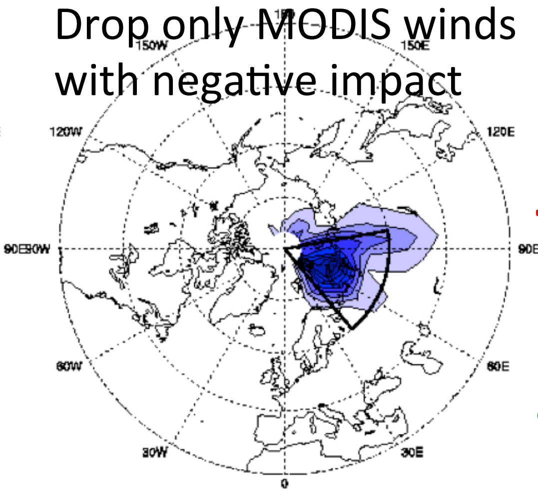
Dr. Daisuke Hotta (2014):
EFSO is able to find whether <u>each</u> observation improves (blue) or makes the 6hr forecast worse (red)



Drop all MODIS winds



Drop only MODIS winds with negative impact

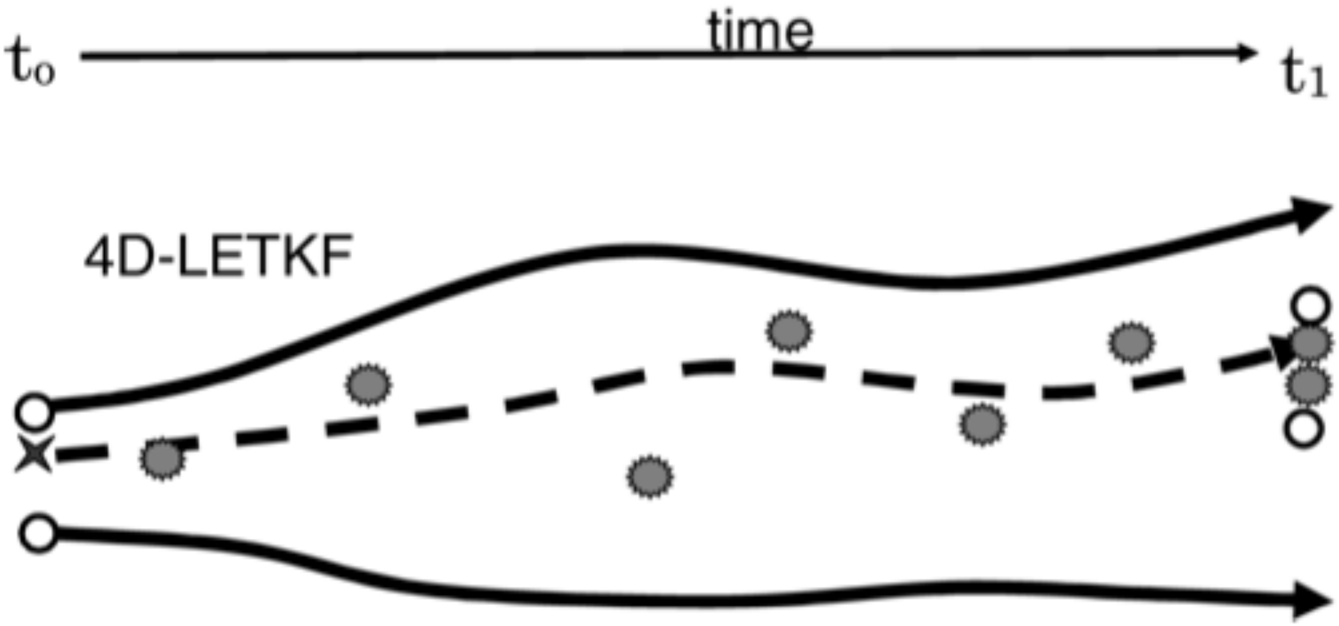


Impact of 6hr PQC on 24hr fcst

PQC with metadata can be used to improve the algorithm!

