Observing System Simulation Experiments for Fun and Profit

Nikki Privé

1 June 2015

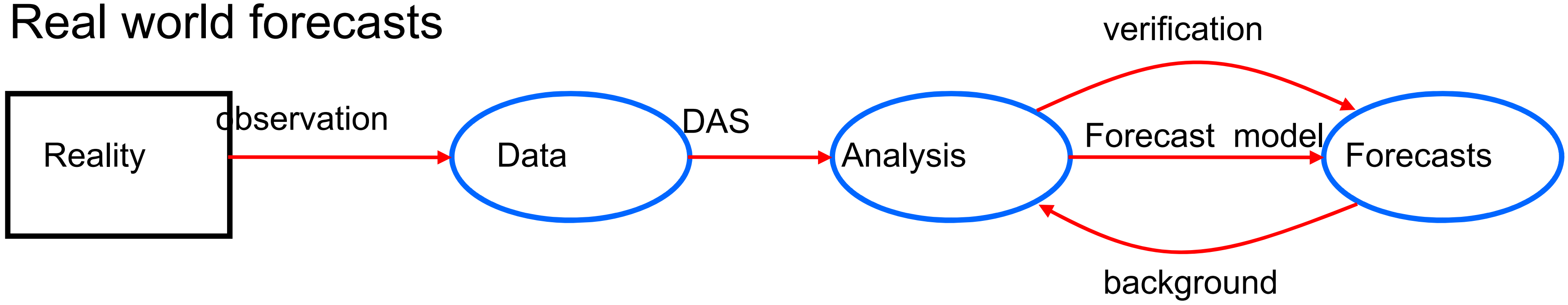
What is an OSSE?

An OSSE is a modeling experiment used to evaluate the impact of new observing systems on operational forecasts when actual observational data is not available.

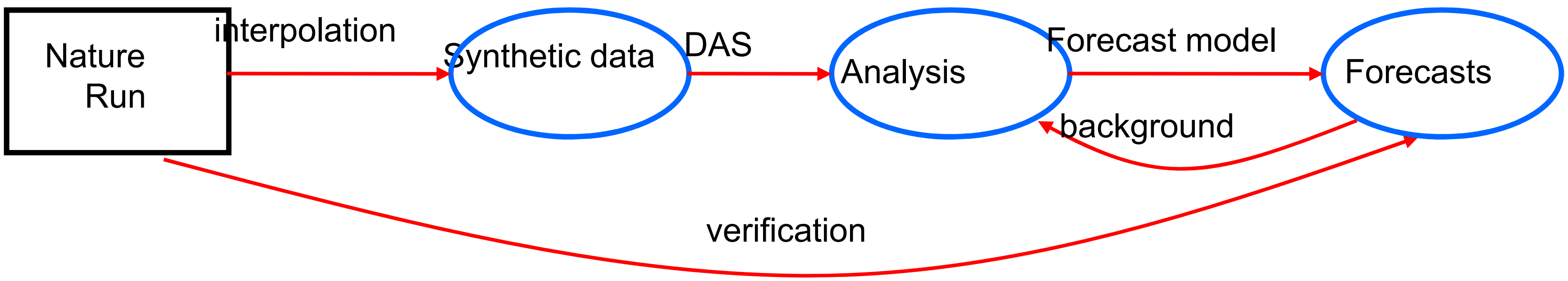
- A long free model run is used as the “truth” - the Nature Run
- The Nature Run fields are used to back out “synthetic observations” from all current and new observing systems.
- The synthetic observations are assimilated into a different operational model
- Forecasts are made with the second model and compared with the Nature Run to quantify improvements due to the new observing system

OSSEs vs. the Real World

Real world forecasts



OSSE forecasts



Why do an OSSE?

1. You want to find out if a new observing system will add value to NWP analyses and forecasts

2. You want to make design decisions for a new observing system

3. You want to investigate the behavior of data assimilation systems in an environment where the truth is known

An OSSE will not....

1. ....show miraculous forecast improvements from new observations

2. ....necessarily show any forecast improvements from new observations

3. ....tell you what will happen in the real world

Nature Runs

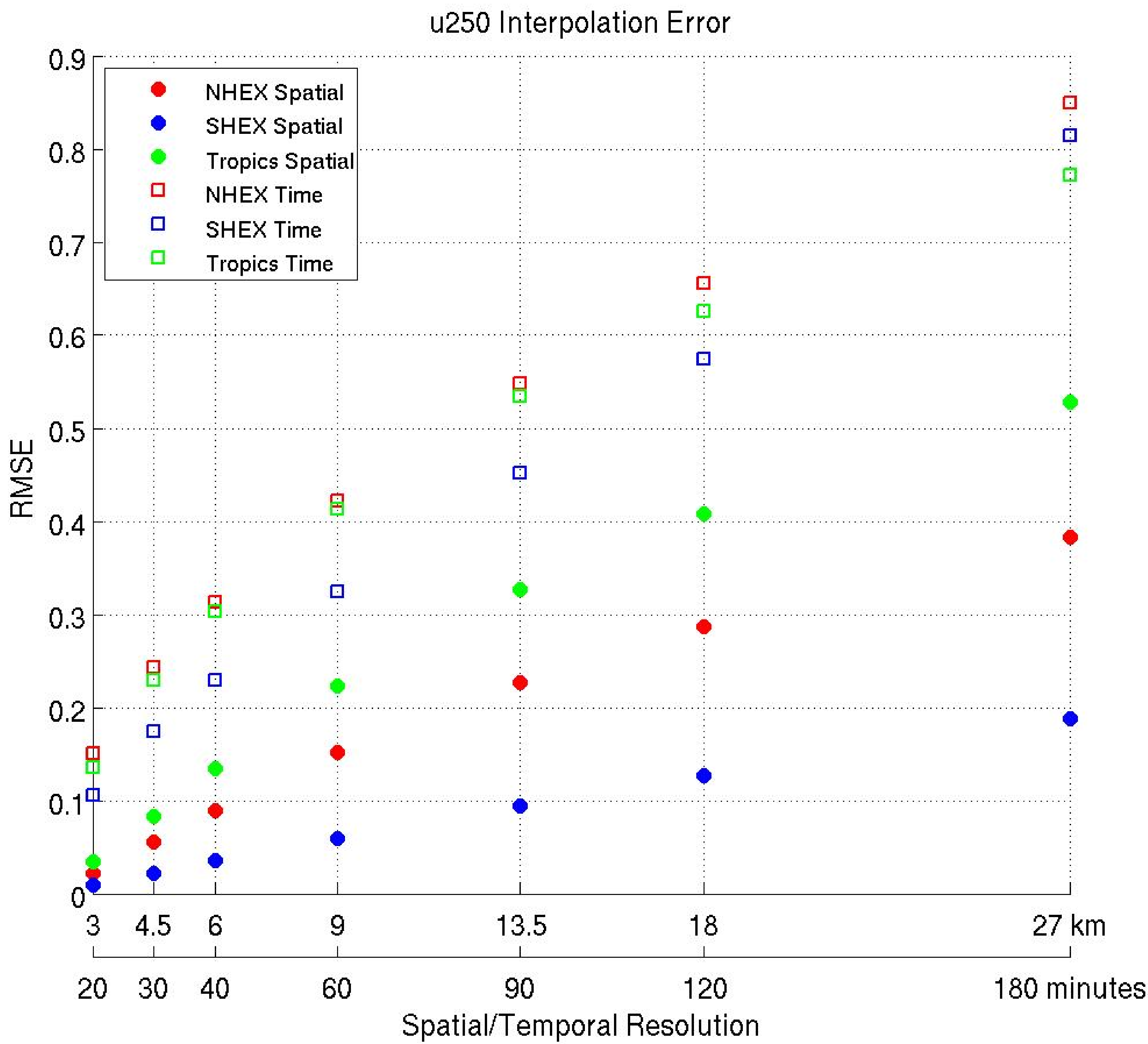
- Nature Runs act as the 'truth' in the OSSE, replacing the real atmosphere.
- Usually, a long free (non-cycling) forecast from the best available model is used as the NR
    - Model forecast has continuity of fields in time
    - Sometimes an analysis or reanalysis sequence is used, but the sequence of states of truth can never be replicated by a model
- Always a push for bigger, higher resolution NR

