Introduction to OSSEs

Nikki Privé

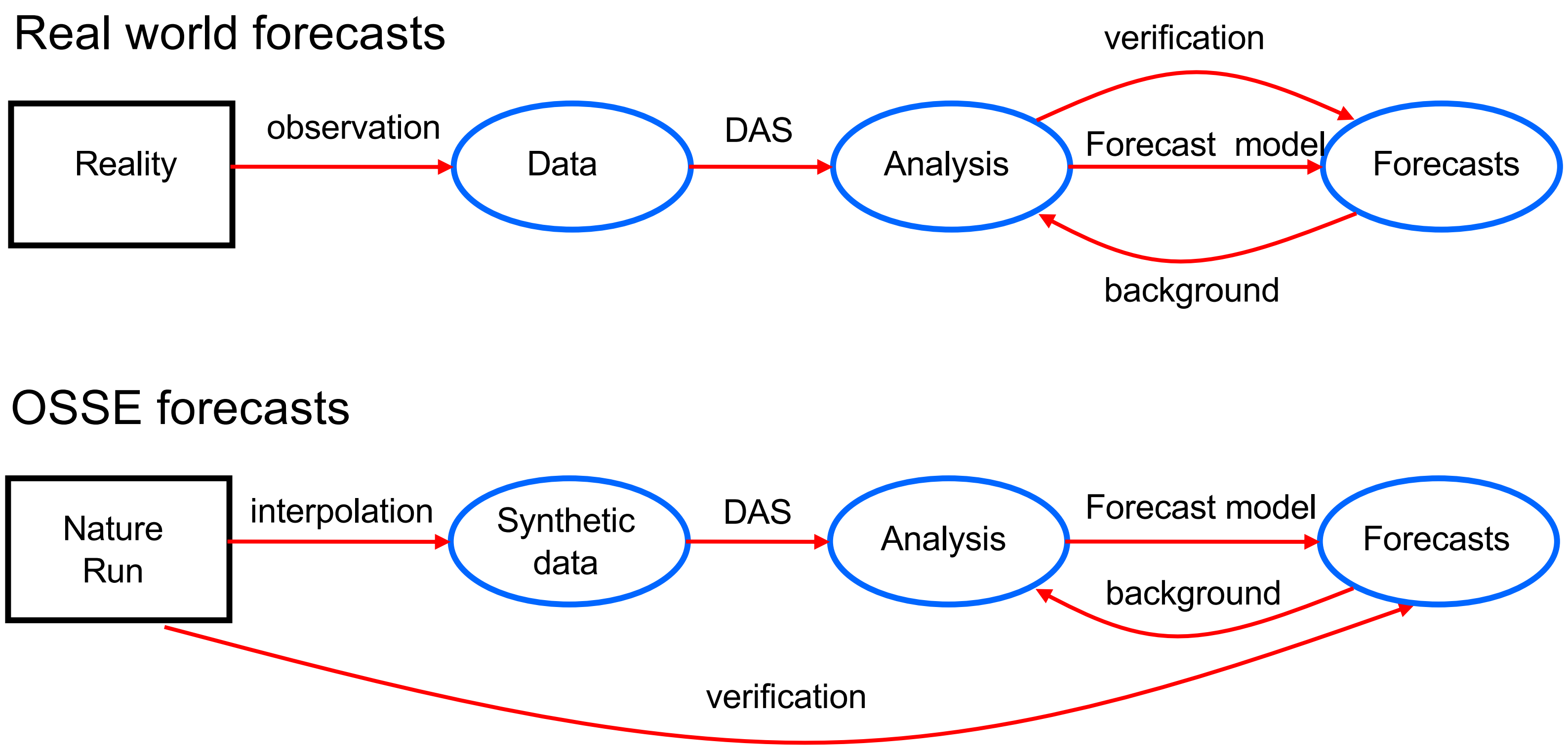
31 October 2022

What is an Observing System Simulation Experiment?

An OSSE is a modeling experiment used to estimate the impact of new observing systems on numerical weather prediction 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
- Realistic errors are added to the synthetic observations
- 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



