# New Applications of Data Assimilation to Improve Models and Observations

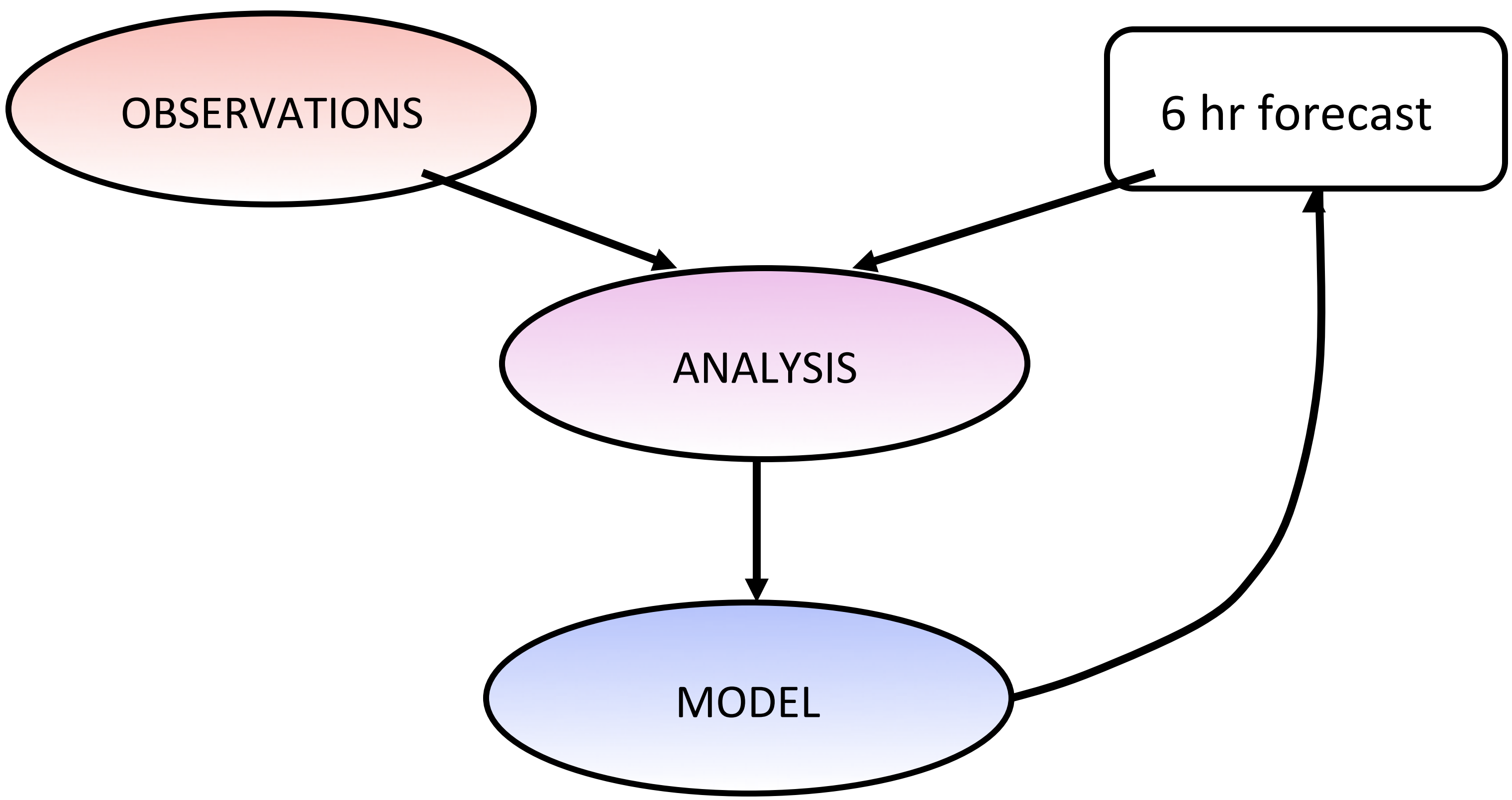
Kriti Bhargava, E Kalnay, J Carton: Models

T.C. Chen, D. Hotta, Y. Ota, E. Kalnay: Observations

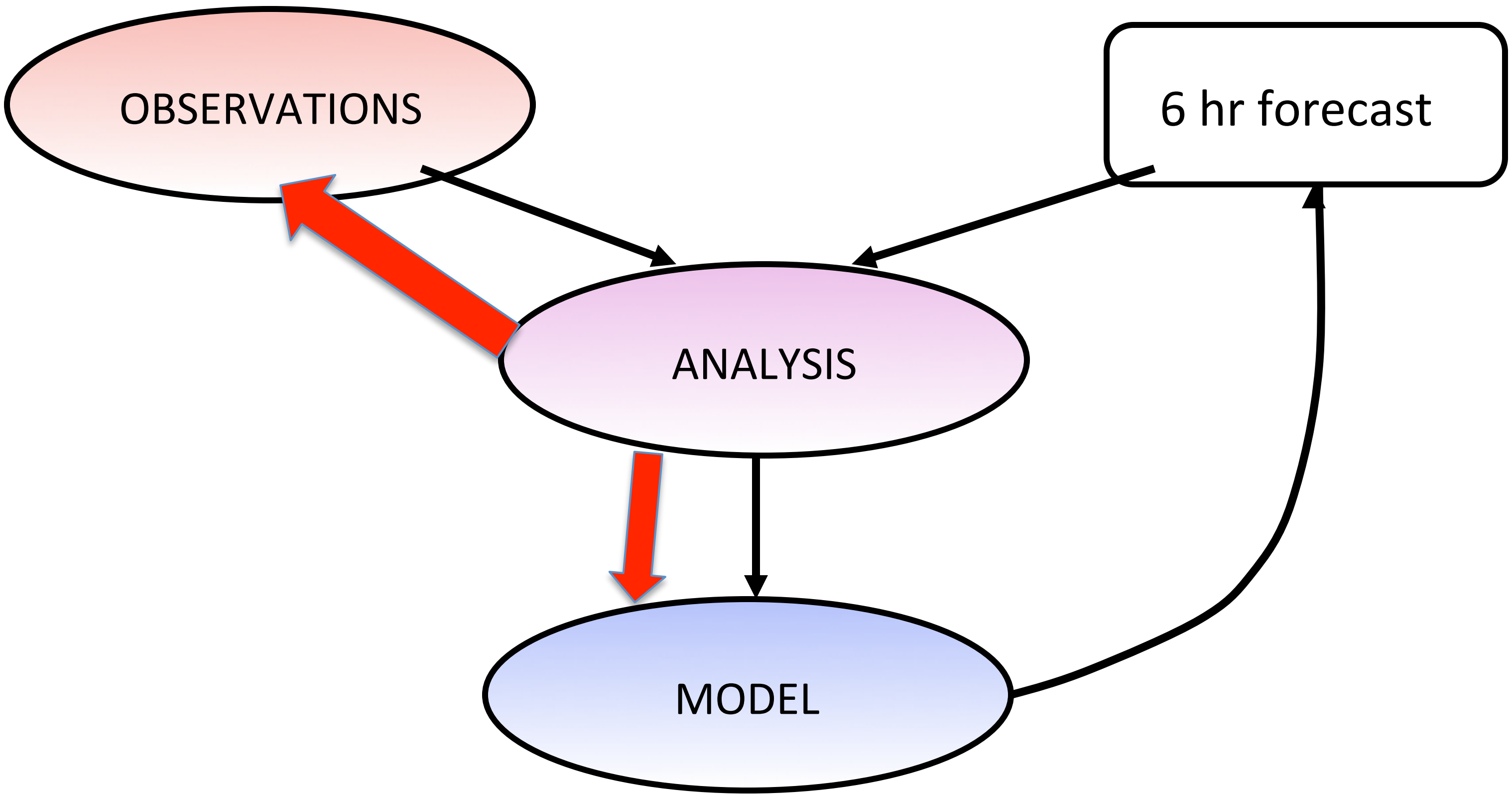
**7th EnKF Data Assimilation Symposium**