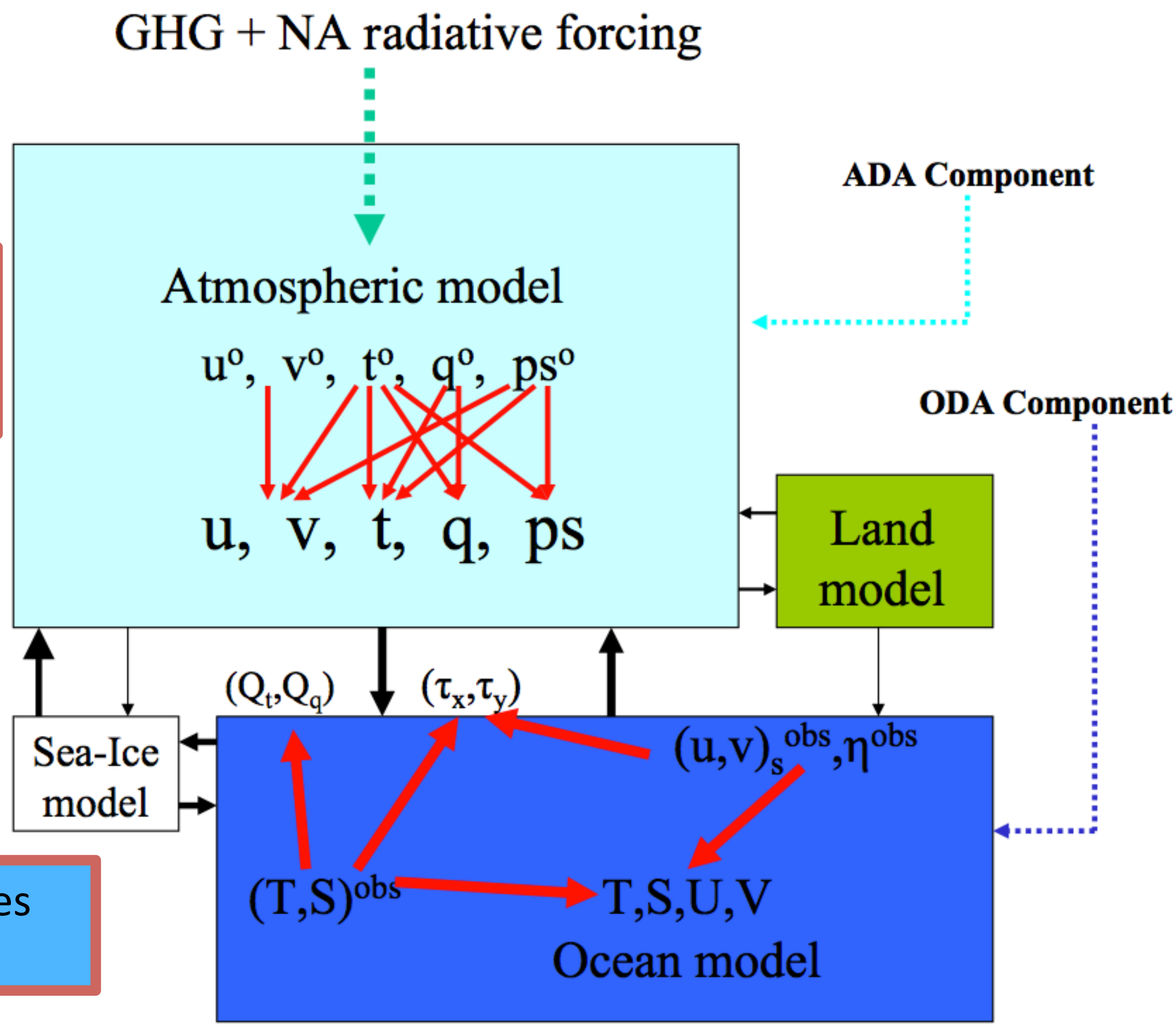
It should accelerate optimal assimilation of new instruments!