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

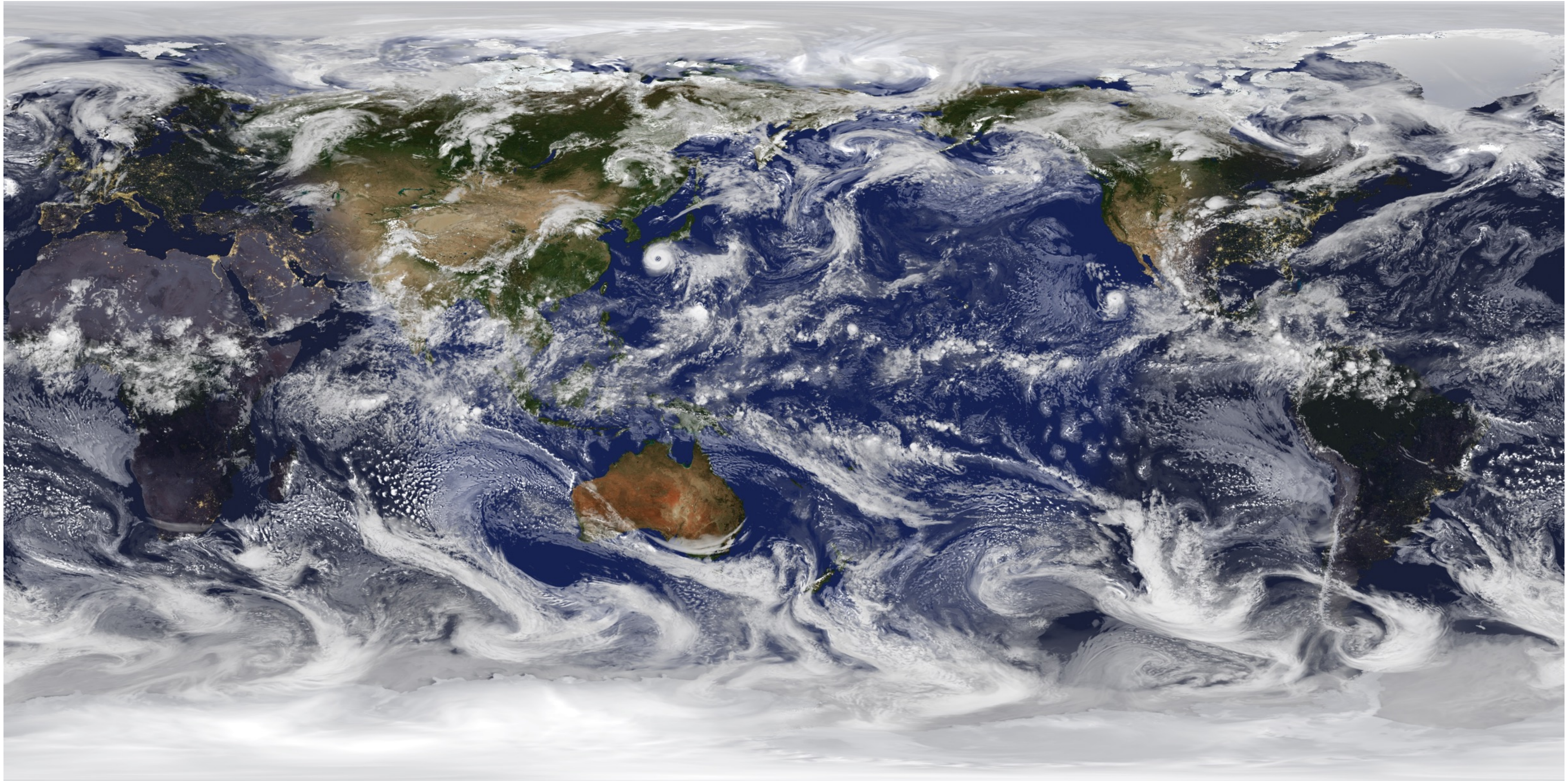
When NOT to run an OSSE

- When you can’t model the phenomenon you are interested in
- When you can’t simulate your new observations
- When you can’t assimilate your new observations
  - (Unless you want to work on development of a new operator)

Nature Runs

- Nature Runs act as the ‘truth’ in the OSSE, replacing the real atmosphere
- Usually, a long free (no DAS cycling) forecast from the best model available is used as the Nature Run
  - ✓ 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
- Must be able to realistically model phenomena of interest
- Ideally is “better” than and different from the operational model used for experiments to replicate model error

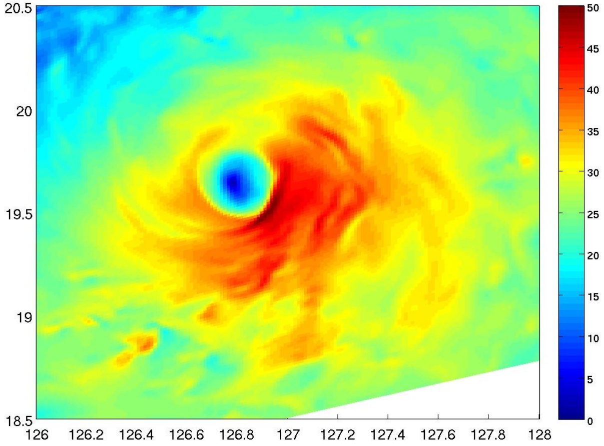
GEOS-5 Nature Run (G5NR)



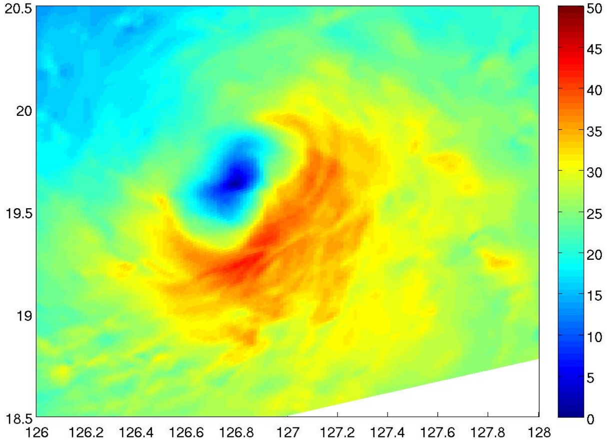
Common Problems with Nature Runs

- Identical or fraternal twin – insufficient model error
- Outdated by the time you get to use them
- Gigantic files and huge computational resource requirements
  - Massive I/O demands
  - Infrequent output with high spatial resolution = trouble

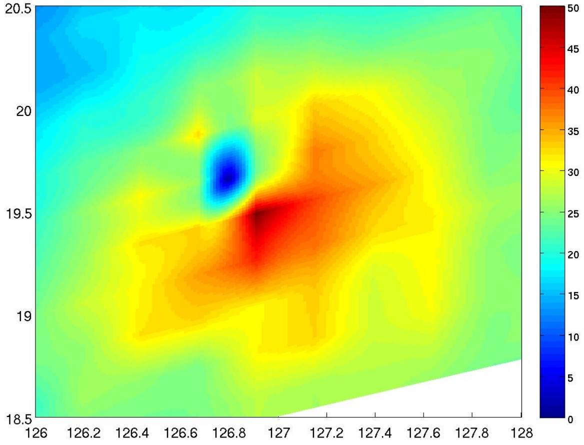
Full resolution (1.5 km, 10 min)



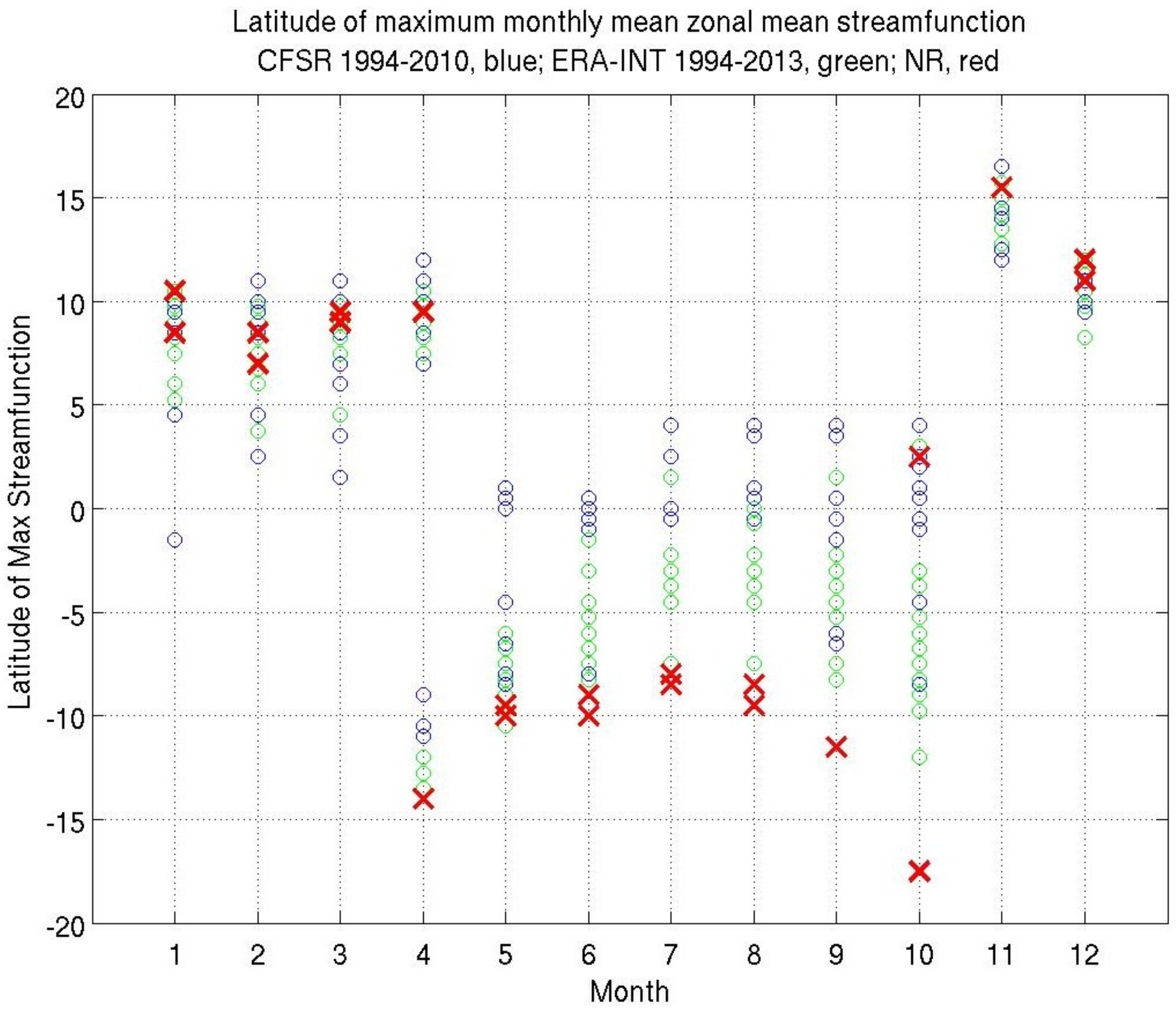
Temporal interpolation (3 hrs)