**PSU**

**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

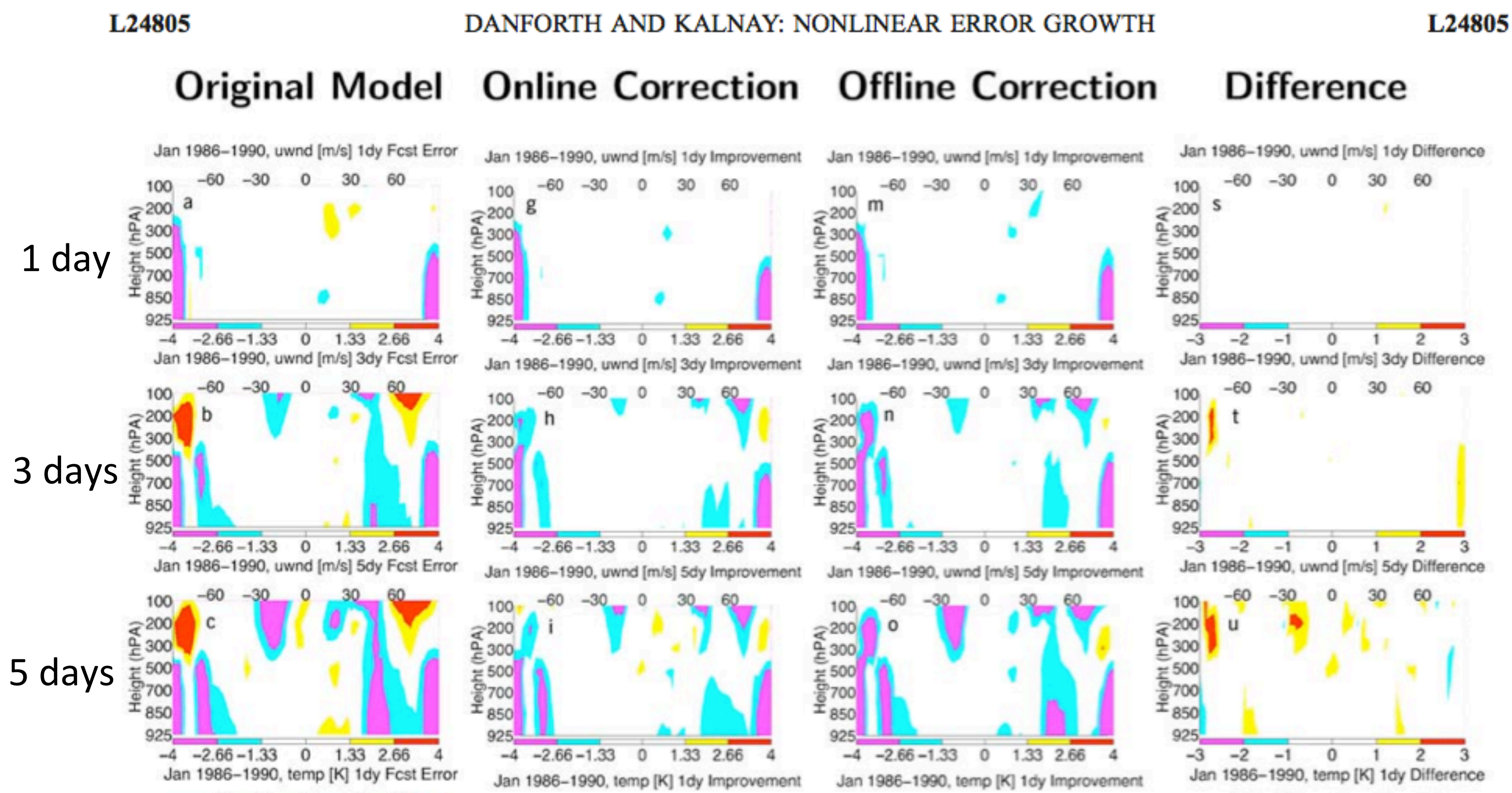


# 1) How can we estimate and correct model bias?

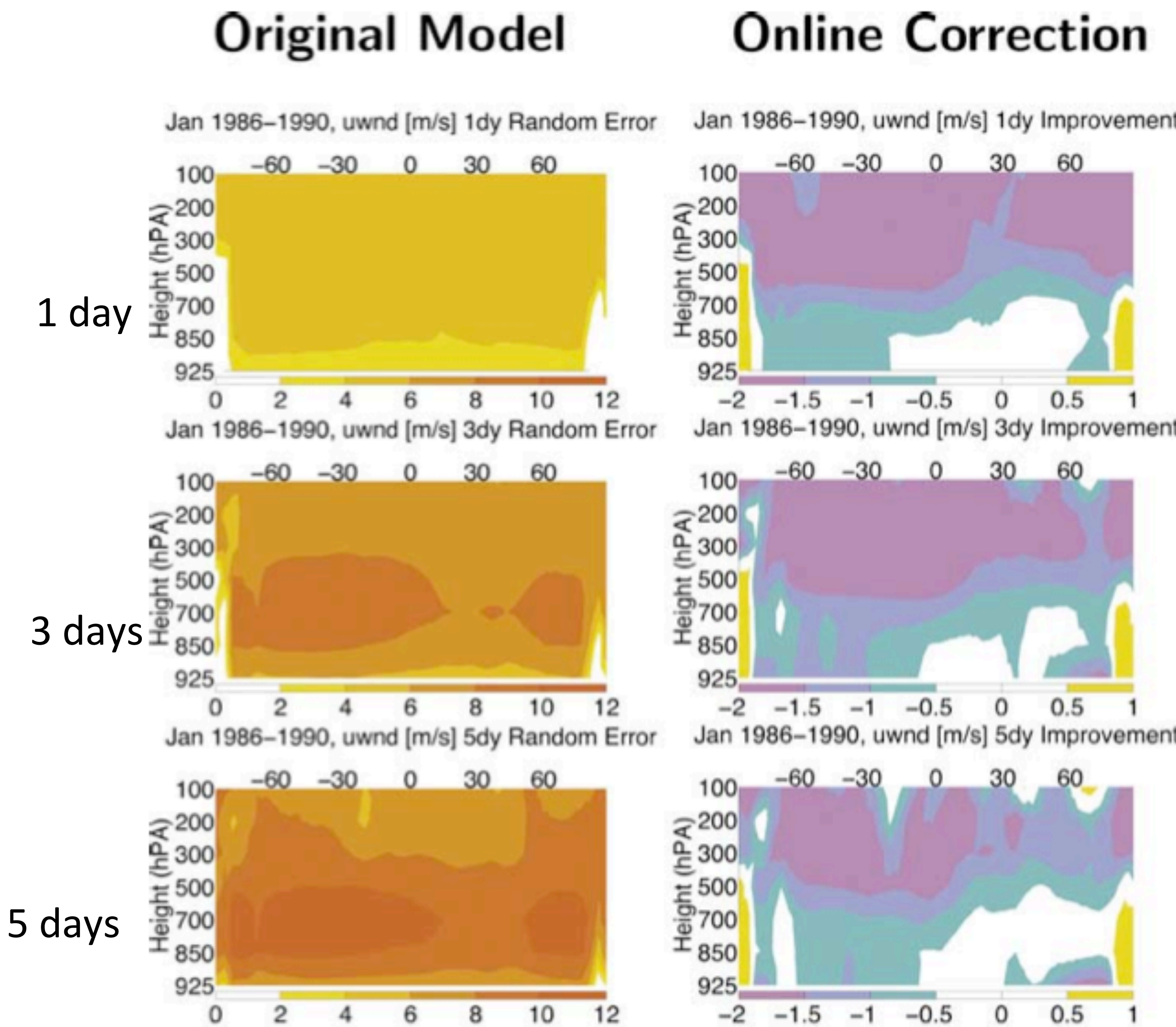
Kriti Bhargava, Jim Carton, Eugenia Kalnay, with Fanglin Yang

- 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.
- **The 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.
- Estimated the average SPEEDY model error (bias) by averaging over several years the 6 hour forecast (started from reanalysis R1) minus the reanalysis.

# The model corrected online did at least as well as the model statistically corrected offline

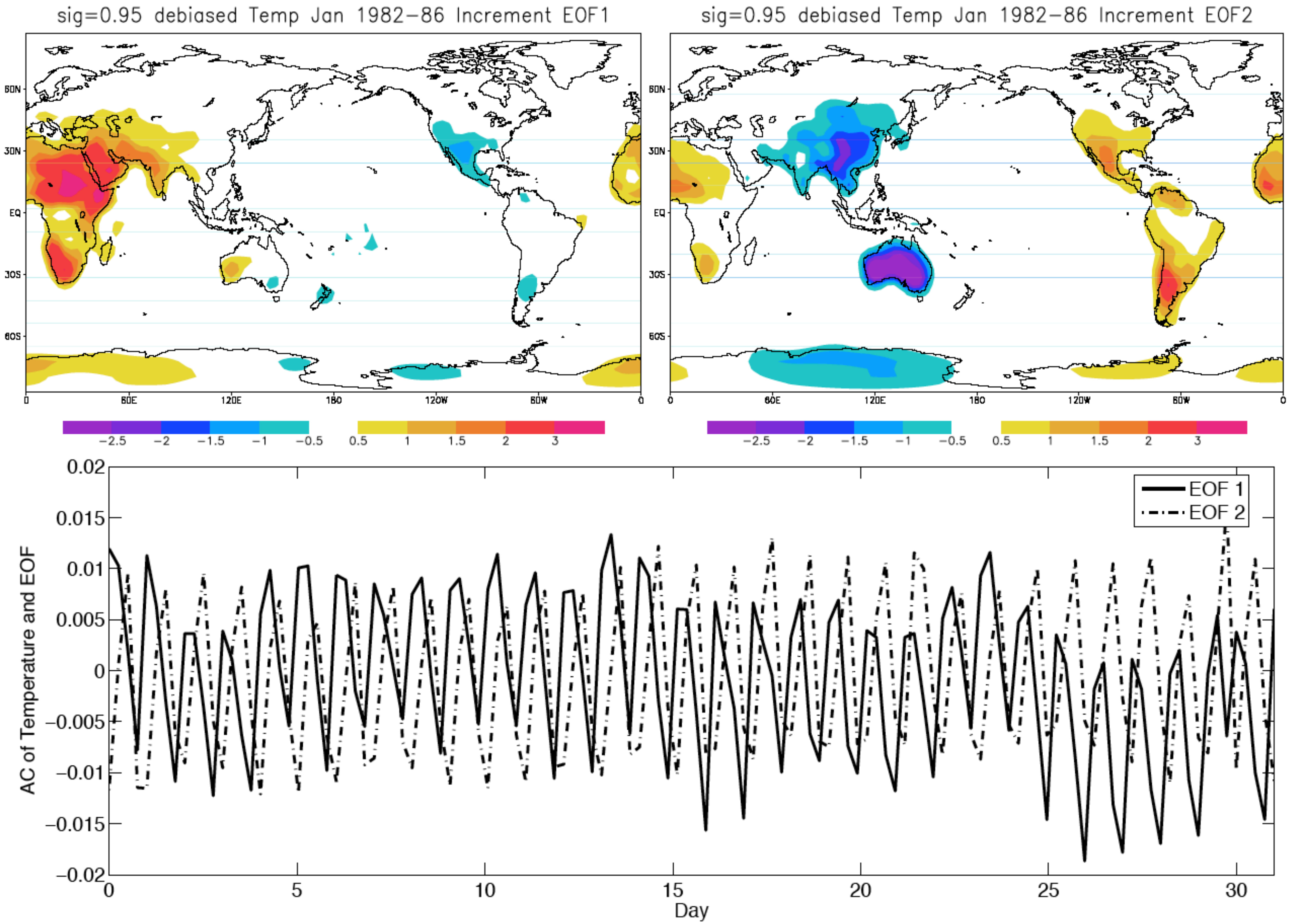


# And the random errors were significantly smaller!




Random errors were reduced by the online correction!

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



# Can we estimate and correct model bias in the GFS?

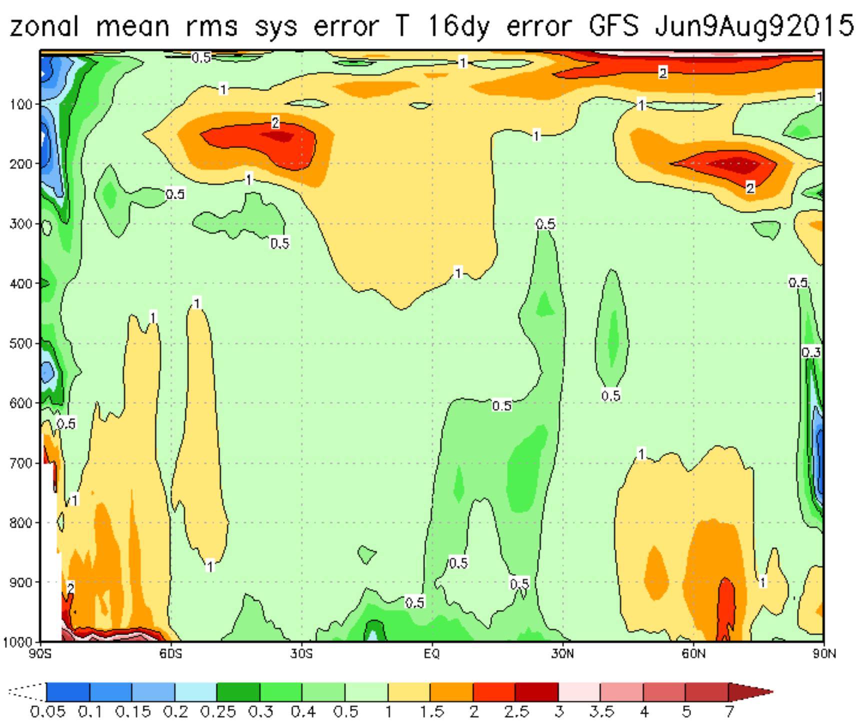
- The systematic errors in the GFS (and all NWP models) are not negligible.
- They are statistically corrected *a posteriori* (offline).
- We aim to correct the GFS (online) adding the average AI/6hr to each forecast variable, like Danforth and Kalnay (2008).
- However, DelSole et al. (2008) attempted this and found it impossible to reduce the random errors.

# Systematic model errors - GFS

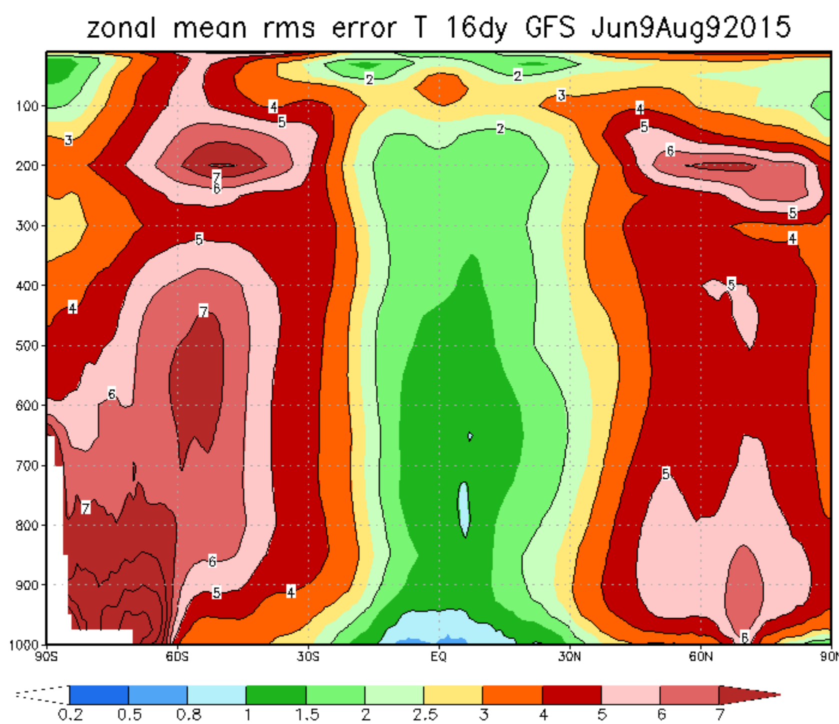
Systematic error range ~1/3 Total error range