Nature Run Requirements

- Must be able to realistically model phenomena of interest
  - Dynamics and physics should be realistic
  - Must produce fields needed for “observations”
  - Should be verified against real world
- Ideally is ‘better’ than the operational model to be used for experiments
- Preferably a different model base is used for the NR and the experimental forecast model to reduce incestuousness

Common Problems with Nature Runs

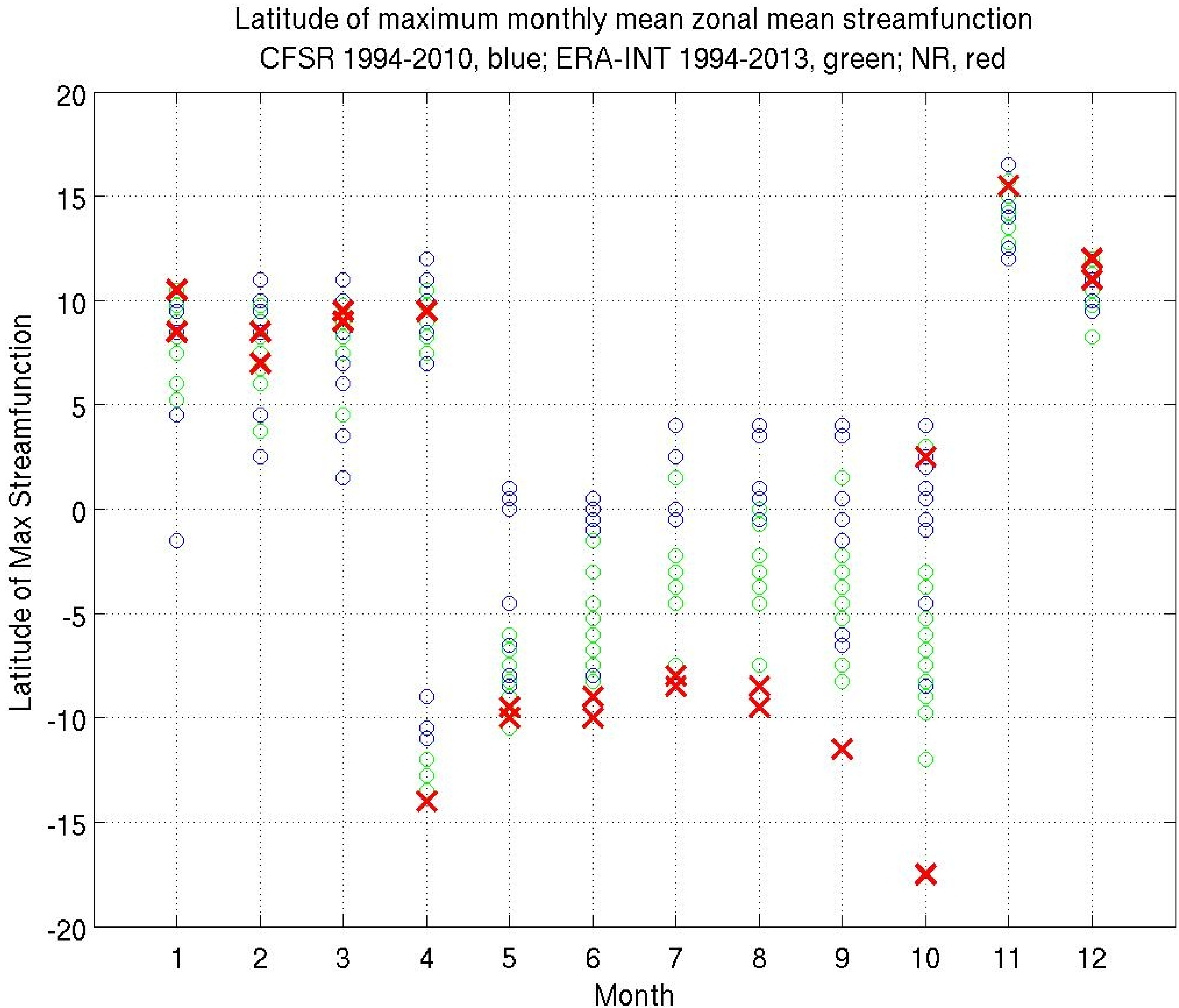
- Nonexistence
- Identical or fraternal twins
- Outdated by the time you get to use them
- Gigantic output files and huge computational resource requirements



u250 Interpolation Error
RMSE
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
NHEX Spatial
SHEX Spatial
Tropics Spatial
NHEX Time
SHEX Time
Tropics Time
3
4.5
6
9
13.5
18
27 km
20
30
40
60
90
120
180 minutes
Spatial/Temporal Resolution