Spatial interpolation to 27 km



Comparison of temporal and spatial Interpolation errors for 1.5 km model run for Typhoon Guchol (2012).



Nature Run Validation

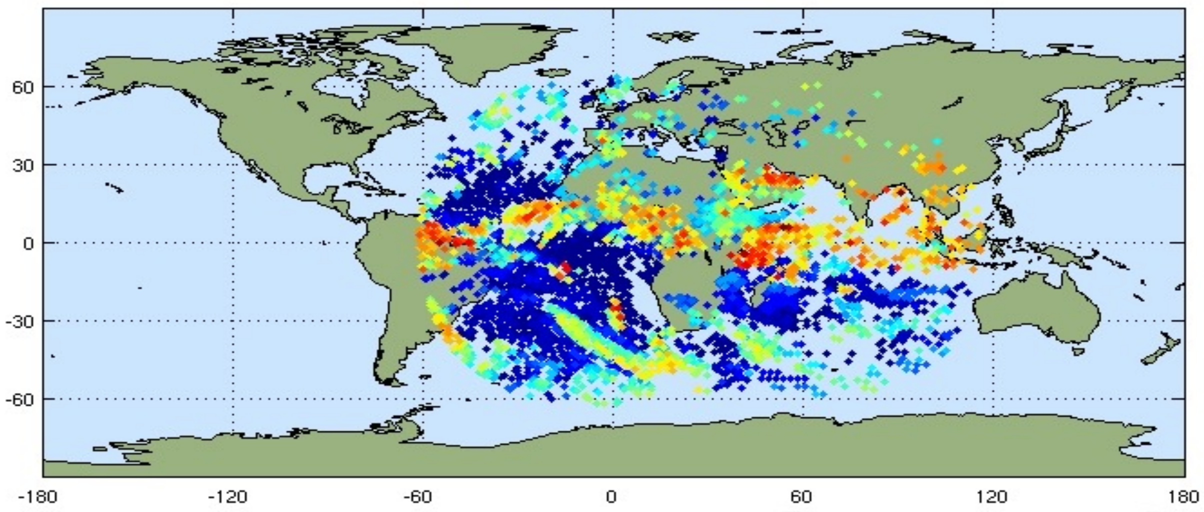
- Evaluate if NR is sufficiently realistic to yield meaningful results
- NR needs to replicate fields needed to generate realistic simulated observations
- Can't validate everything – corollary – don't expect a NR to come pre-validated for your needs

Simulated Observations

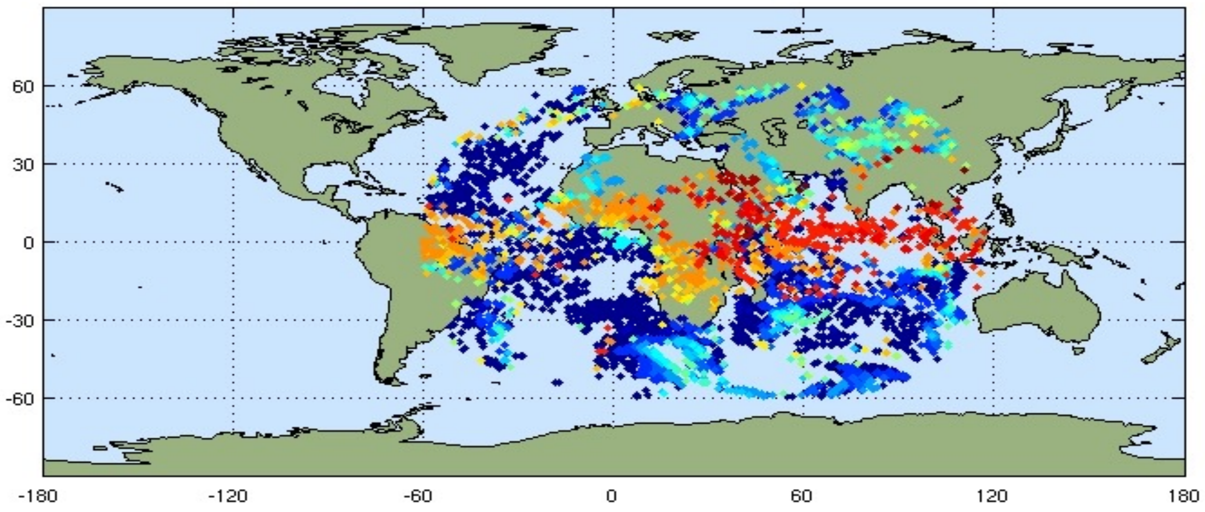
Goal: produce simulated observations that are statistically indistinguishable from real data as perceived by the DAS

- Spatiotemporal distribution
- Observation Count
- Co-location with features
- Error characteristics

Real



Simulated



Example of METEOSAT AMV observations at 0000 UTC 10 July

Observation Errors

- Simulated observations contain some intrinsic interpolation/operator errors, but less than real observations (usually)
  - Representativeness Error
  - Operator Error
- Simulated observation errors are created and added to the simulated observations to compensate
- Observation error is complex and poorly understood
  - Error Magnitude
  - Biases
  - Correlated Errors