after 2 weeks

RMS Systematic errors GFS




RMS Total errors GFS




ΔT(systematic) ~ 0.5 -3K

ΔT(total) ~ 1.5 -9K

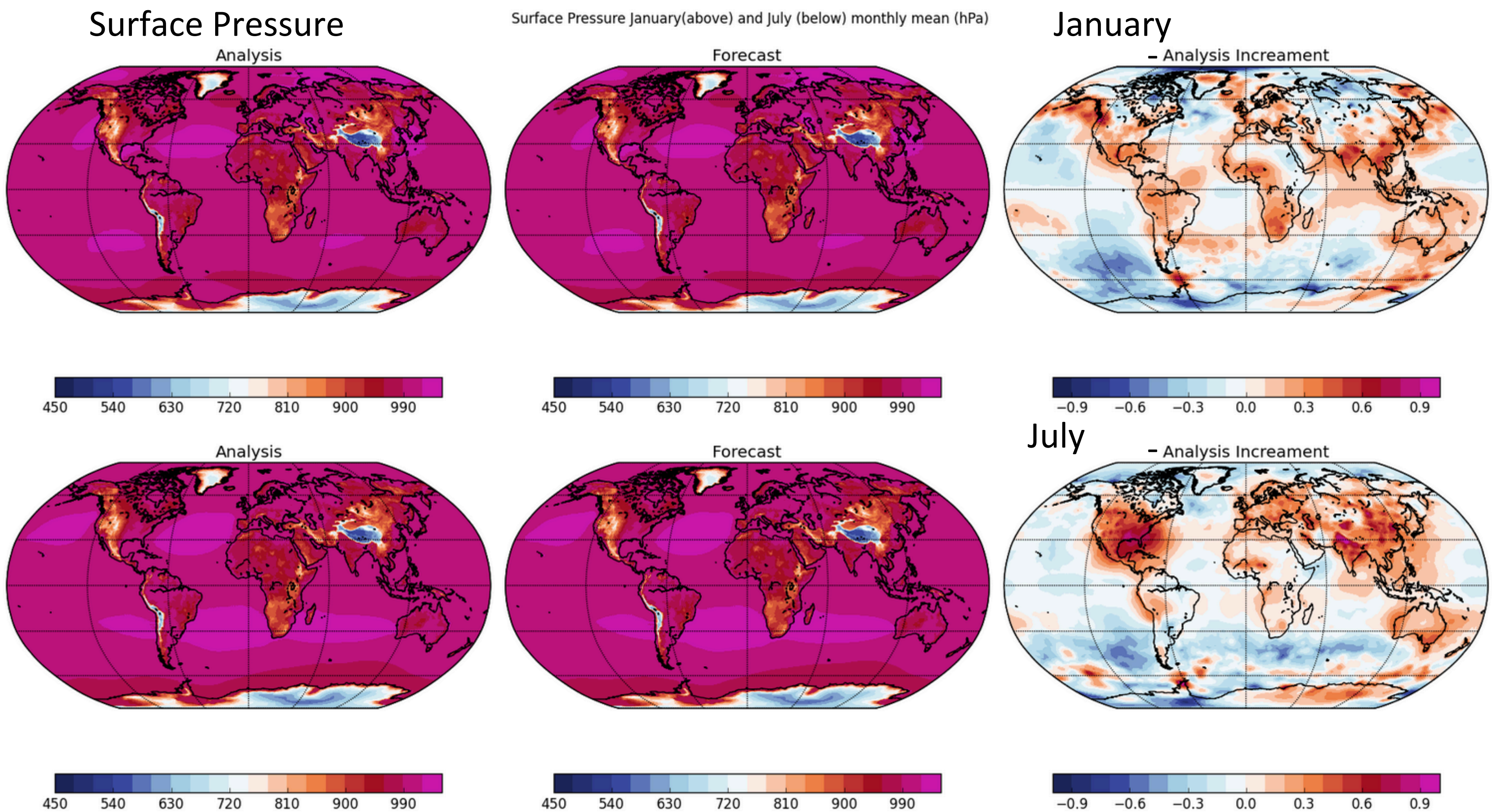
Image courtesy: Glenn White

# Application to GFS

Bhargava, Kalnay, Carton

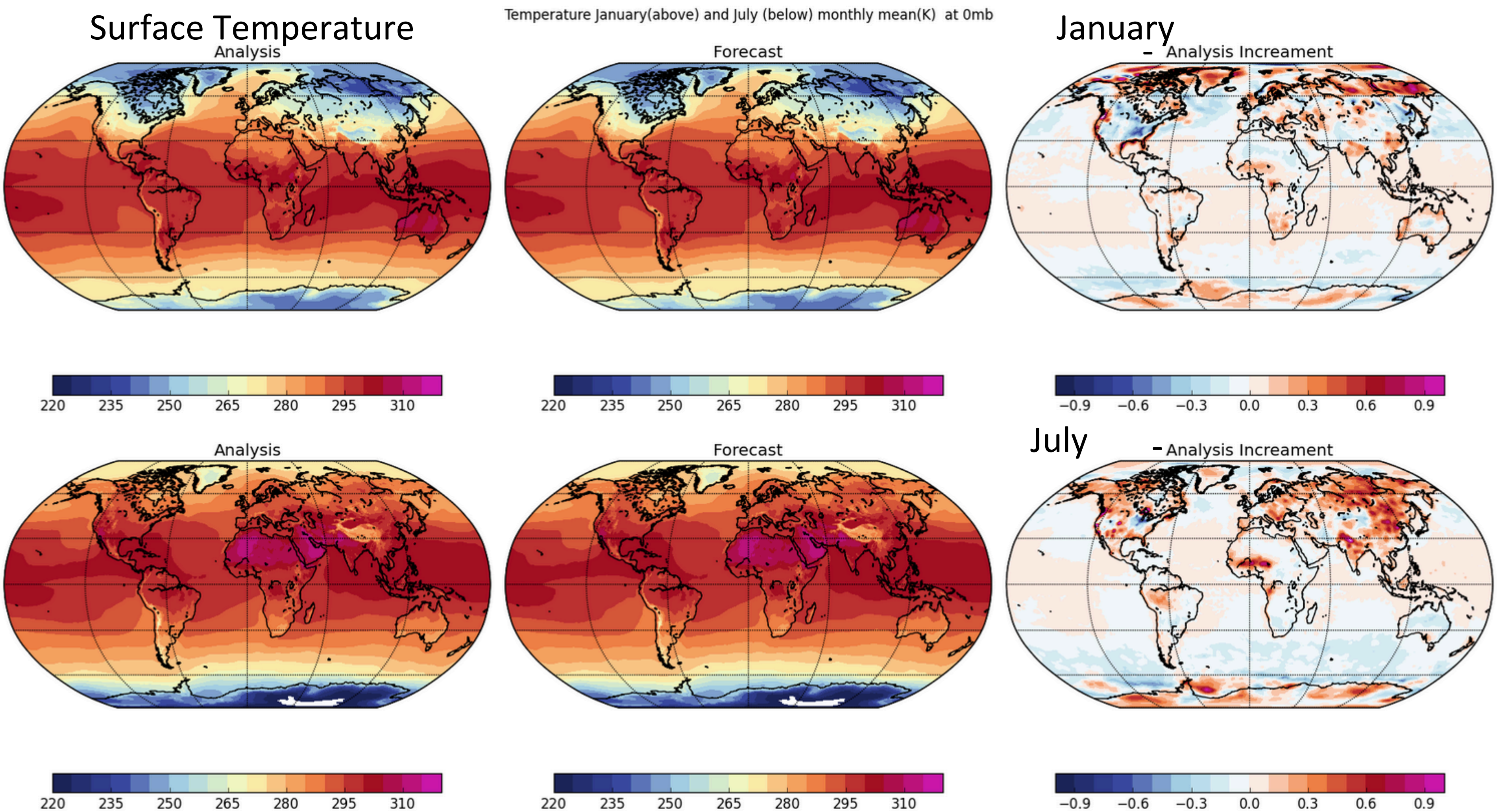
- We obtained T254 6hr forecasts and analyses for 2012, 2013, 2014 from Dr. Fanglin Yang
- We estimate the GFS systematic errors
  - Mean
  - Diurnal
- Check robustness: compare 2012, 2013, 2014
- Explore low dimensional approaches (e.g. diurnal cycle)
- Explore error sensitivity to resolution

# First result: 2014 Analyses, Forecasts and Bias



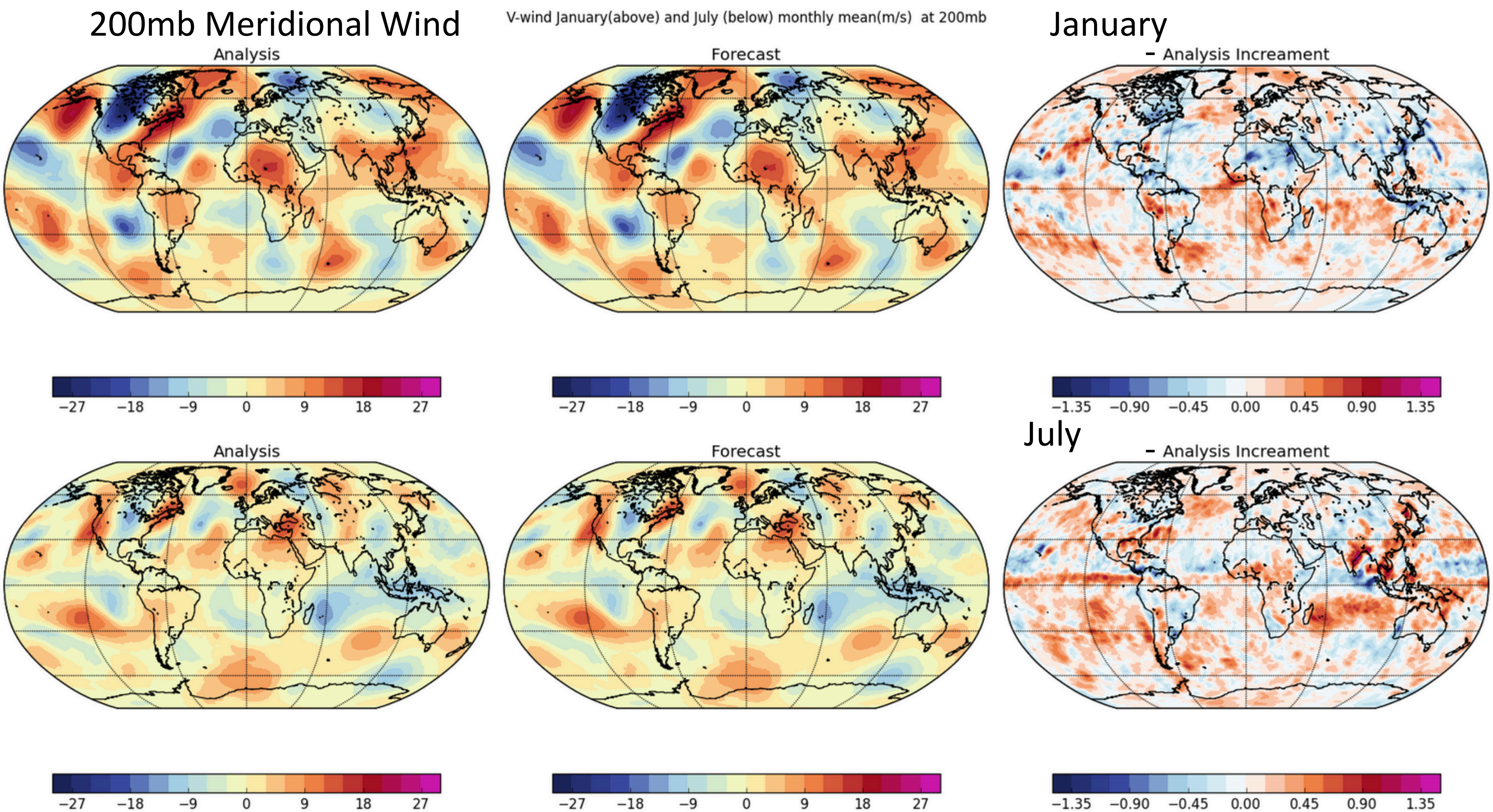

$P_s$ is too high over continents, too low over oceans in both winter and summer.