More accurate analysis by using future and past data



No-cost smoother: The weights are valid throughout the window. The original analysis uses only past data. The cross corrects it by using the final weights. Since it uses both past and future data, it should be significantly more accurate than the original analysis (like second order differences compared to first order differences). **COST-FREE!**

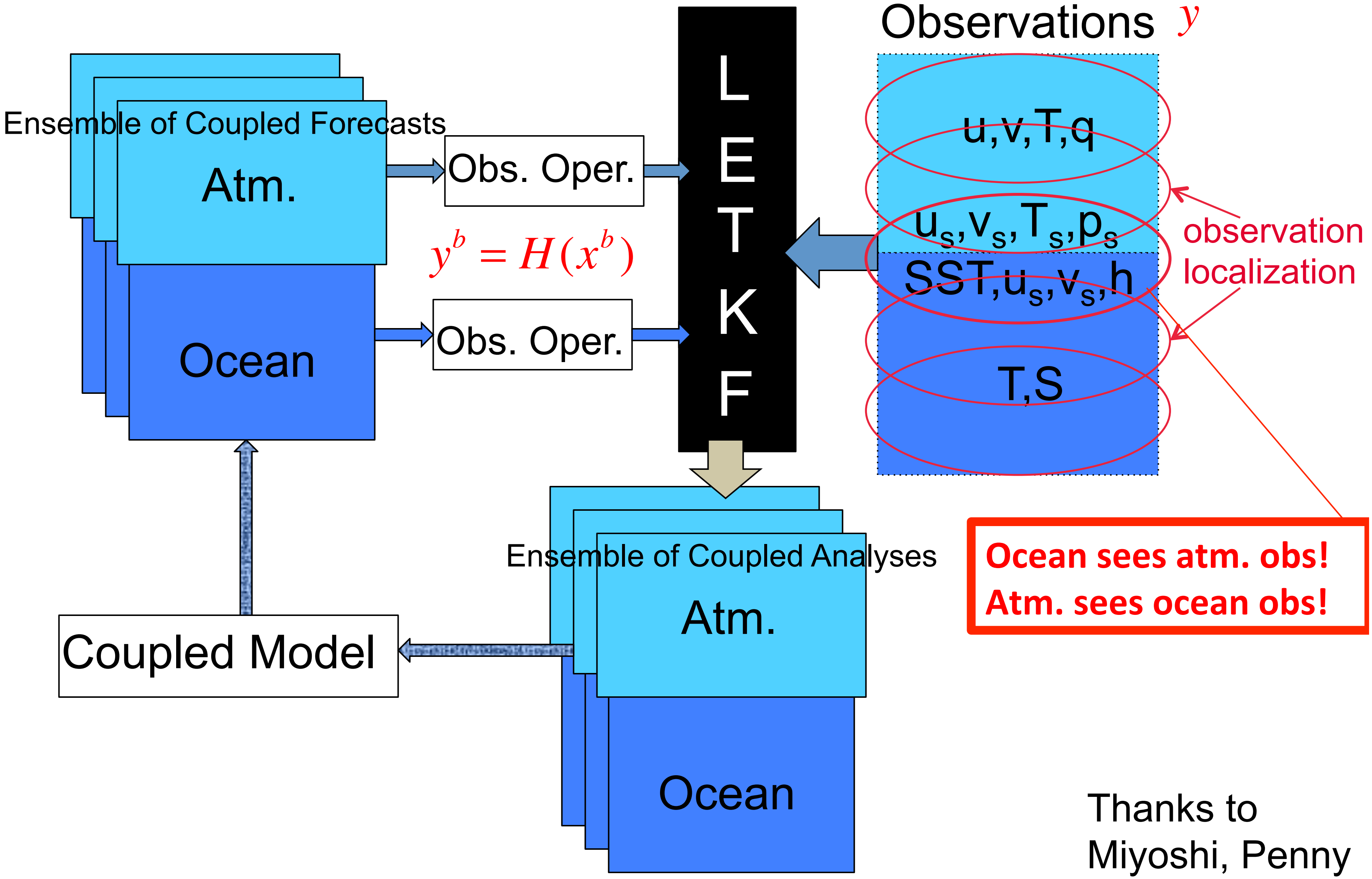
STANDARD (WEAK) COUPLED Data Assimilation
S. Zhang et al.: GFDL Coupled Ocean-Atm EnKF



Atmosphere assimilates only atm. obs.!

Ocean assimilates only ocean obs.!

Our strongly coupled LETKF assimilation



Ocean sees atm. obs!
Atm. sees ocean obs!