Nature Run Validation

- Evaluate if NR is sufficiently realistic to yield meaningful results
- NR does not have to be “average”, but should fall within the envelope of possible real scenarios. ie, the NR should be indistinguishable from a random period drawn from the real world
- In addition to the phenomena of interest, the NR needs to realistically replicate fields needed to generate synthetic observations


Latitude of maximum monthly mean zonal mean streamfunction
CFSR 1994-2010, blue; ERA-INT 1994-2013, green; NR, red
Latitude of Max Streamfunction
Month
20
15
10
5
0
-5
-10
-15
-20
1
2
3
4
5
6
7
8
9
10
11
12

Nature Run Validation

- Can't validate everything; corollary – don't expect a NR to come pre-validated for your needs
- Validate the NR for the OSSE that **you** want to do

Synthetic Observations

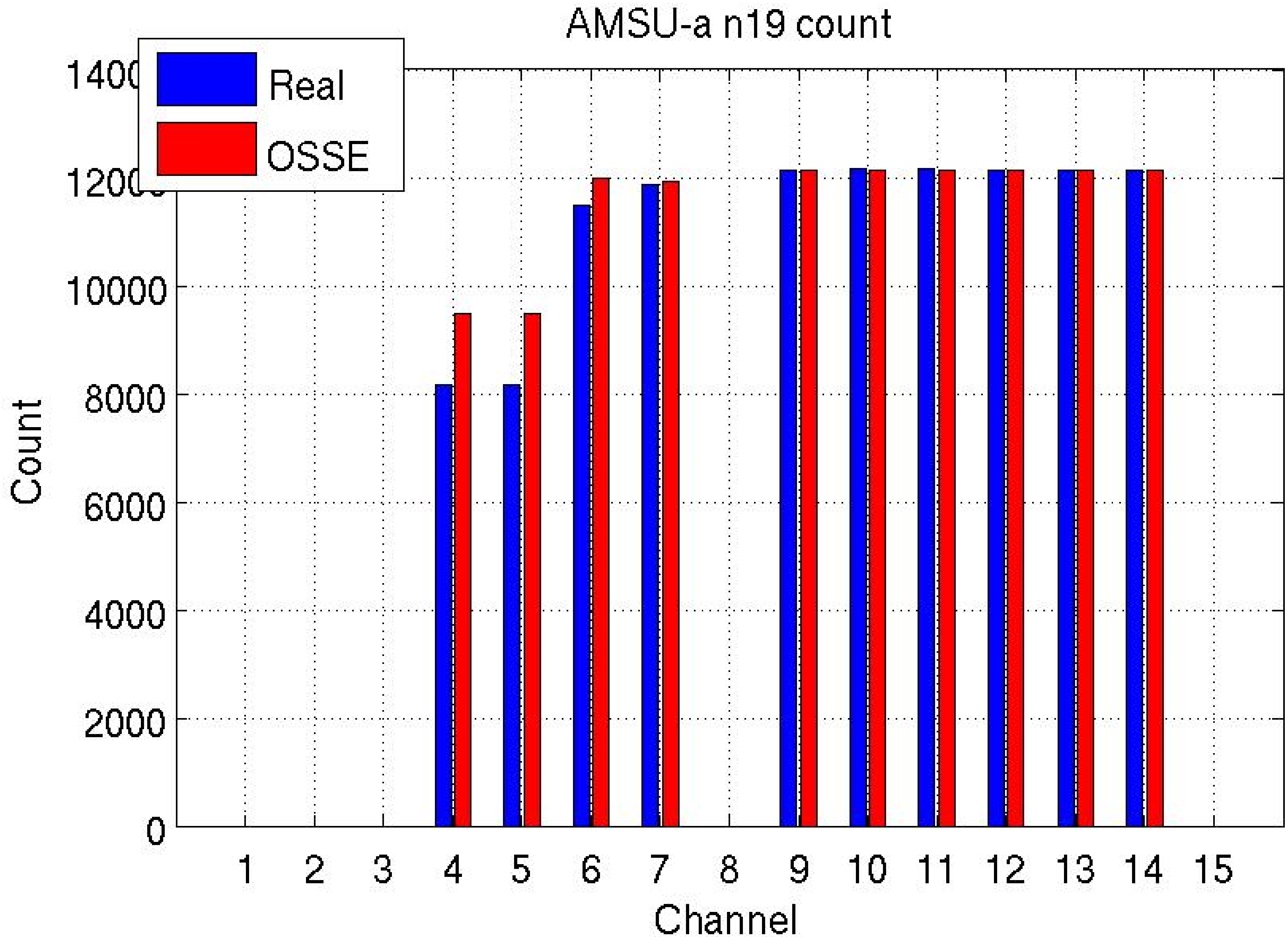
- Synthetic observations are “backed out” from the NR
  - Direct interpolation of NR fields (conventional)
  - Observation operator (radiance, GPS)
  - More art than science (AMVs)
- Observation locations/frequency can be based on archived real data, or simulated