# First results: 2014 Analyses, Forecasts and Bias



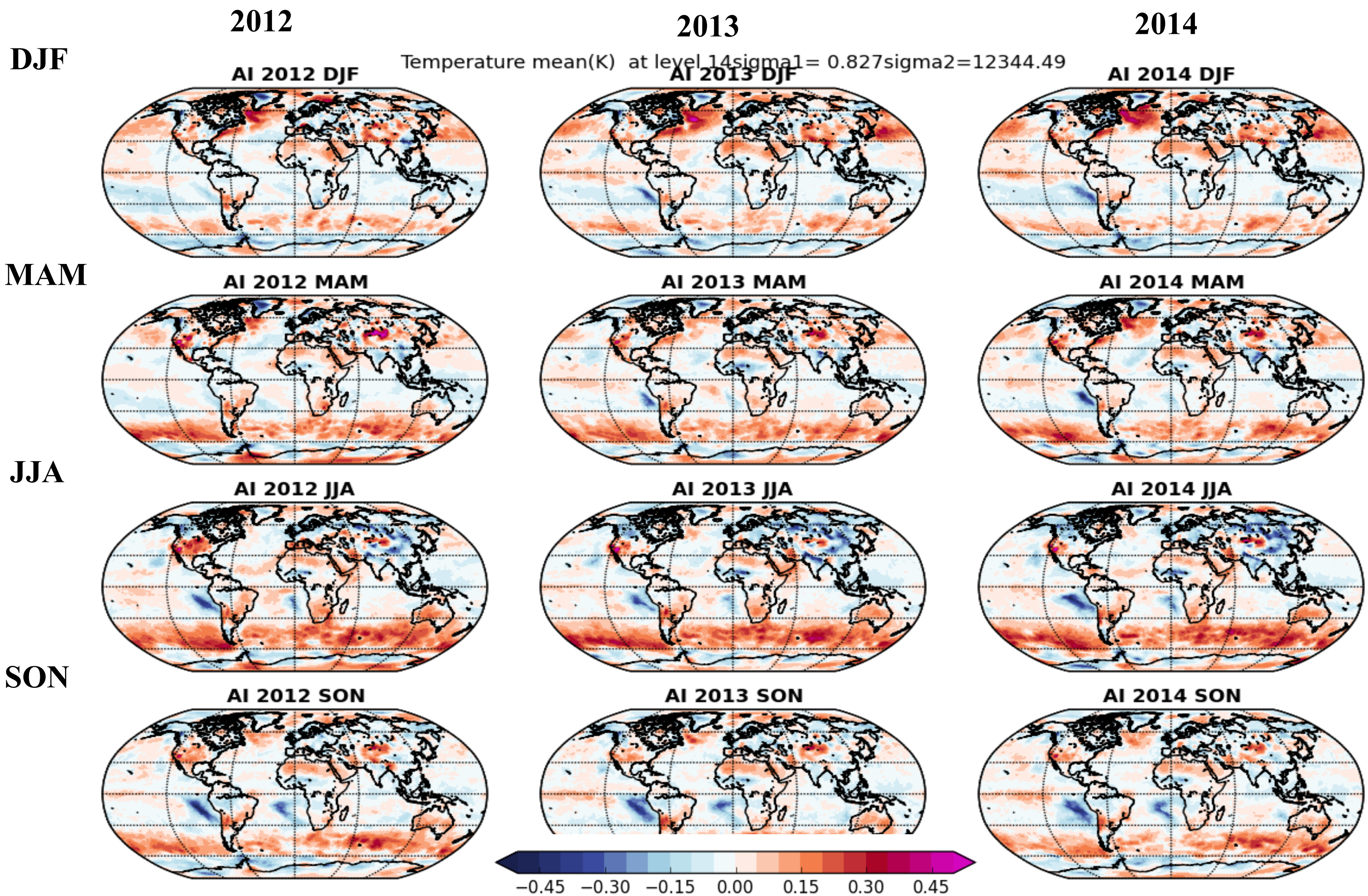

$T_s$ is too high over continents in the summer, too low in the winter.

# First results: 2014 Analyses, Forecasts and Bias

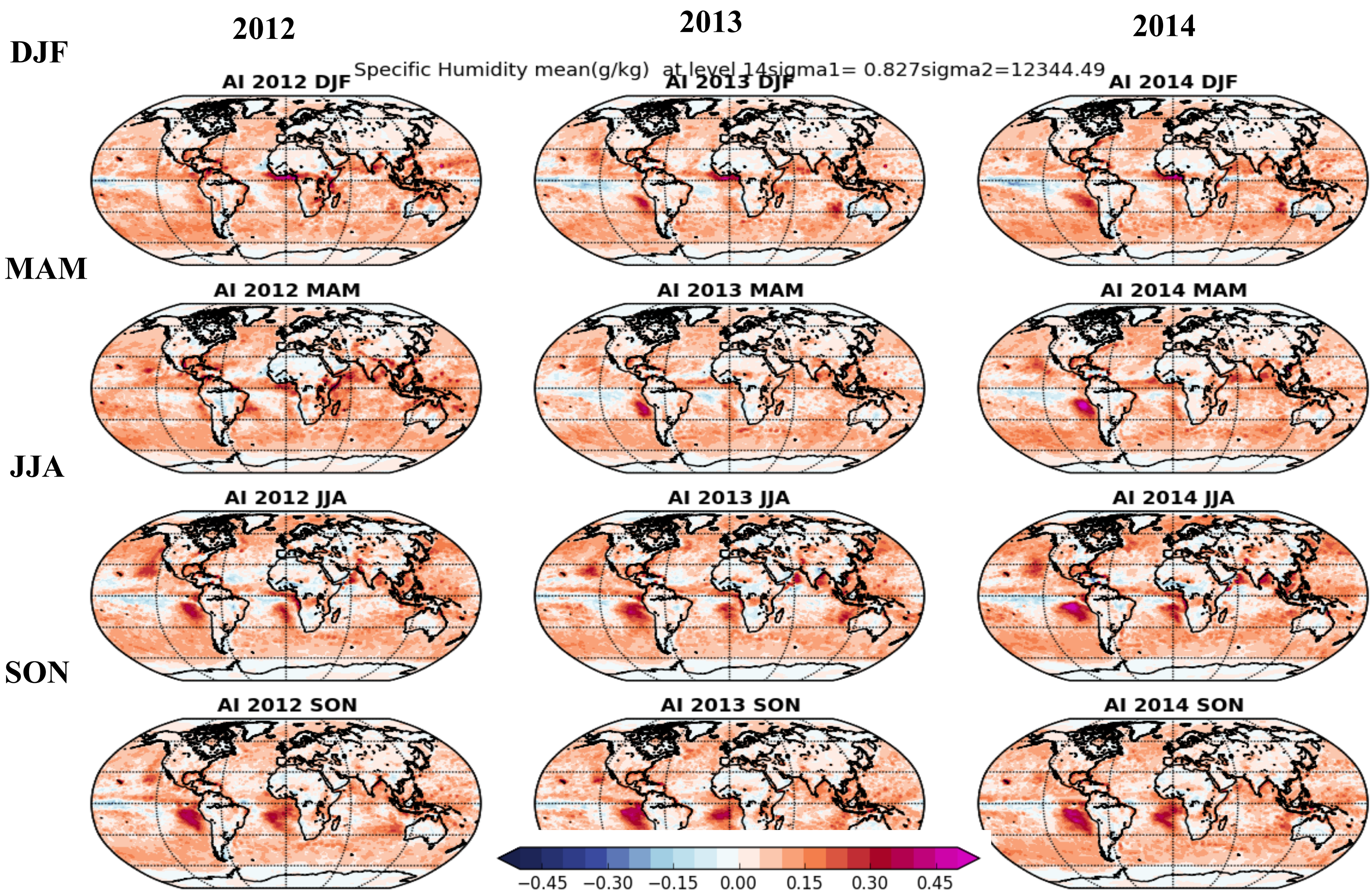



At 200mb the systematic errors have equatorial convergence.

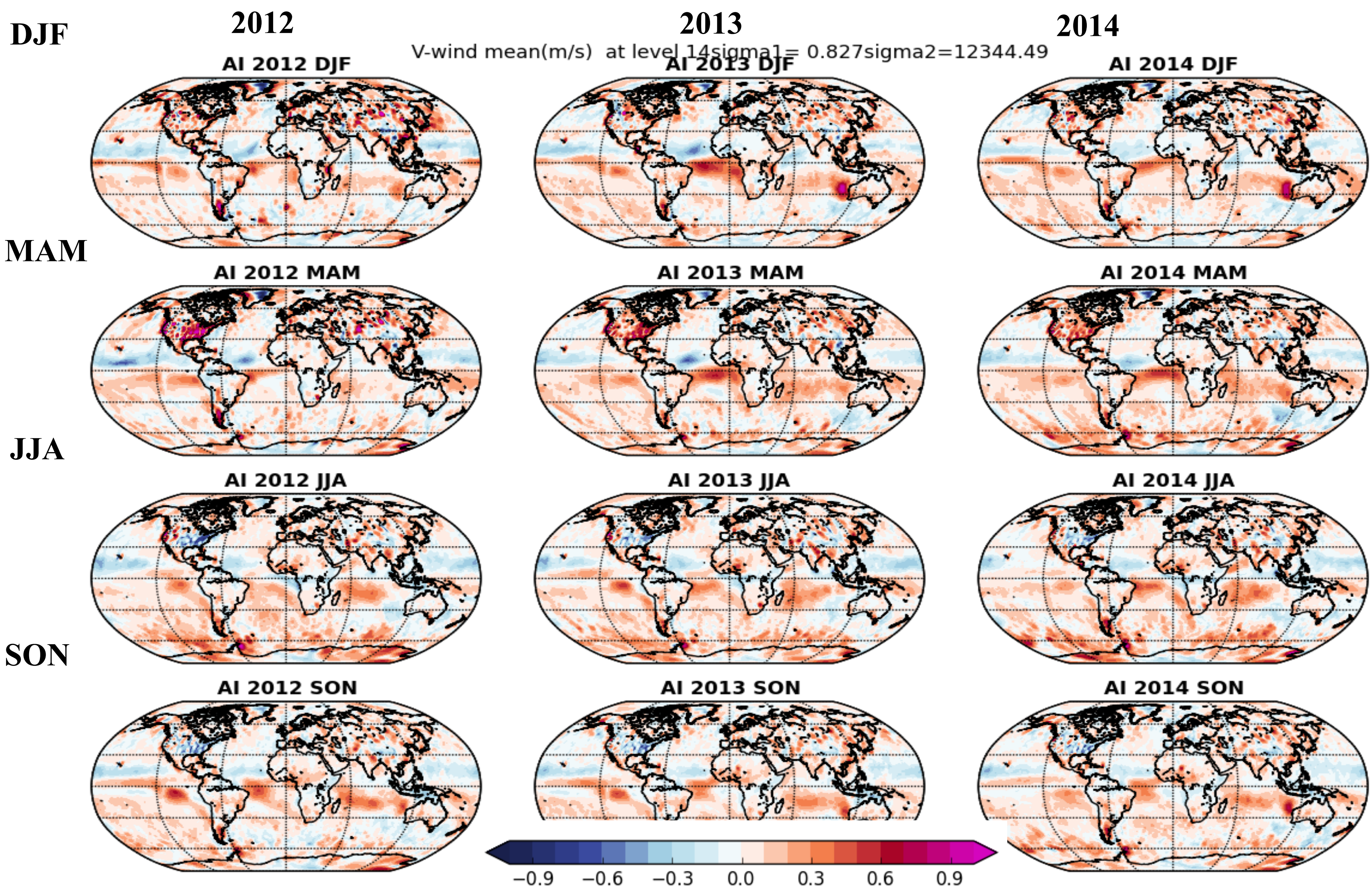
# Seasonal Mean Bias: T (K) at ~850 mb for 2012, 2013, 2014