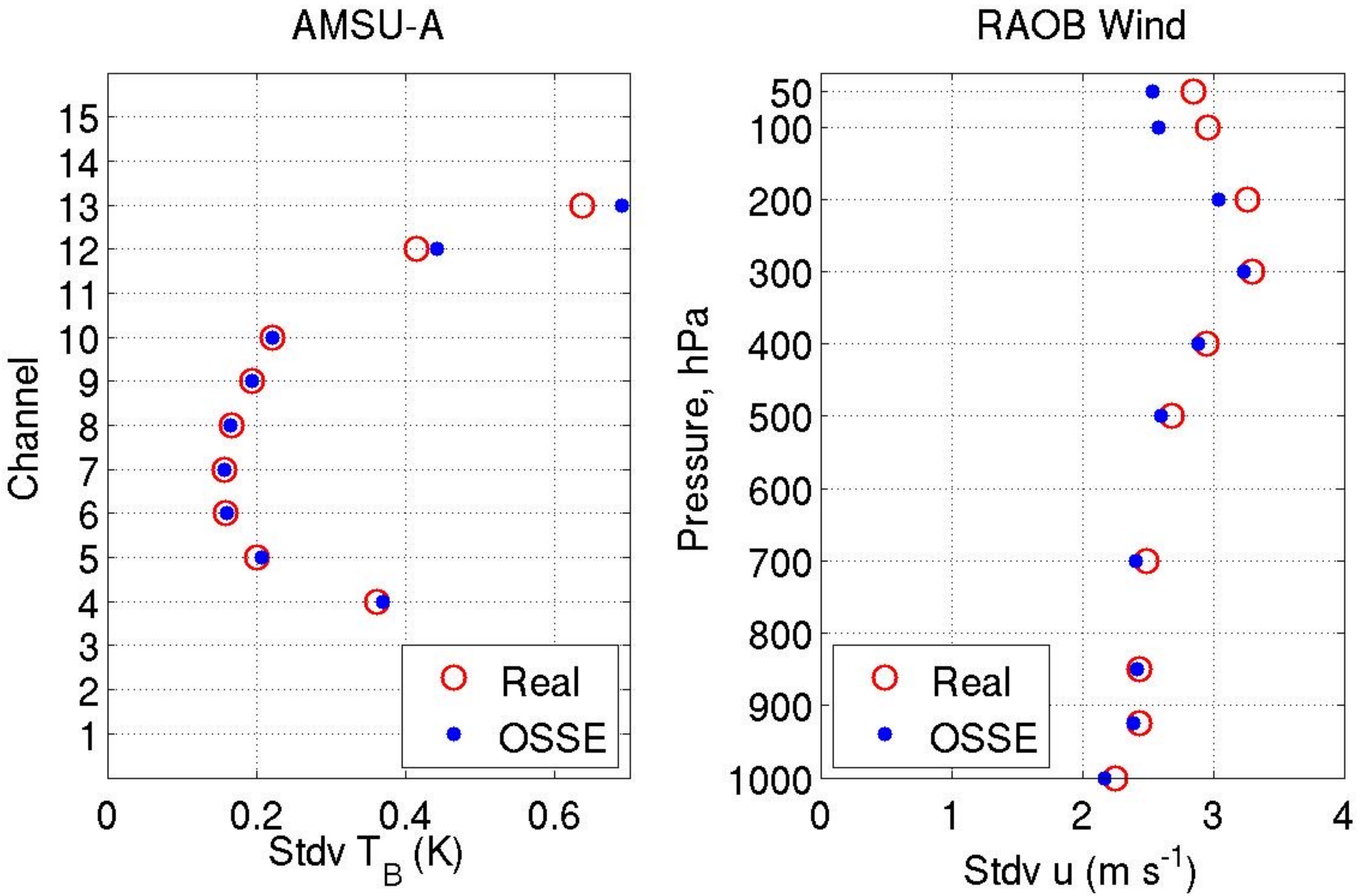
Calibration and Validation

- Goal – match statistics of the DAS behavior in the OSSE to that of the real world observations/DAS
- Choose some metric(s) to use as a basis for calibration
- Run analog case with real observations
- Compare the OSSE metrics to real data metrics
- Adjust the OSSE framework to match the metrics
- Compare other important metrics (not used for calibration) to validate your OSSE
- Describe your calibration and validation in any publications

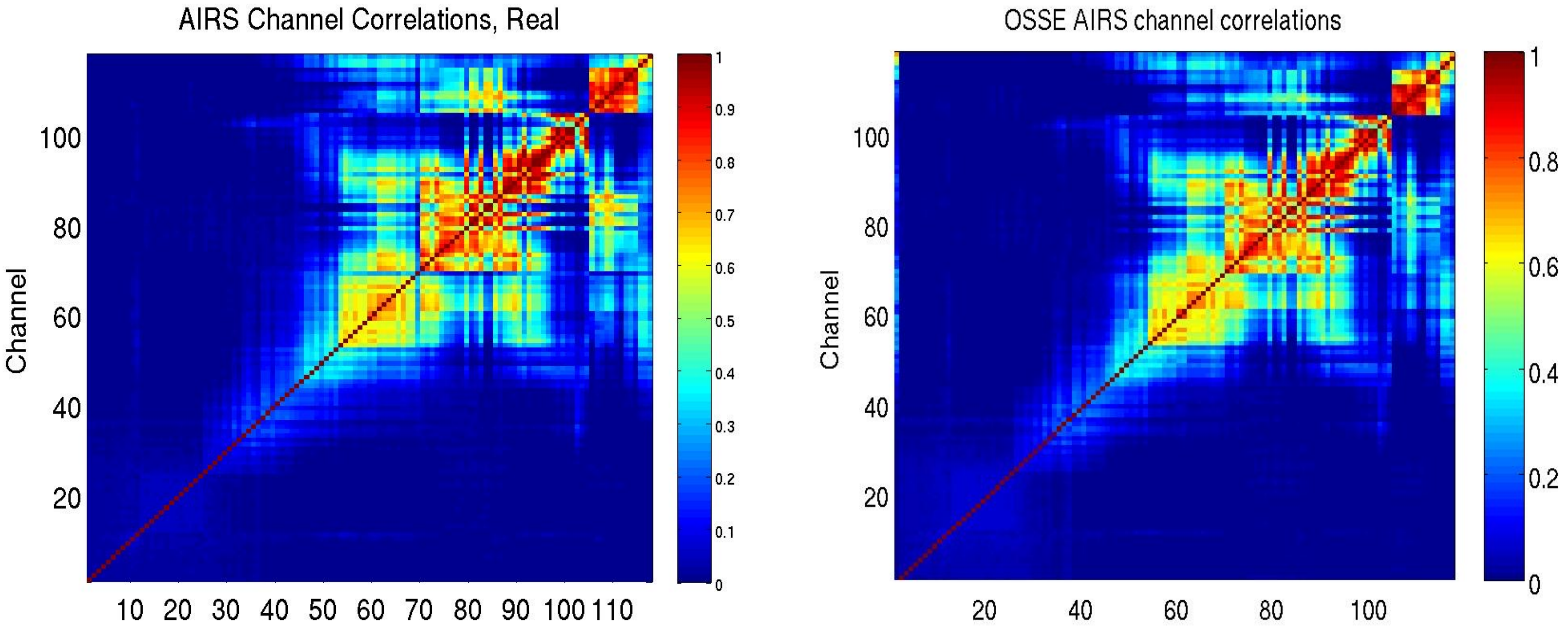
Calibration: Observation Innovation

Observation innovation (O-F) is fairly easy to calibrate because you can manipulate O directly