Observation Errors

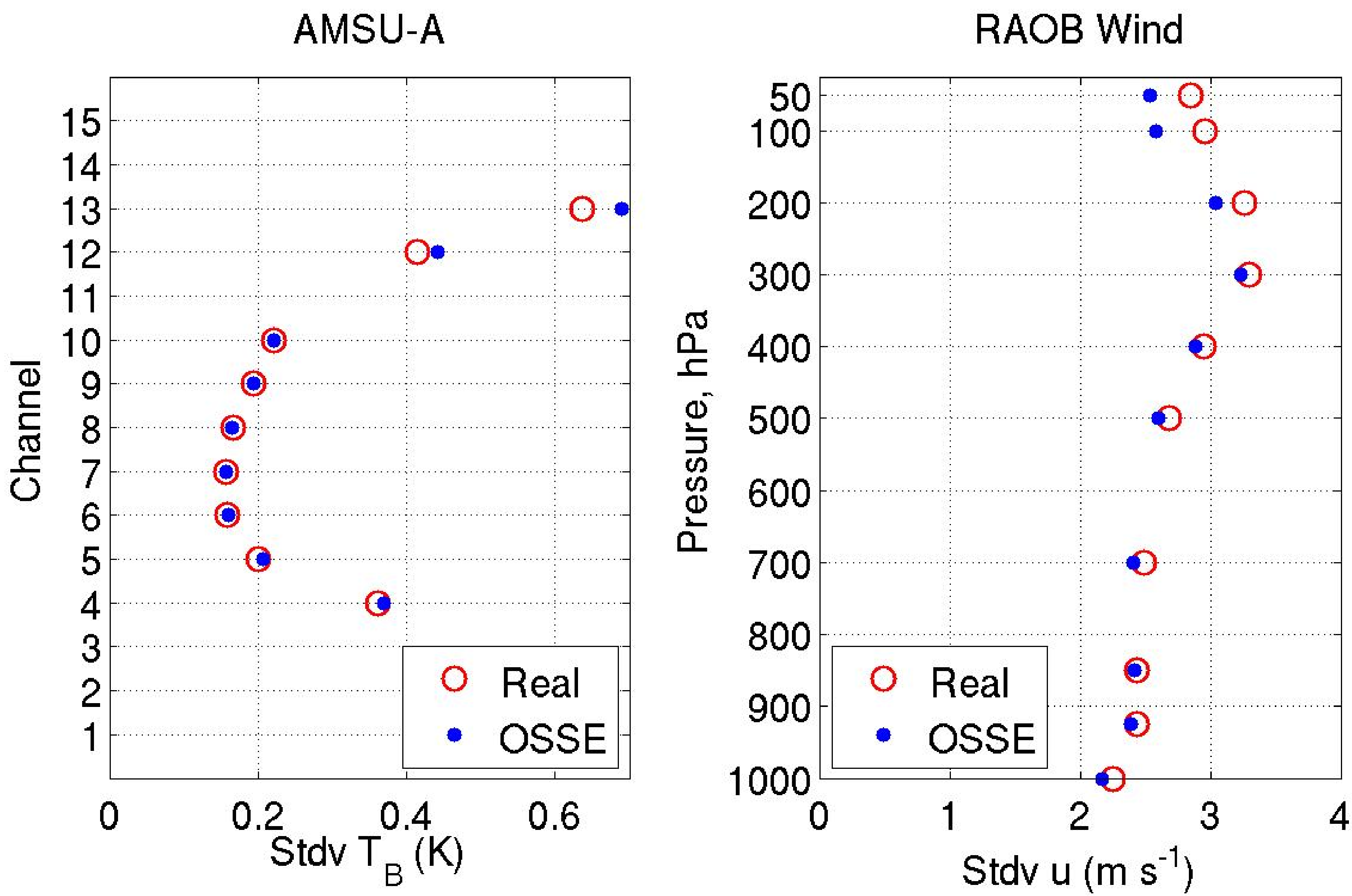
- Synthetic observations contain some intrinsic interpolation/operator errors, but less than real observations (usually)
- Synthetic errors are created and added to the synthetic observations to compensate
- Error is complex and poorly understood
    - Error magnitude
    - Biases
    - Correlated errors

Calibration

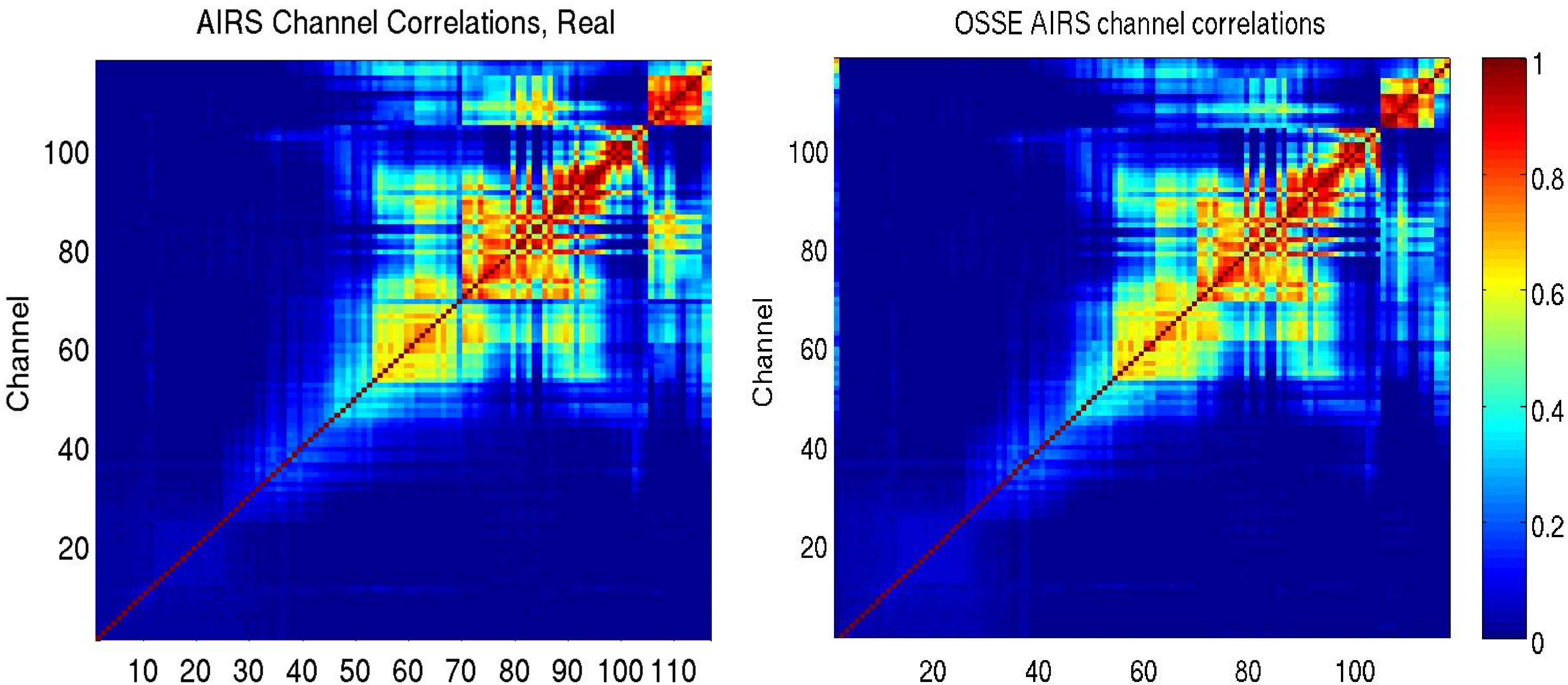
- Adjust synthetic observations and their errors to increase realism of the OSSE in a statistical sense
  - Compare OSSE statistics to statistics using real data in the same DAS/forecast system
- Need to decide what statistical metrics to use for the calibration, depending on your needs
- Calibrating new observation types?
  - Find an analogous data type if possible