# Seasonal Mean Bias: Q (g/kg) ~850 mb



# Seasonal Mean Bias: V (m/s) at ~850 mb



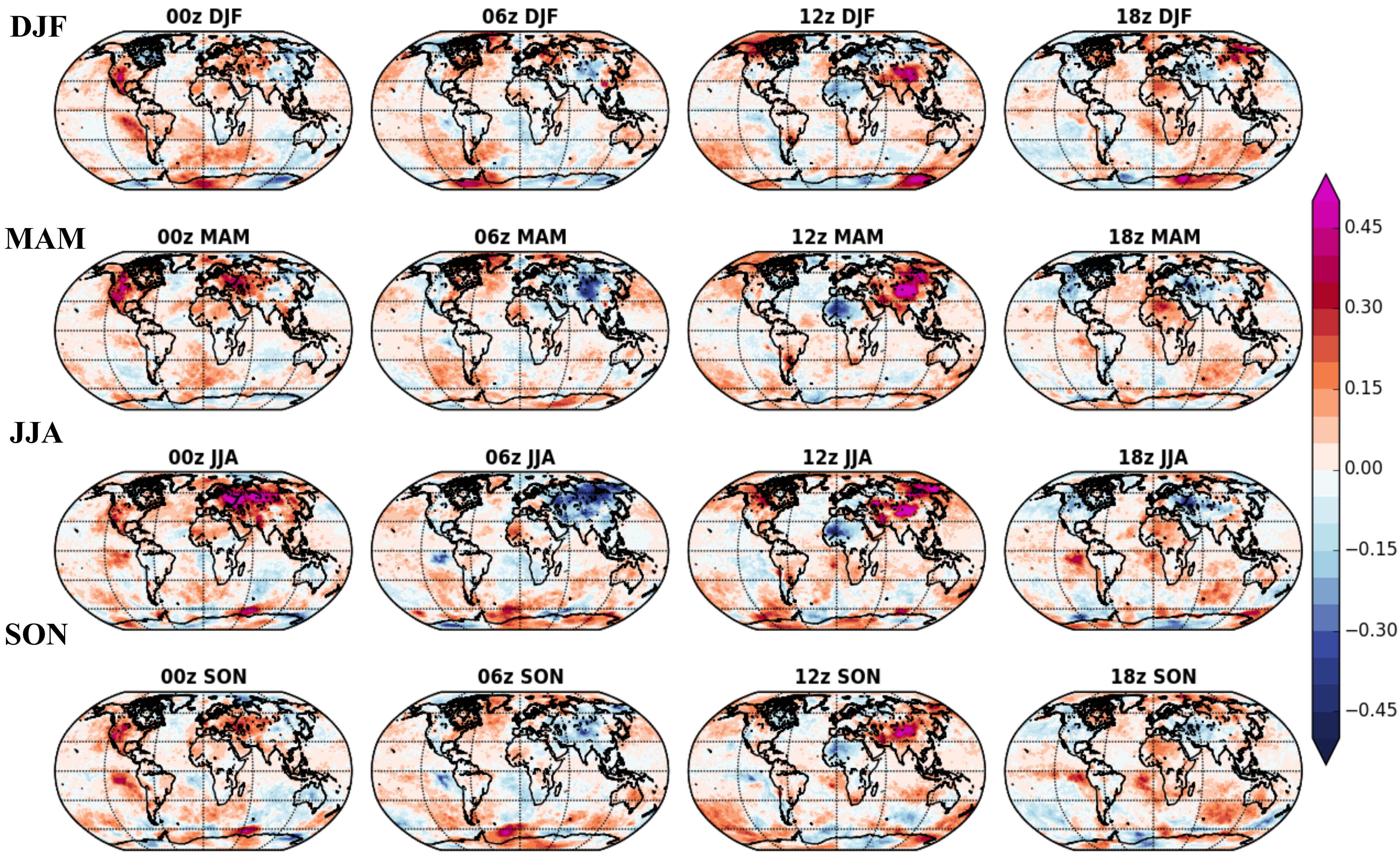
# Findings

- 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

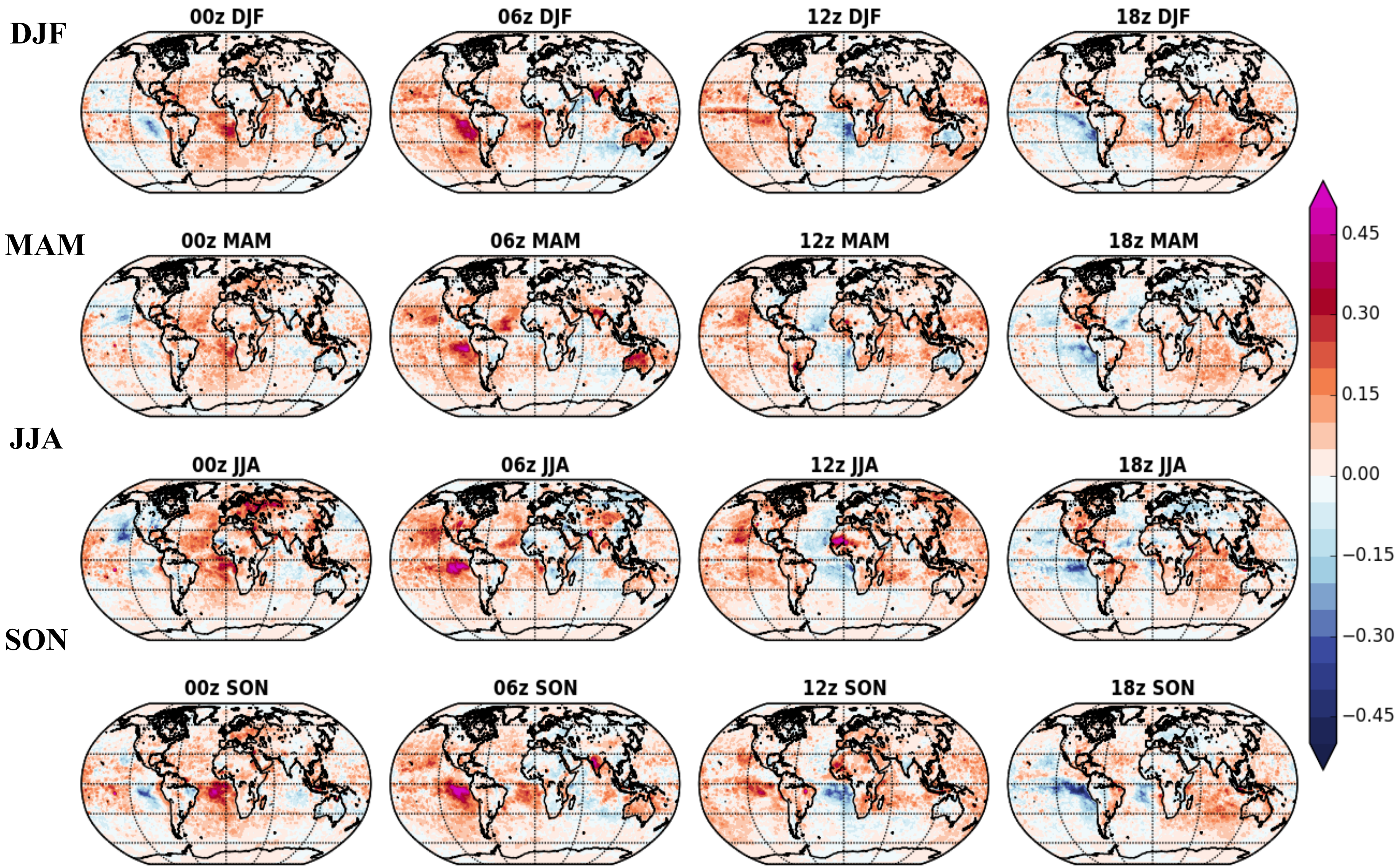
# Diurnal cycle error estimation

- 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

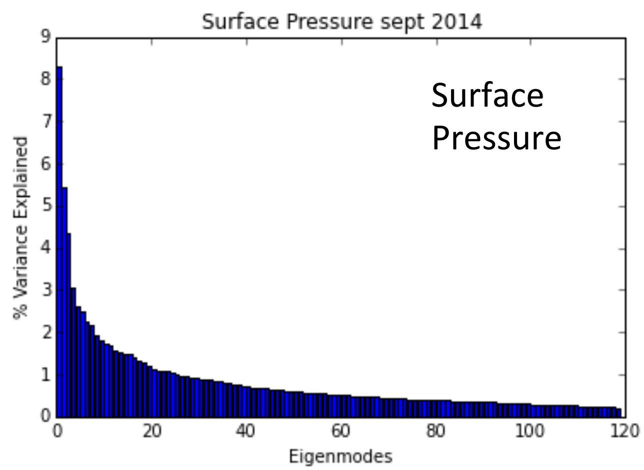
# Mean diurnal cycle error: T (K) Sept '14 at ~850mb

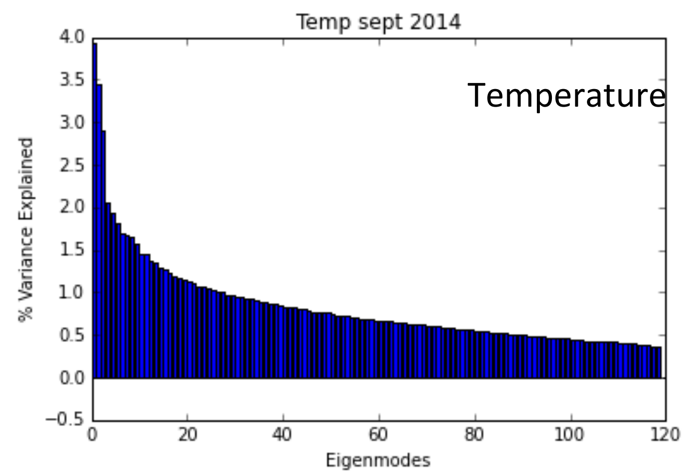


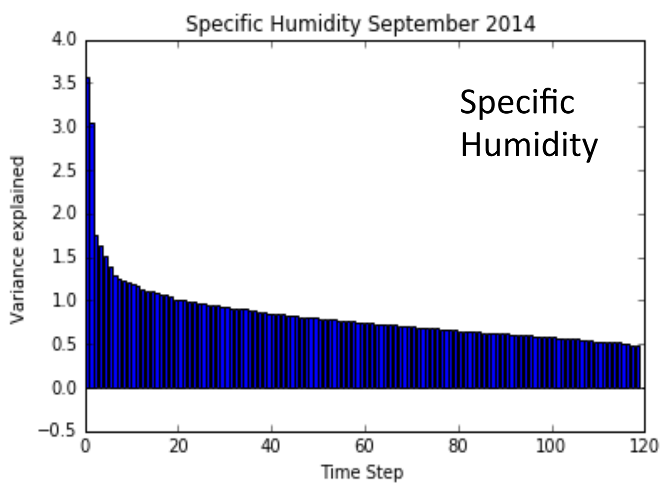
# Mean diurnal cycle error: Q (g/kg) Sept ’14 at ~850mb



# Variance Explained by Eigenmodes




**Variance explained by first 4 modes**

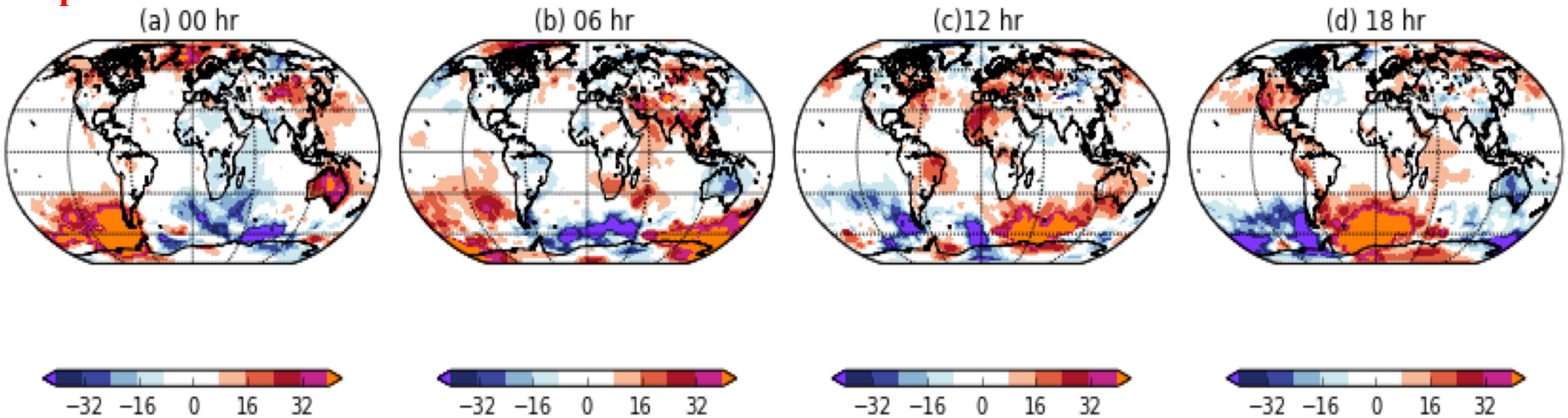
- Ps- 24%
- T- 11%
- Q- 10%

We check whether leading 4 modes capture the errors in the diurnal cycle. The rest of the modes explain errors due to other sources.