Thanks to
Miyoshi, Penny

Impact of strong coupling of the ocean-atmosphere LETKF (Travis Sluka)

- 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

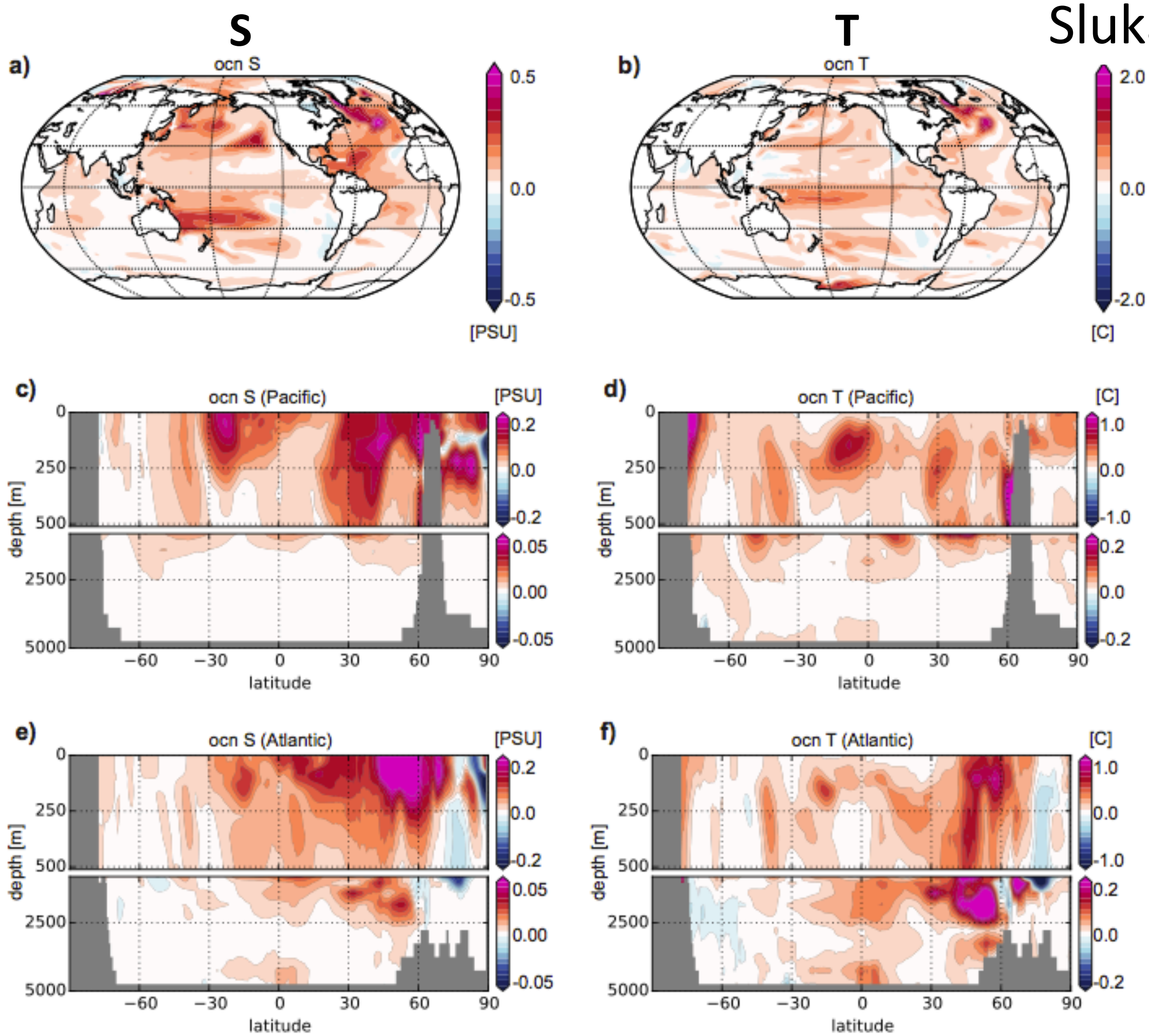
Experiments: 1) Only atmos. obs.

(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).

Results: Red means STRONG DA is better!