Be careful not to overcompensate by adding too much observation error when the background error is too small



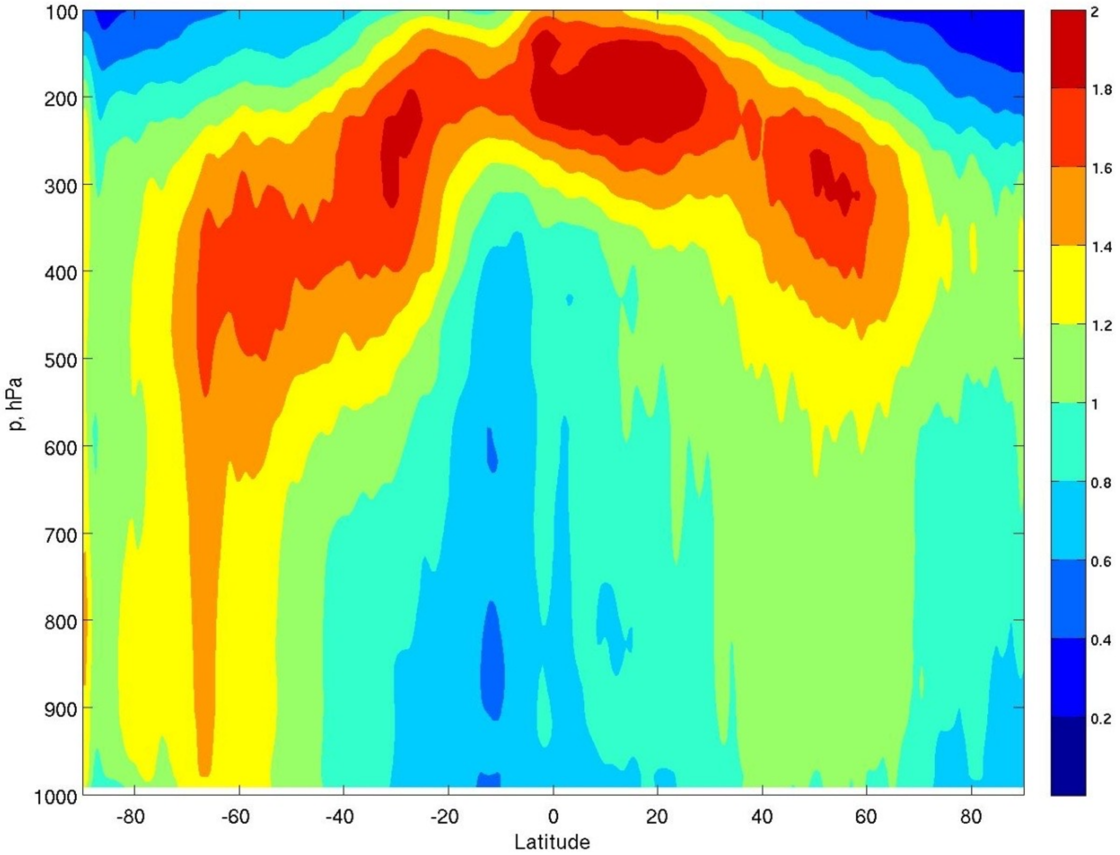
Calibration: Correlations



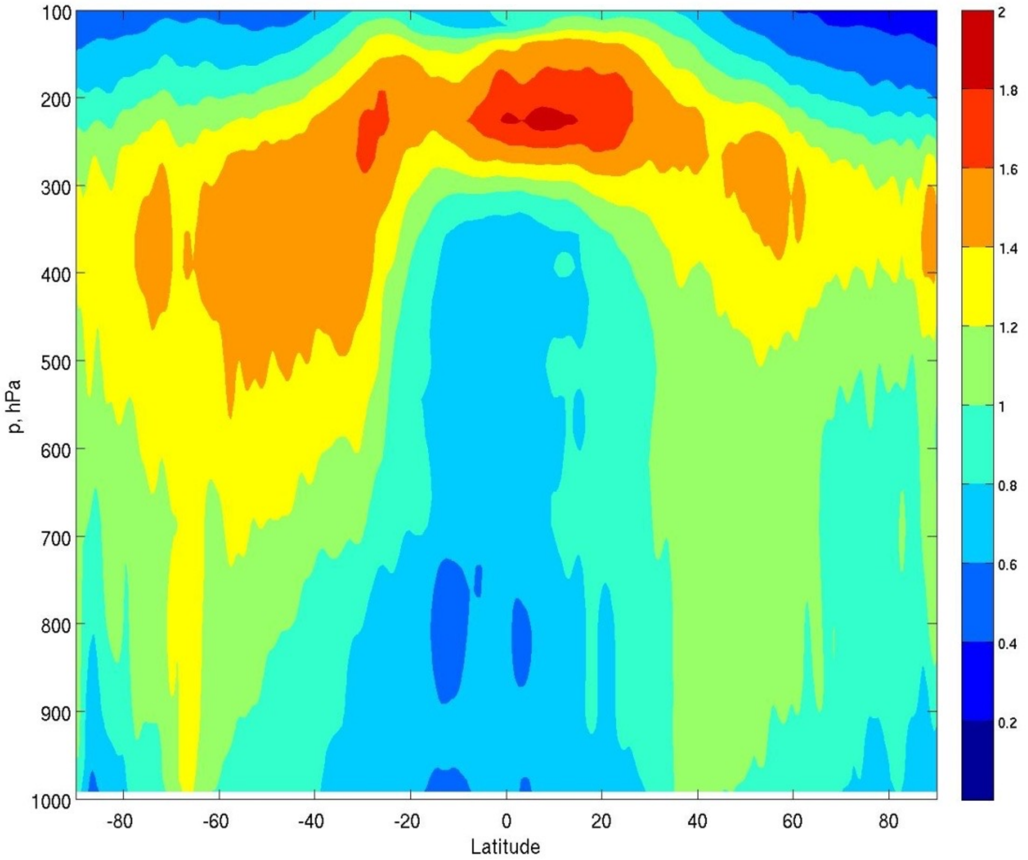
Some types of observation correlations are relatively easy to calibrate