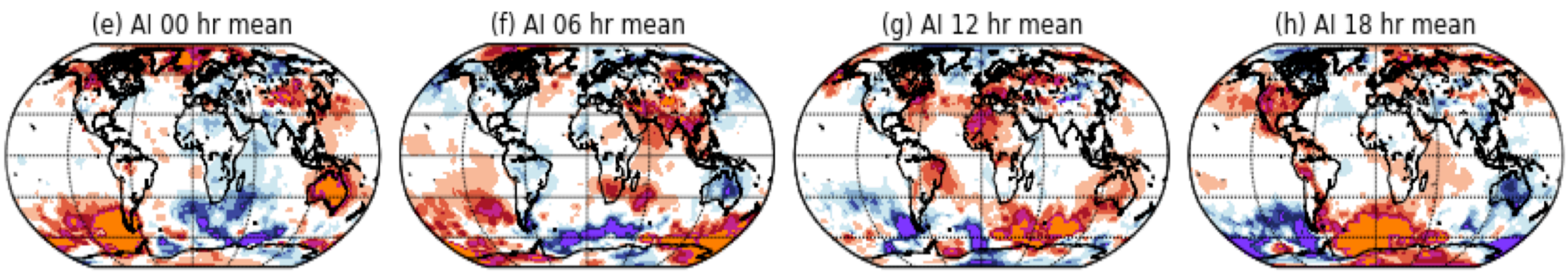
# First four vs 120 modes: $P_s$ (mb) Sept'14

**First 4 modes capture the diurnal cycle errors almost perfectly**

**Top: 4 modes**



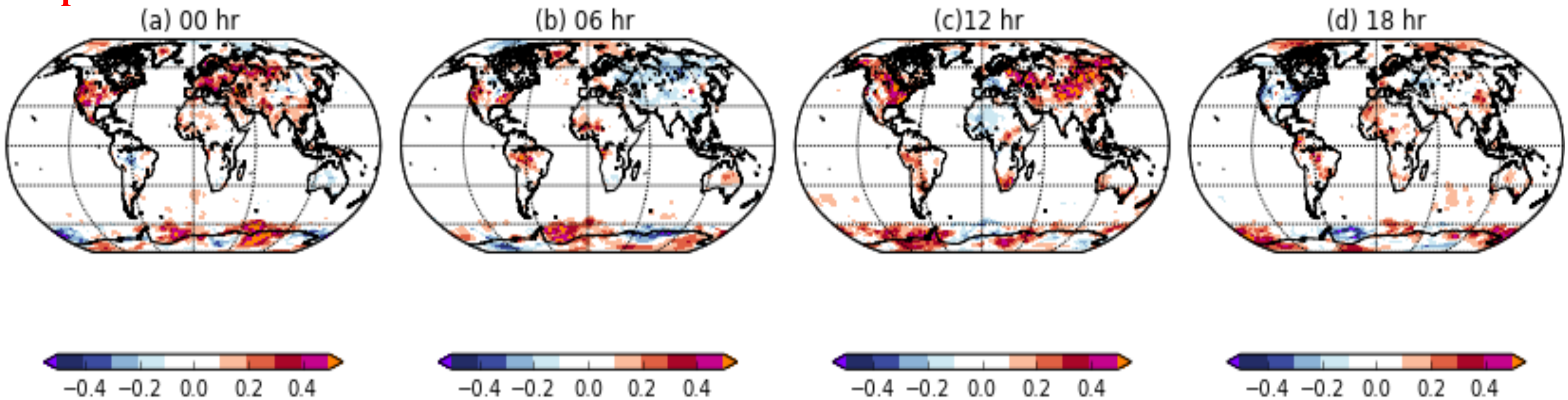

**Bottom: 120 modes**



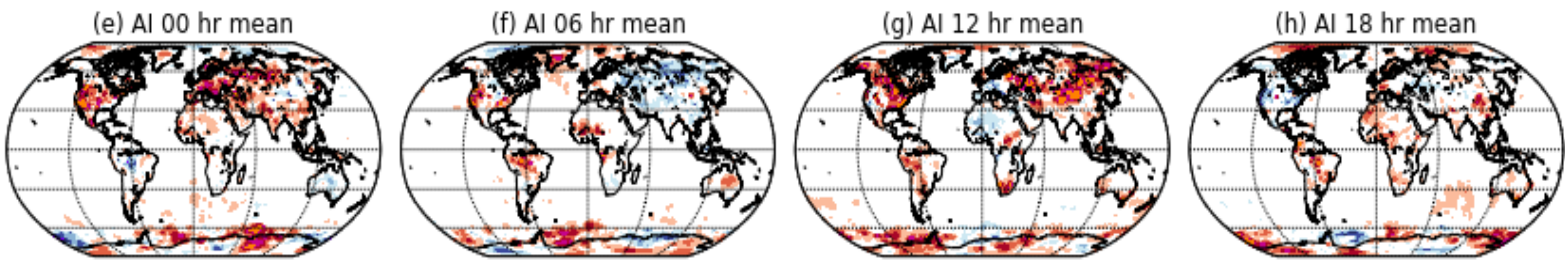
# First four vs 120 modes: T(K) Sept’14

## First 4 modes capture the diurnal cycle errors almost perfectly

Top: 4 modes



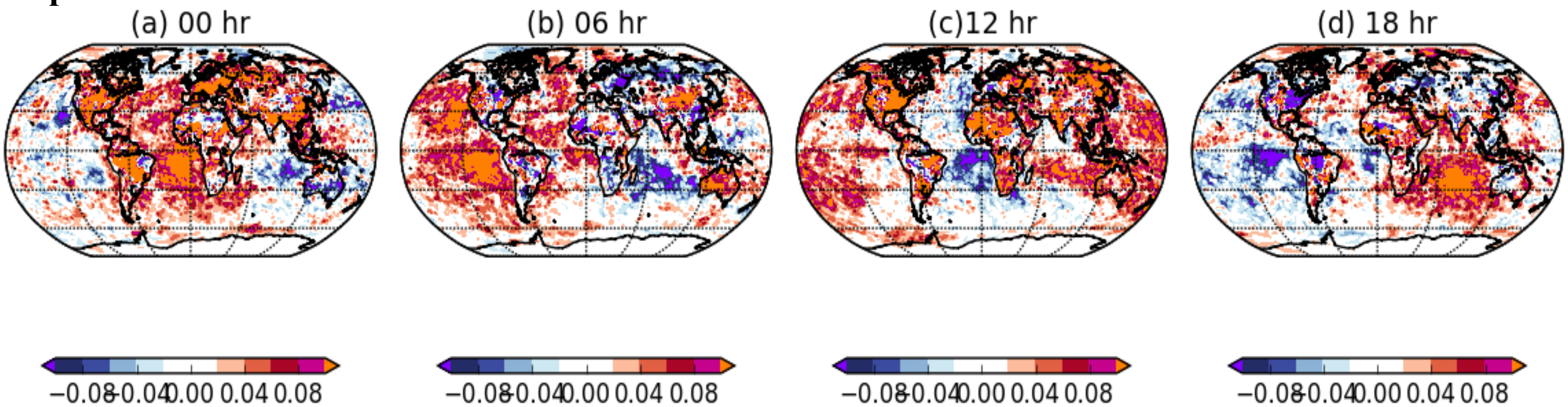

Bottom: 120 modes



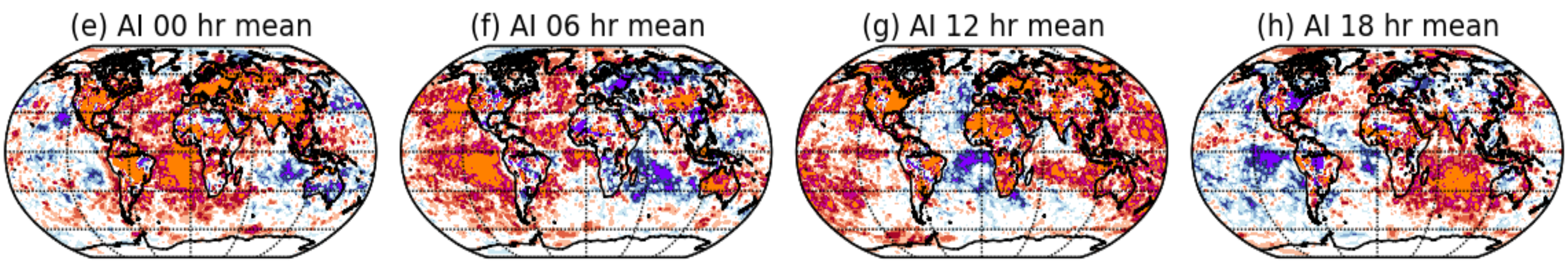
# First four vs 120 modes: Q (g/kg) Sept'14

**First 4 modes capture the diurnal cycle errors almost perfectly**

**Top: 4 modes**



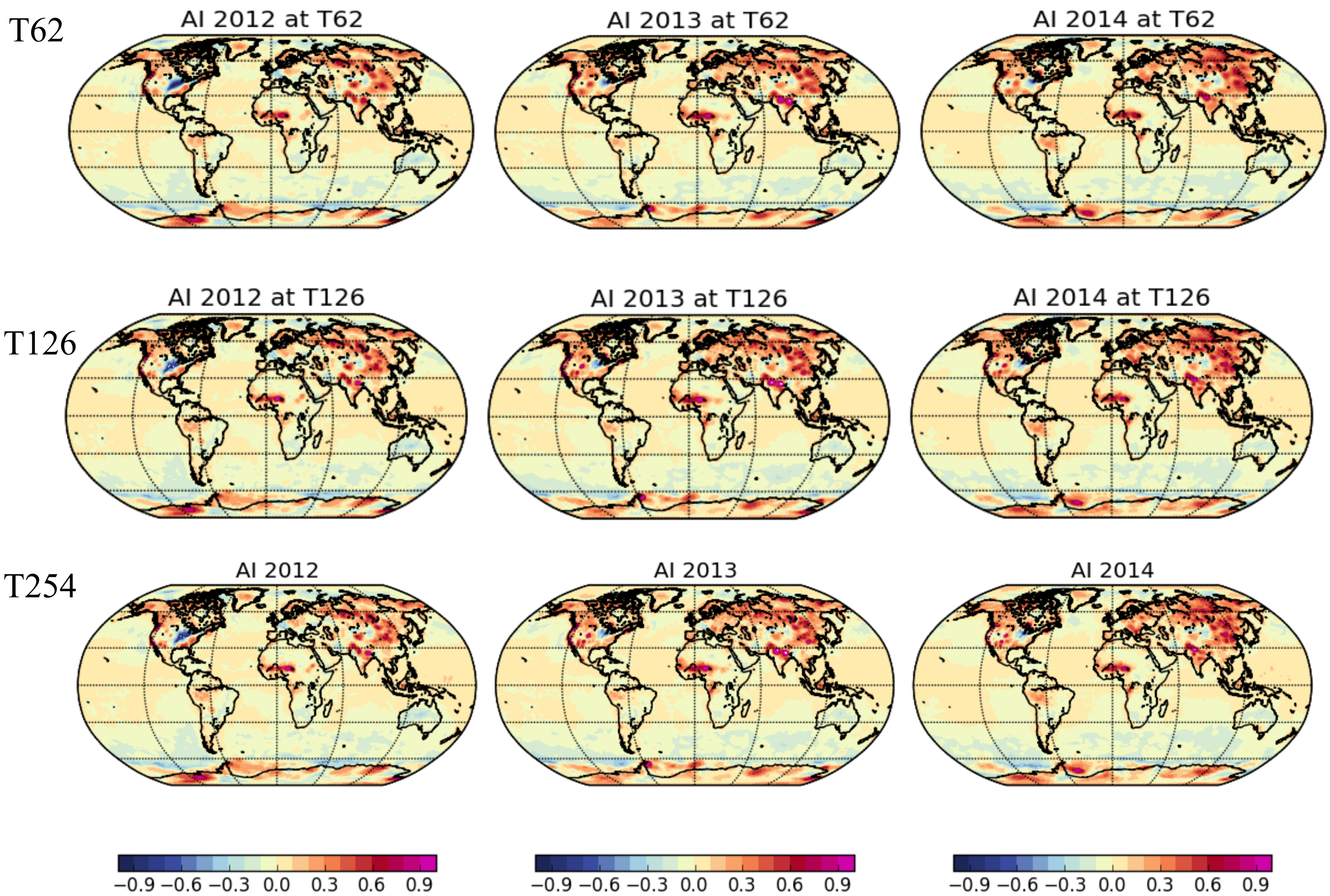

**Bottom: 120 modes**



# 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




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

# 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

# Proposed plans for GFS correction in collaboration with EMC

- Apply online corrections to GFS
- Examine improvements in bias and random error
- Compare online correction results with standard operational statistical bias correction
- Check the impact of correcting diurnal cycle

- Use ensemble members as a testbed for corrections
- Work with the EMC scientists on how to facilitate testing impacts of new parameterizations