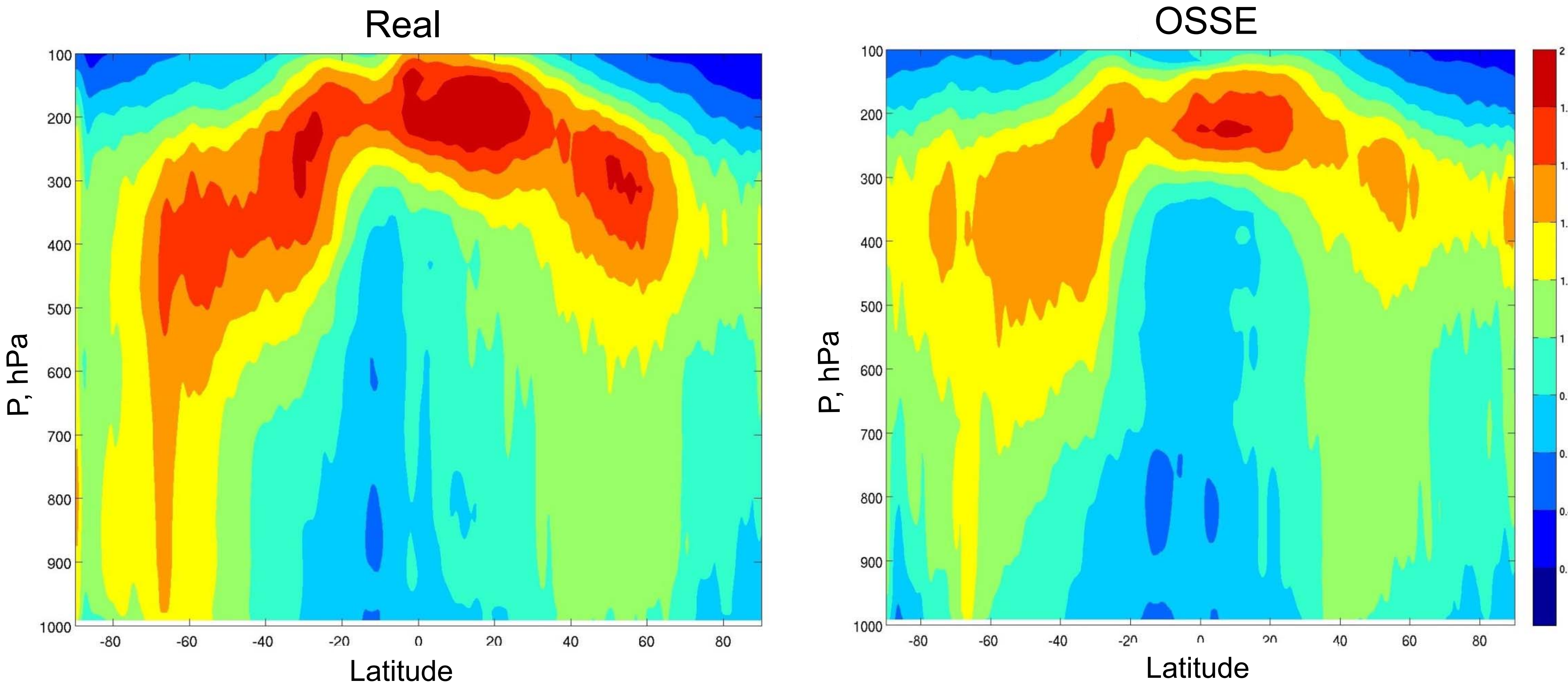
Observation count is easy to calibrate



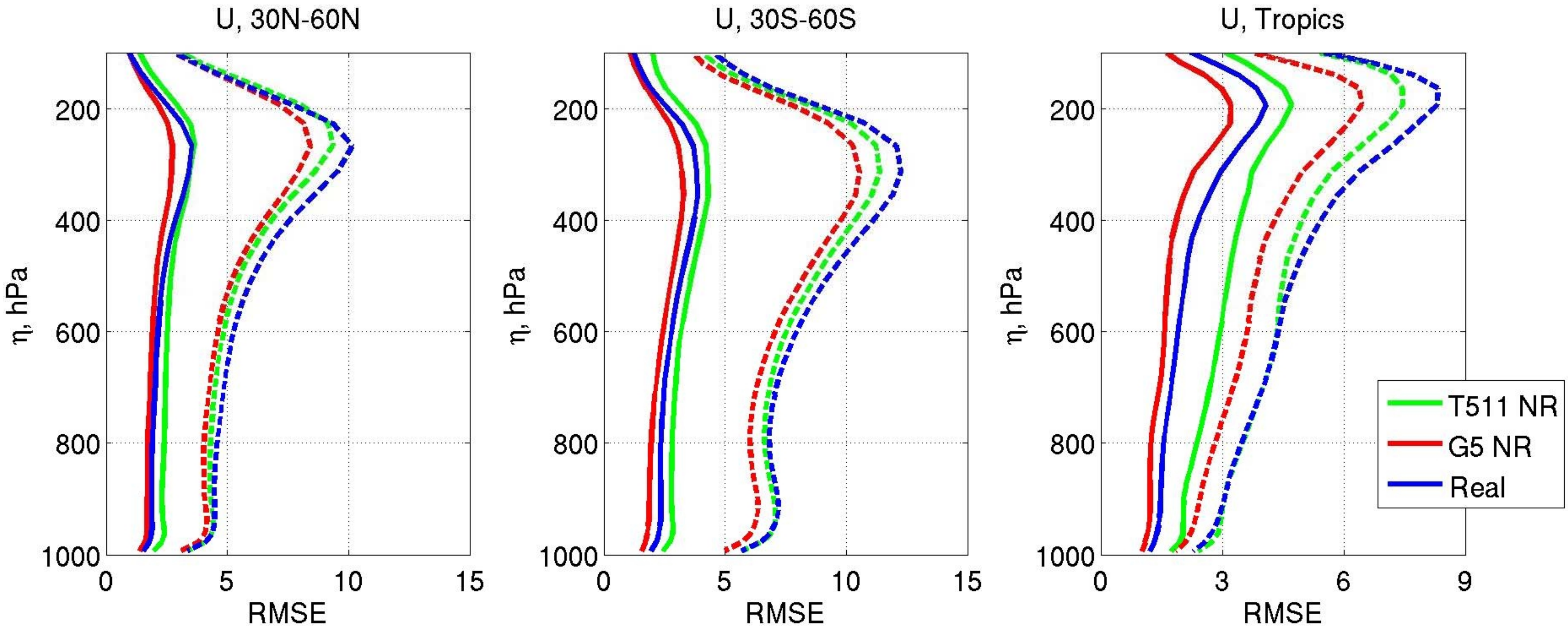
O-F is fairly easy to calibrate because you can manipulate O directly.