Calibration: Analysis Increment

Real



OSSE



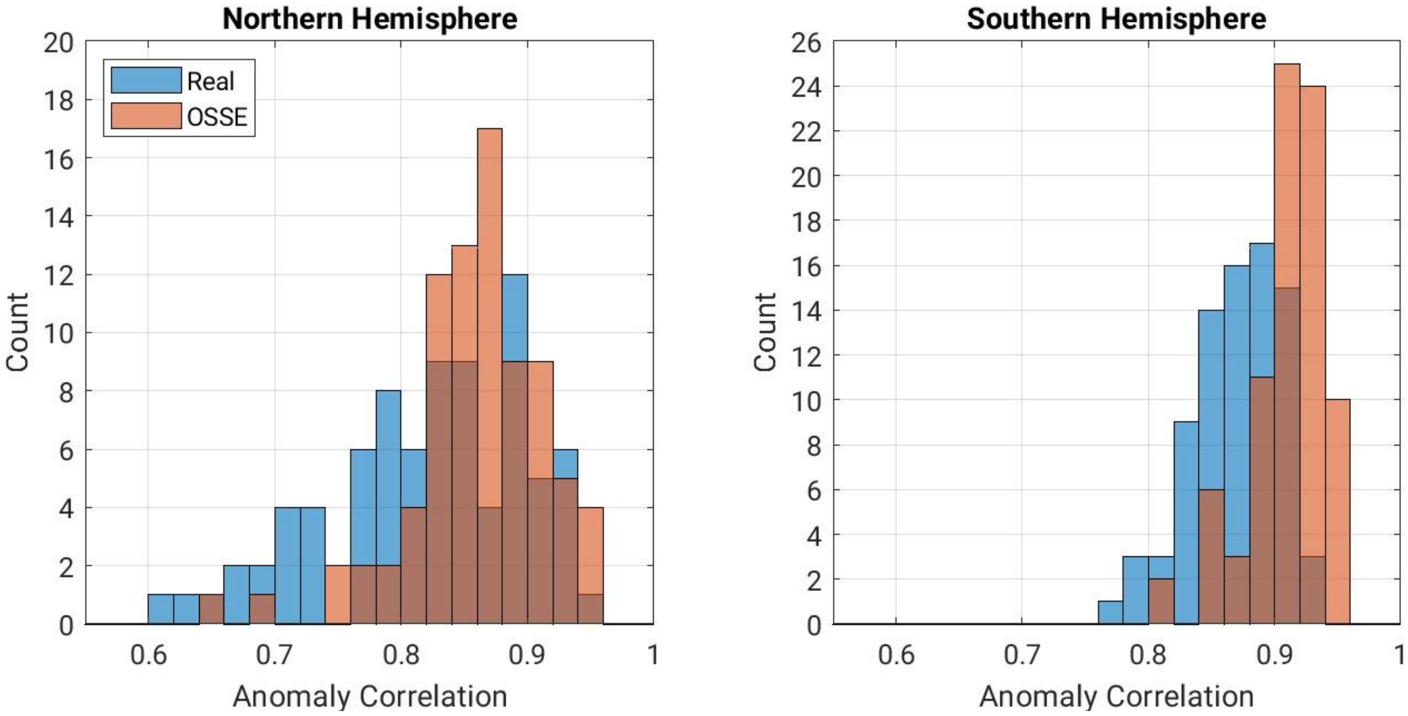
StDev(A-B) for Temperature, K

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

Calibration: Forecast

Model error strongly influences forecast skill in the medium/long term, so calibration is not possible (unless you want to mess up your model)

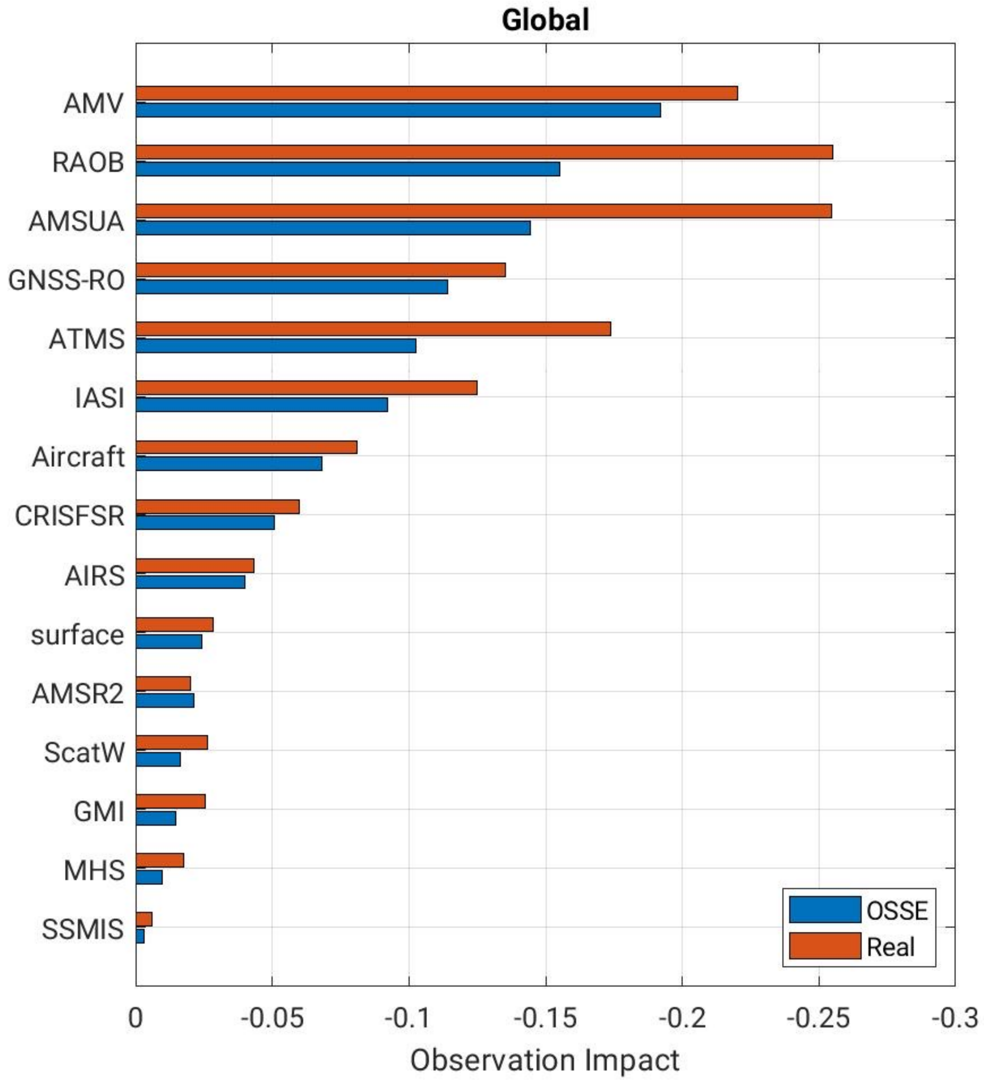
500 hPa anomaly correlation of geopotential height