- Work with EMC scientists on R2O implementation

# Carton et al. are working on similar approaches for SODA 3.4 (ERA Interim fluxes) with great results

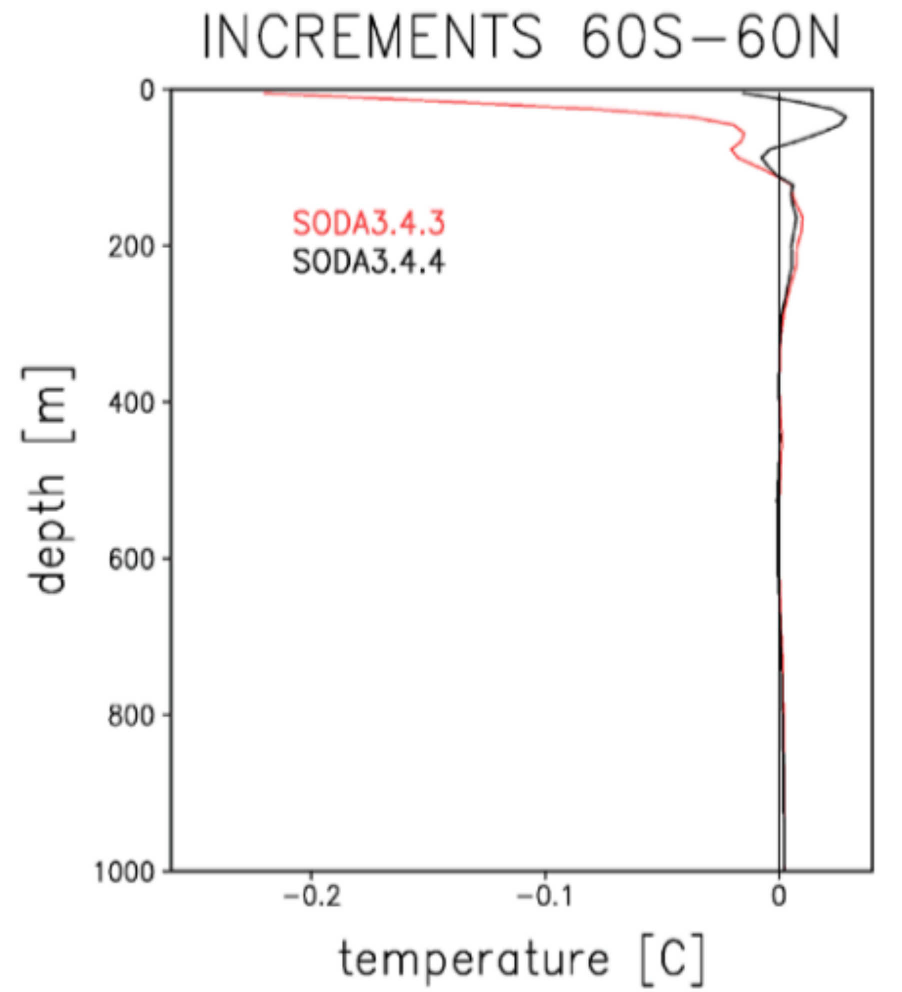



**Figure 1** spatial and time average of the increments for SODA3.4.3 and SODA3.4.4. The bias correction is knocking town the increments to a much  more reasonable 0.01C level.

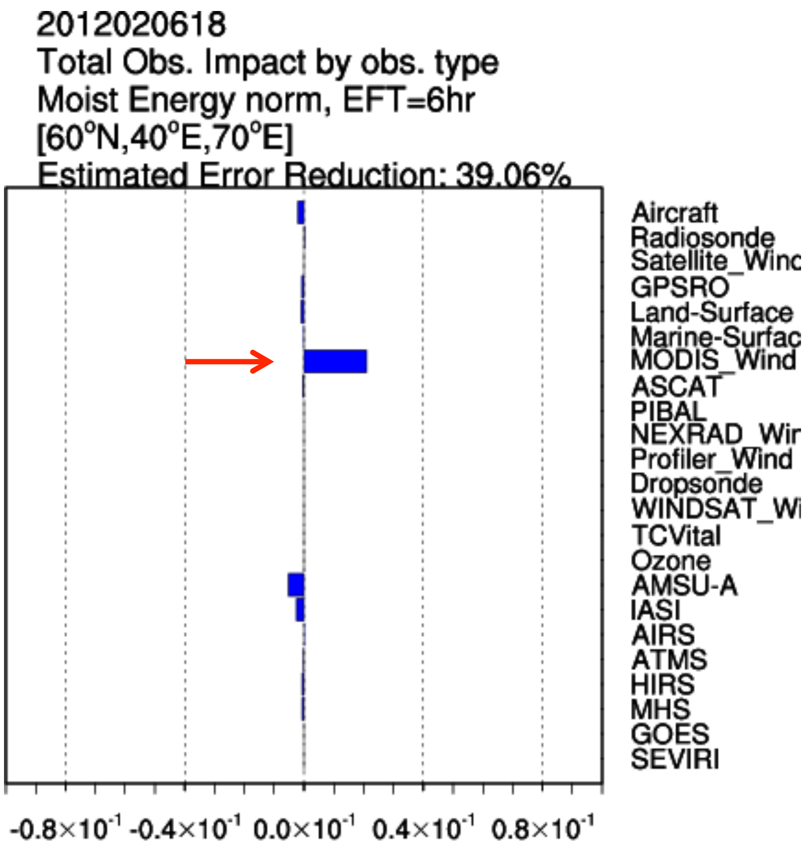
# 2) Improve the observations: Ensemble Forecast Sensitivity to Observations and Proactive QC

- Kalnay et al. (2012) derived EFSO.
- Ota et al. (2013) tested 24hr GFS forecasts and showed EFSO could be used to identify detrimental obs.
- D. Hotta (2014): **EFSO can be used after only 6 hours**, so that the detrimental obs. can be withdrawn and collected with useful metadata to be improved. The analysis is corrected with EFSO (see Chen's poster).
- 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: **GOOD RESULTS!**

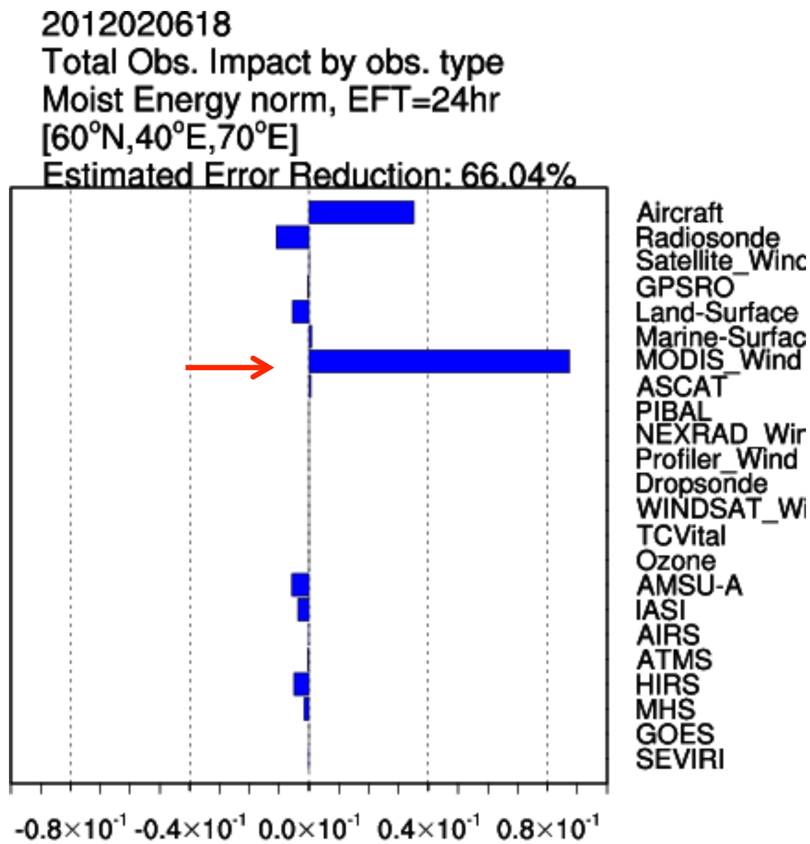
# Hotta (2014)

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

**FT=06 hr.**




**FT=24 hr.**

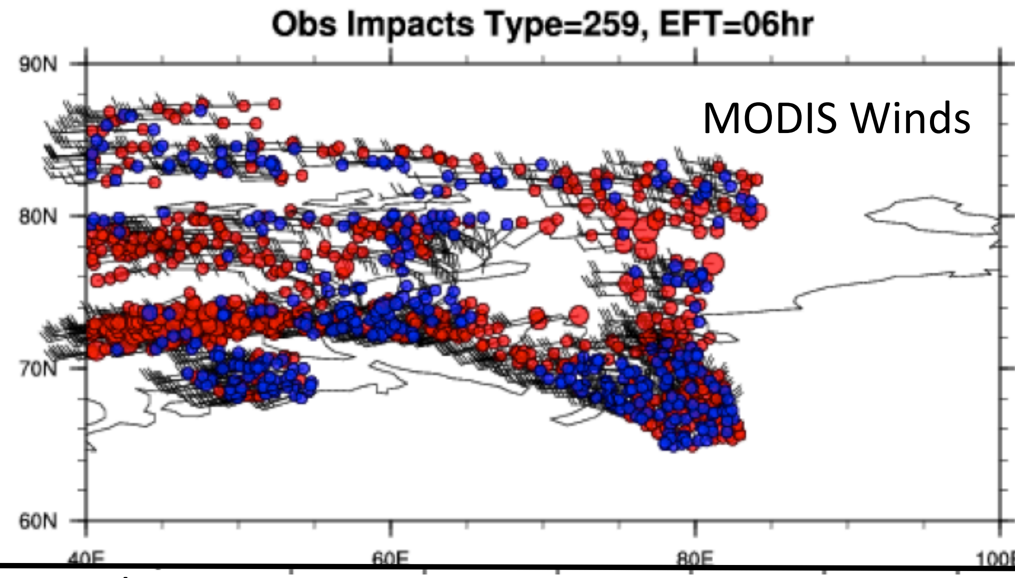



Can identify the bad observations after only 6 hours!

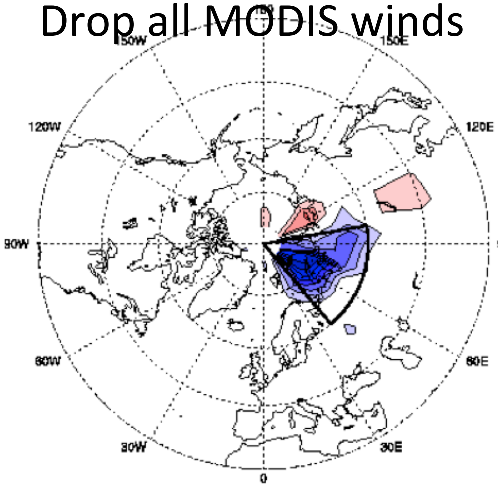
# Improve observations:

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

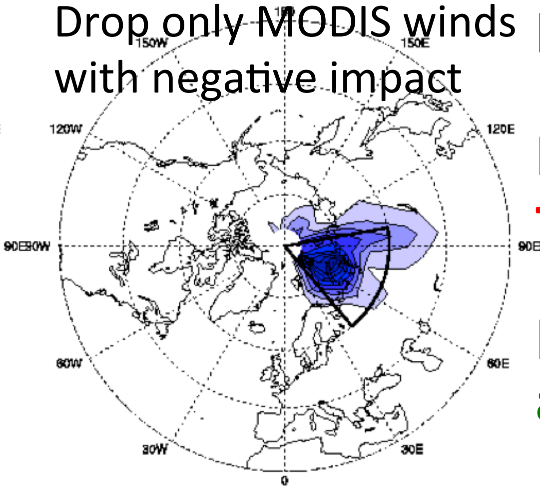
**Dr. Daisuke Hotta (2014):**
**EFSO is able to find whether each observation improves (blue) or makes the 6hr forecast worse (red)**