Sluka et al., in preparation

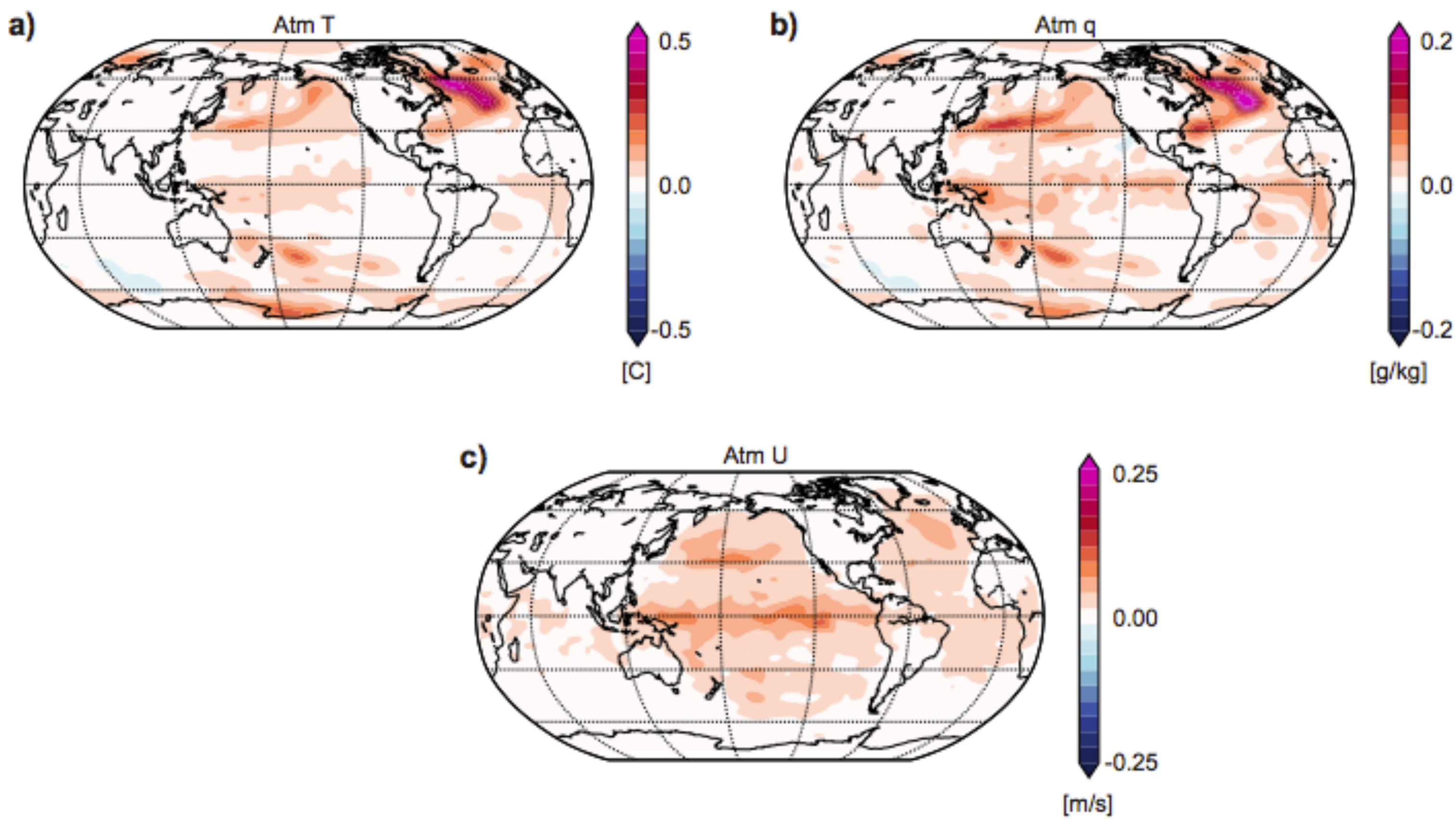


- With Strongly Coupled DA, the errors in temperature and salinity decrease by about 50%.
- The improvements reach the lower levels.

Results: Red means STRONG DA is better!

Sluka et al., in preparation

In turn, with Strongly Coupled DA, **the ocean improved by assimilating atmospheric observations improves the atmosphere**!



Summary

- We should take advantage of the opportunities that advanced DA provide!
- Estimate and correct the jumps introduced by new observing systems
- The best method is DKM2007 (Yan Zhou's thesis). The correction can be trained in 1-2 yrs. Low cost.
- Proactive QC: capture and delete flawed observations that survived the regular QC.
- Use no-cost smoother to improve the analysis at the beginning of the time window using future observations.
- Do strongly coupled data assimilation!!!!