Calibration: Forecast Sensitivity to Observation Impact

Net impacts are lower in the OSSE due to insufficient model error

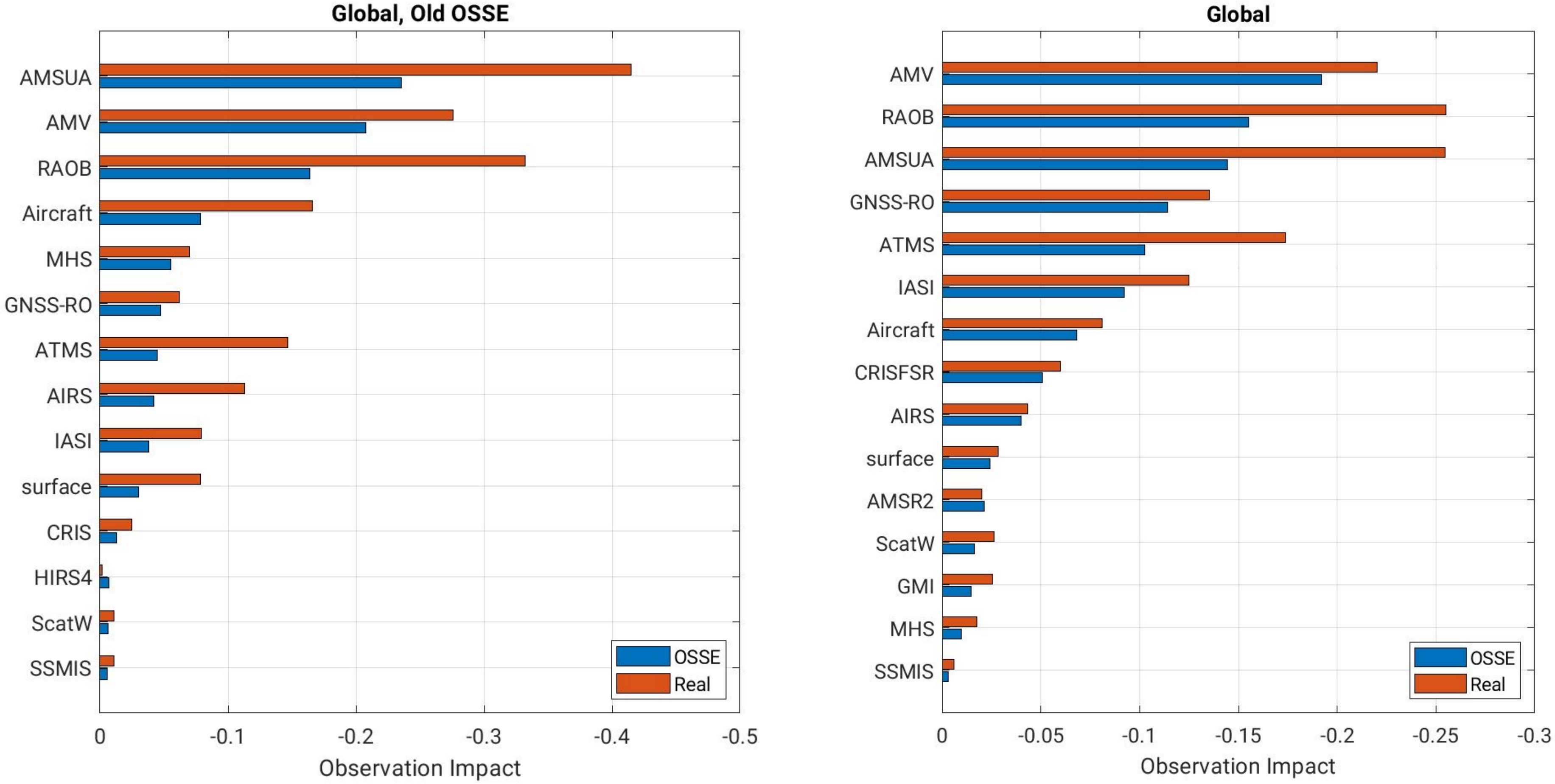
But new observations can be put into context relative to existing observation impacts



'Twinning' Affects Everything

Lack of model error between the NR and forecast model affects every aspect of the OSSE, from calibration to observation impacts.

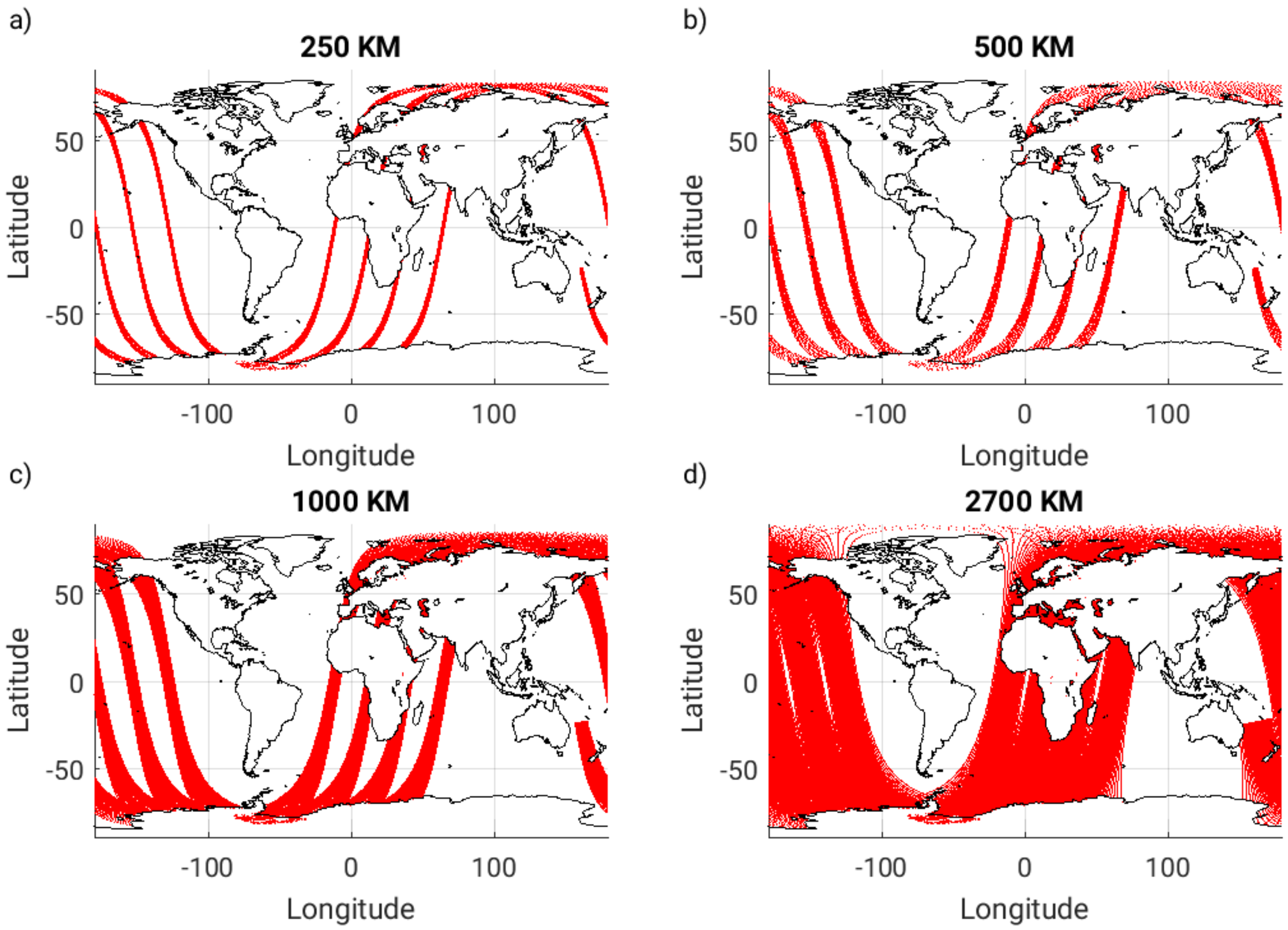
Updated OSSE with more differences between the NR and forecast model shows better agreement with Real in FSOI.