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!**

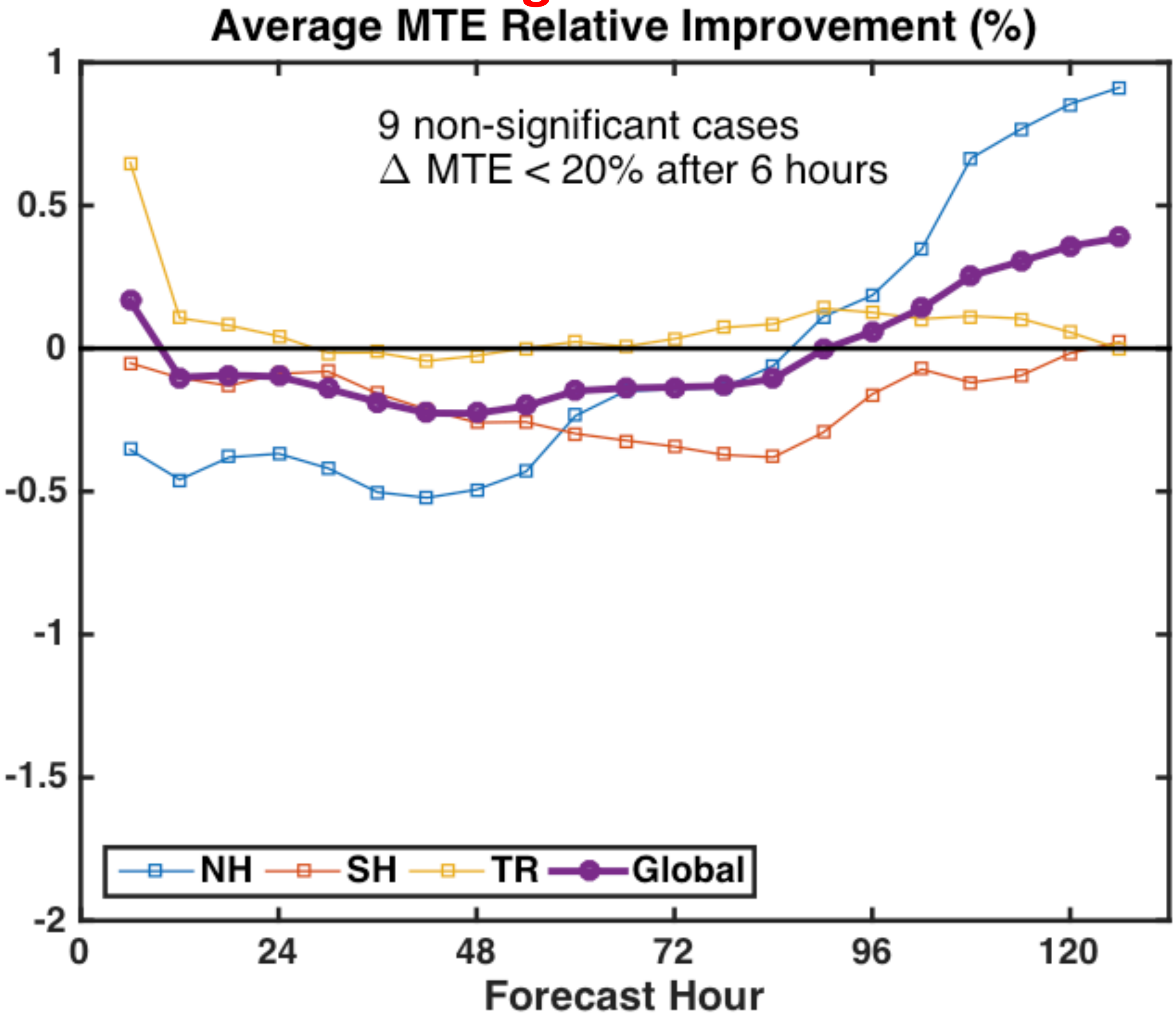
**It should accelerate optimal assimilation of new instruments!**

# Current testing of PQC on JCSDA S4
# Tse-Chun Chen

- Before operational testing at NCEP, we need to show that:
  - Denying detrimental observations improves the forecasts.
  - Denying detrimental observations works in a cycled way (we tested case by case so far).
  - The EFSO approximation (constant K) can be used to replace the full analysis without the flawed observations (much faster).
  - We can use the 6hr early forecast to check the final analysis.
- Prof. Daryl Kleist has kindly offered to help test PQC operationally once we have good results.
- So, let's look at the results: We tested 20 cases of withdrawing flawed observations, re-computing the analysis and performing 5-day forecasts with and without the flawed observations.

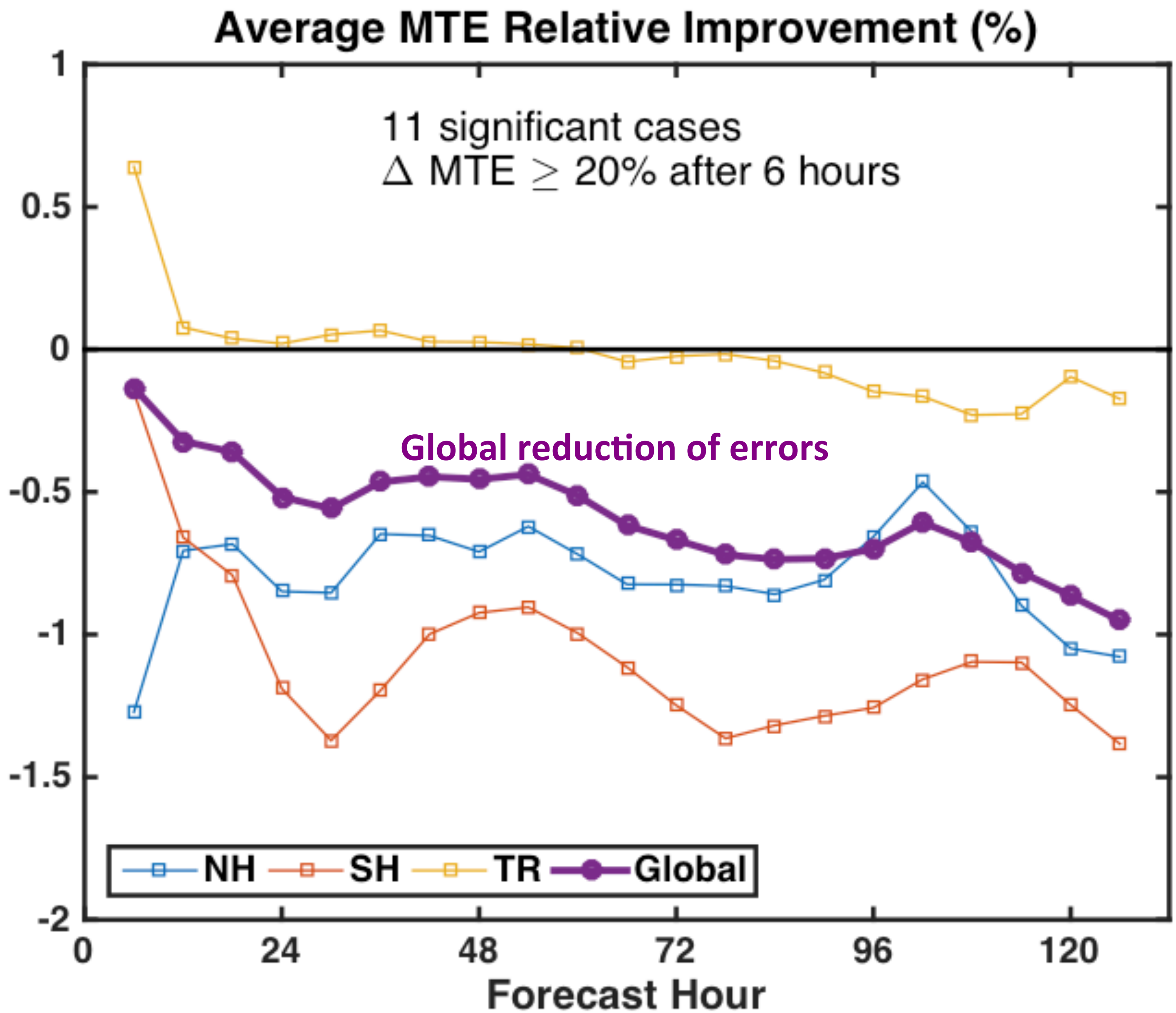
# 5-day reduction of Moist Total Energy of the forecast error

## 9 non-significant cases

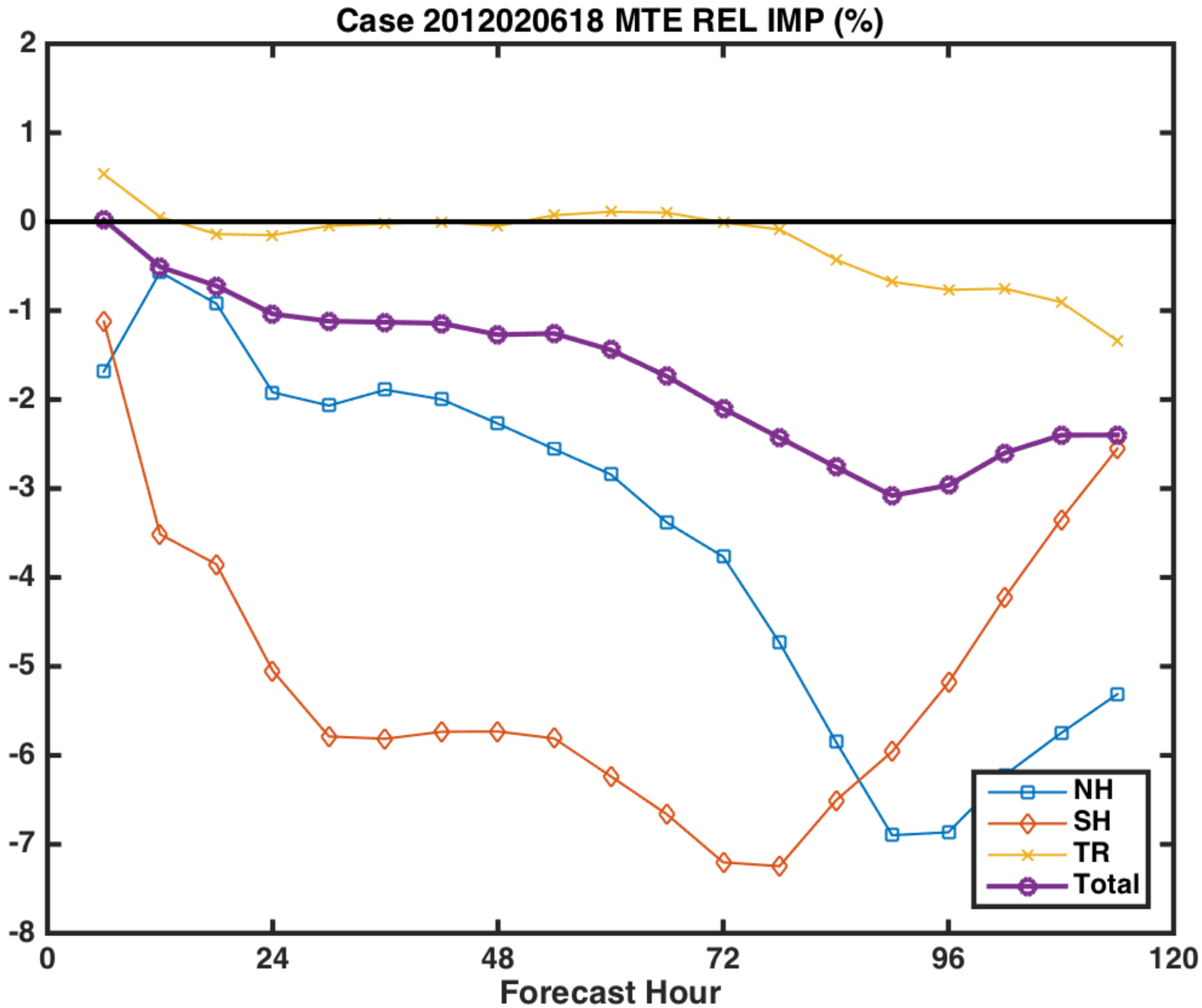


# 5-day reduction of Moist Total Energy of the forecast error

## 11 significant cases



# Results: we measure the % change in forecast error (Moist Total Energy) when withdrawing flawed obs.




**<u>BEST CASE</u>**

Detrimental obs in both NH and SH

**7% reduction of error in both NH and SH! ("Skill-dropouts")**

Analysis changes increase initial errors in the tropics, but **the tropics also improve with time.**

# SUMMARY

- Future applications of EnKF-based data assimilation for improving both weather and climate prediction
  - 0) Combine model forecast and observations to create the best initial conditions ✓
  - 1) Do strongly coupled data assimilation ✓
  - 2) Improve observations with EFSO and PQC ✓
  - 3) Improve models by using the Analysis Increments to correct the models' bias ✓
  - 4) Assimilate non-Gaussian observations (e.g., precipitation, ice cover, clouds) with a Gaussian Transformation ✓

**THANKS!**