Some observation correlations are relatively easy to calibrate

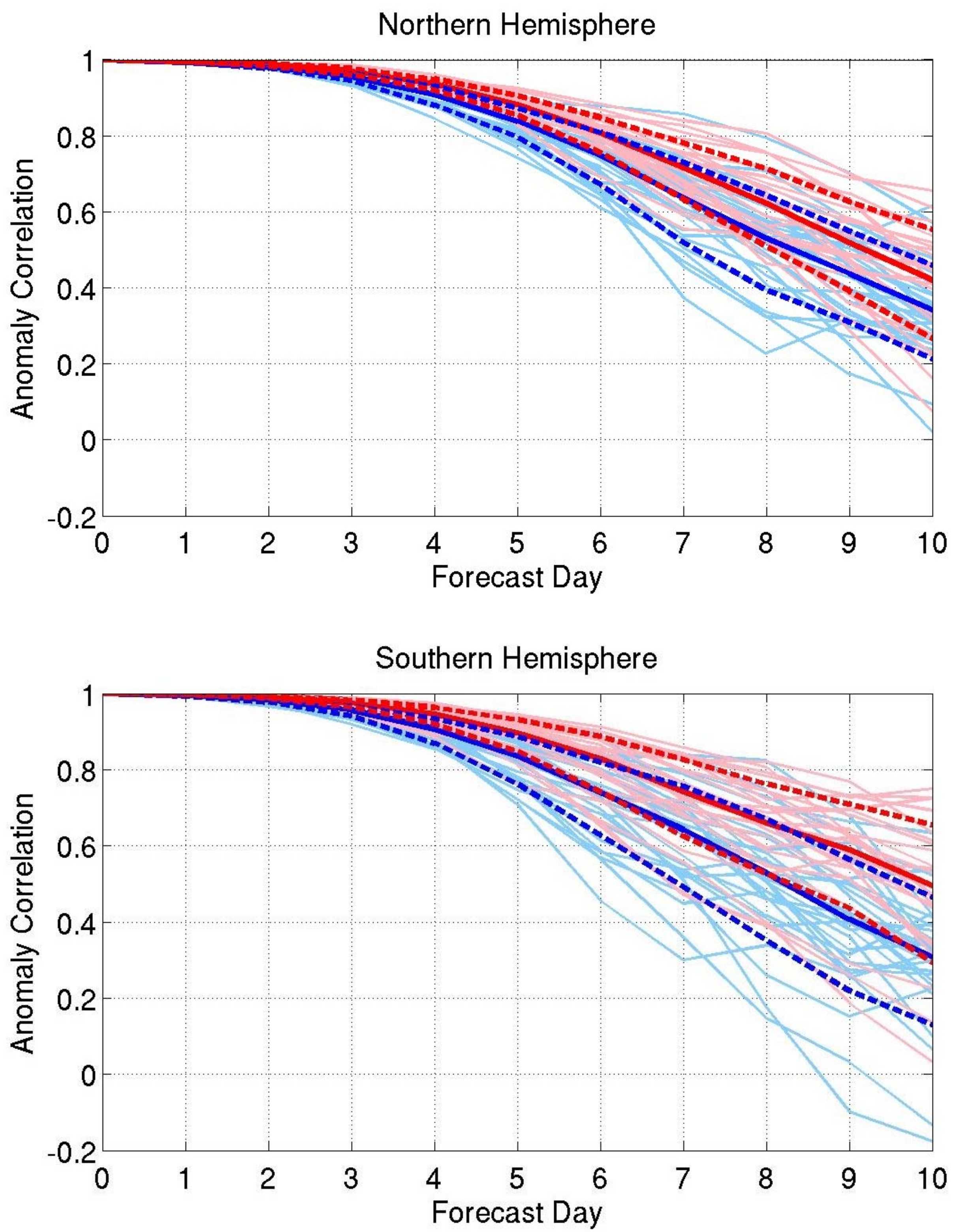


A-B (analysis increment) is a little harder to calibrate, as A and B are not directly controlled



Forecast errors are harder to calibrate, especially for longer forecasts. Matching of this statistic by manipulation of observations is difficult to impossible beyond ~24 hour forecasts.

Model error determines forecast skill in the longer term forecast, so calibration is not possible (unless you want to mess with your model).