Designing an OSSE

Clearly determine the question you want to answer with your experiments

- Proof of concept?
- Realistic impact study?
- Idealized observation network?
- DAS development?



Designing an OSSE

Consider what to use as your Control observation network

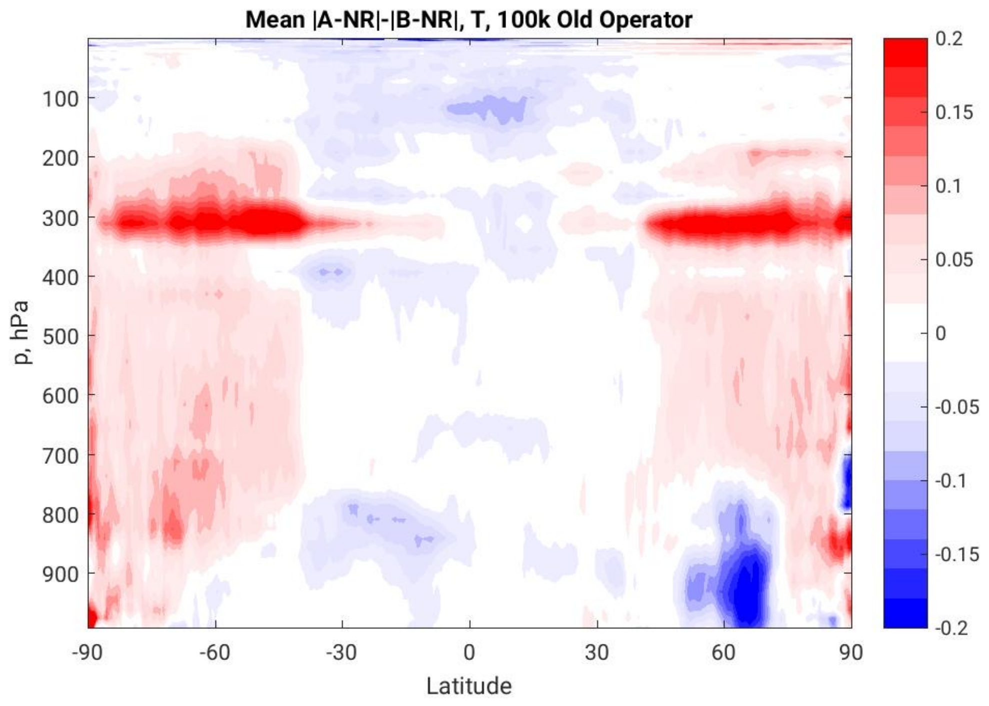
- An idealized Control to illuminate your observation impact?
- A realistic “current” Control to replicate the existing global network?
- A potential “future” Control to emulate expected changes to the global network?

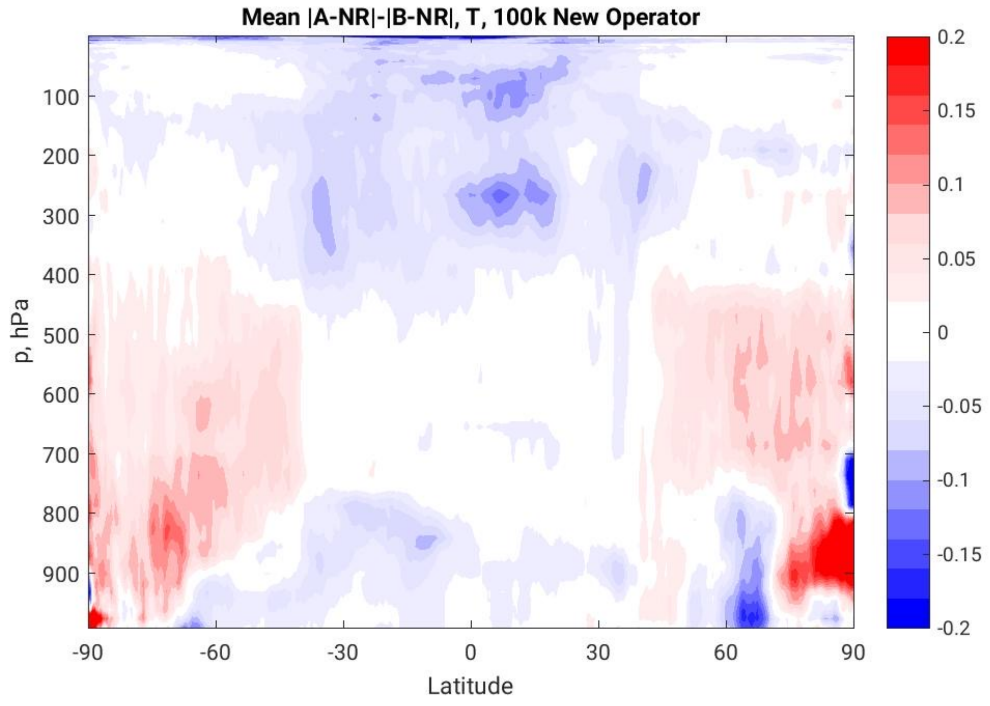
The choice of Control will in part dictate what will be learned in your experiments.

Designing an OSSE: Metrics

Remember that you have the “truth” available!

- The analysis and background errors can be precisely calculated
- You are not confined to traditional metrics used with real data – the OSSE framework gives you more (better!) options





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
- OSSE DAS and model always lag behind operations
- 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

Takeaway

- OSSEs can provide useful information about new observational types
- OSSEs can be used in the development and testing of data assimilation systems
- Careful consideration of research goals should guide each step of the OSSE process
- Validate your OSSE!
- OSSEs are hard, good OSSEs are harder