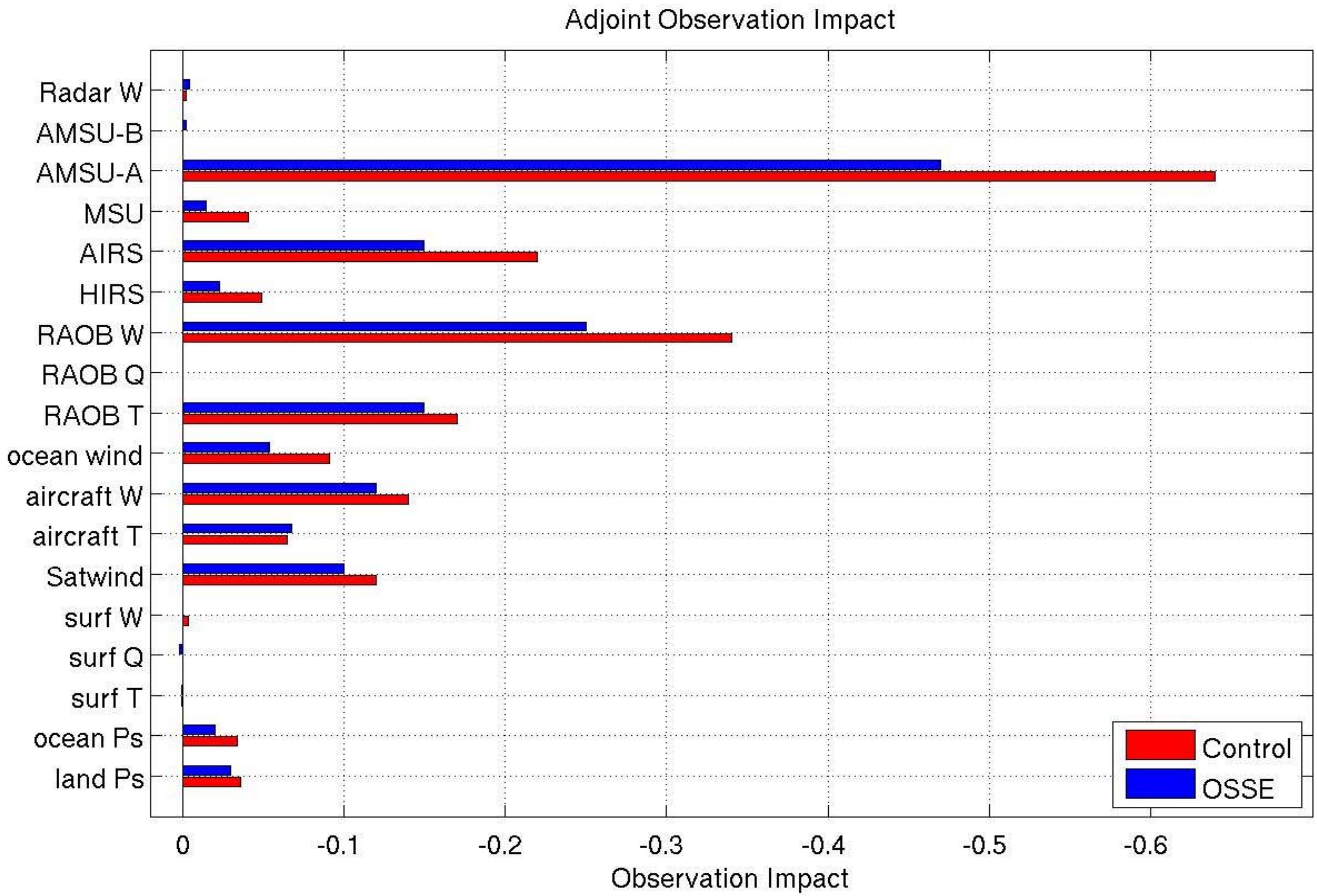
Red: OSSE
Blue: Real

500 hPa anomaly correlations of geopotential height

Criticisms of OSSEs

- Results only apply within the OSSE system – no concrete connection to the real world
- Even the best OSSEs are far from perfect: incestuousness, difficulty in generating observations and errors, deficiencies of the Nature Run
- By the time the new instrument is deployed, both the global observing network and the forecast models/DAS will be different
- Examples of sloppy or unsuccessful OSSEs

Why believe OSSE results?



New observations can be put into context relative to existing observation impacts

When not to run an OSSE

- When you can't model the phenomena you are interested in
- When you can't simulate your new observations
- When you can't assimilate your new observations

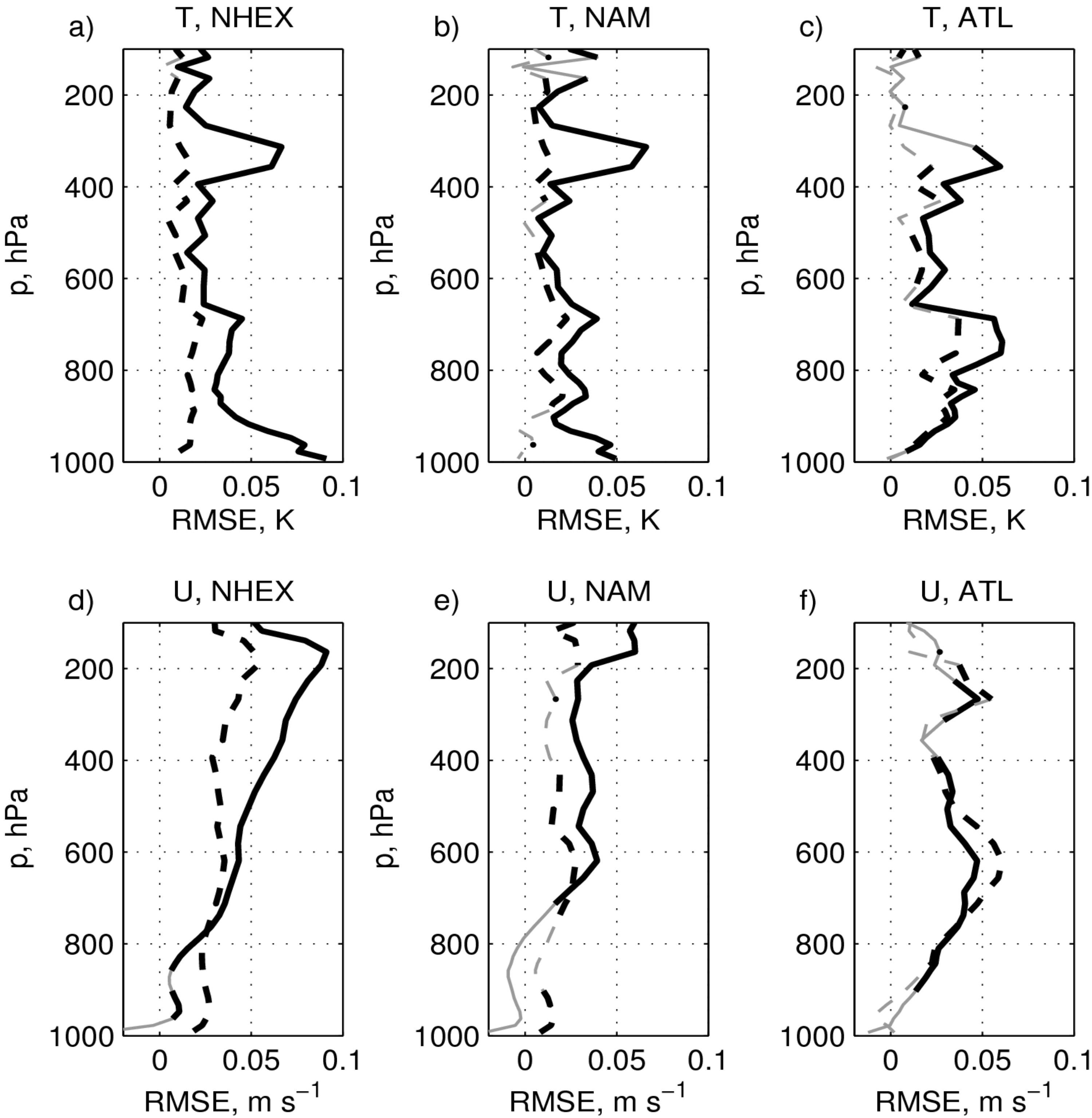
Common Pitfalls

- Very reduced baseline of assimilated observational data (ex. no radiance data)
- Other artificial degradation of analysis state
- No validation or calibration of OSSE framework
- Obtaining robust results from case studies is very challenging

Choosing Metrics

- Long cycling periods necessary to get statistically significant results for most new observations
- Anomaly correlation is a difficult metric to show appreciable impacts
- What fields do you expect the instrument to improve?
- Largest impacts found at analysis time or short-term forecasts

OSSE with added RAOBS at 06z and 18z



Two months of DAS cycling

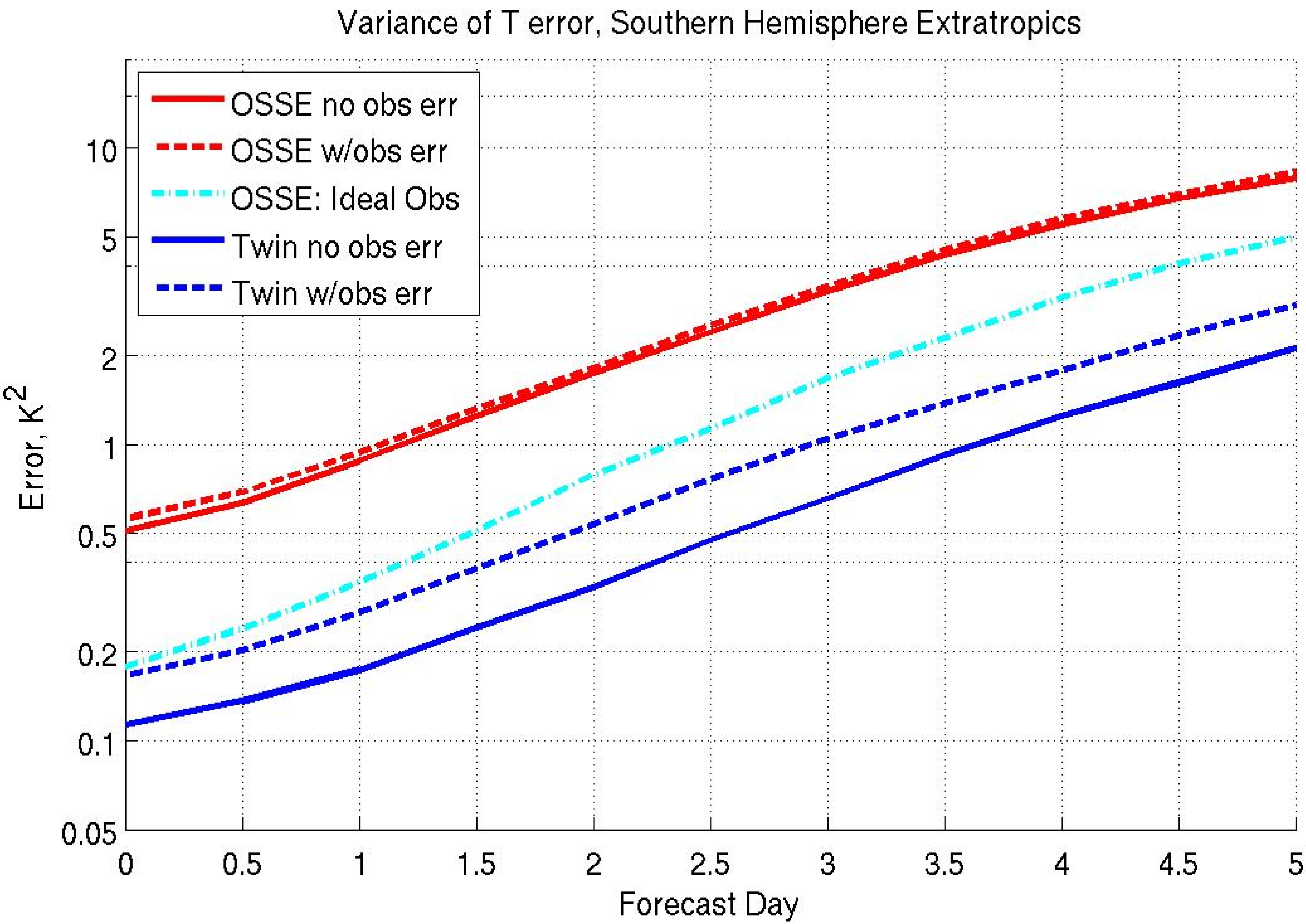
Regional RMSE

Metric of state variable field and not case studies (even though TCs are of interest in this case)

Idealized Studies

- Identical twin experiments
- Idealized observations
- Manipulation of observation errors
- Experiments with **B**, **R**

- Make use of available “Truth”


Variance of T error, Southern Hemisphere Extratropics
OSSE no obs err
OSSE w/obs err
OSSE: Ideal Obs
Twin no obs err
Twin w/obs err
Error, K^2
Forecast Day
10
5
2
1
0.5
0.2
0.1
0.05
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

Takeaway

- OSSEs can provide useful information about new observational types and the workings of data assimilation systems
- Careful consideration of research goals should guide each step of the OSSE process
- OSSEs are hard, good OSSEs are harder